The 42nd International Conference on

MATHEMATICAL METHODS IN ECONOMICS

MME 2024

ČSOV

UNIVERZITA J. E. PURKYNĚ V ÚSTÍ NAD LABEM

UNIVERZITA J. E. PURKYNĚ V ÚSTÍ NAD LABEM
Přírodovědecká fakulta

PROCEEDINGS OF THE 42ND INTERNATIONAL CONFERENCE ON **MATHEMATICAL METHODS IN ECONOMICS**

Proceedings of

The 42nd International Conference on Mathematical Methods in Economics (MME 2024)

September 11–13, 2024
Jan Evangelista Purkyně University
Ústí nad Labem, Czech Republic

The 42nd International Conference on Mathematical Methods in Economics (MME 2024).
September 11–13, 2024.
Hosted by the Department of Computer Science, Faculty of Science, Jan Evangelista Purkyně University in Ústí nad Labem.
Pasteurova 3632/15, 400 96 Ústí nad Labem
https://mme2024.ujep.cz/

Mathematical Methods in Economics (MME 2024)
Proceedings of the 42nd International Conference on Mathematical Methods in Economics
Publisher: The Czech Society for Operations Research
Winston Churchill Square 1938/4, 130 67 Prague 3, Czechia
www.csov.eu
Editors: Hossein Moosaei, Josef Jablonský, Helena Brožová
This publication is not subject to a language check.
All papers have passed a blind peer review process.
All papers adhere to the ethical guidelines of the publisher.
ISBN ??????

Contents

Conference Organization

General Chair:

- Hossein Moosaei, Jan Evangelista Purkyně University, Czech Republic

Organizing Committee:

- Hossein Moosaei (chair), Jan Evangelista Purkyně University, Czech Republic
- Jiří Škvor (chair), Jan Evangelista Purkyně University, Czech Republic
- Regina Herma, Jan Evangelista Purkyně University, Czech Republic
- Pavel Beránek, Jan Evangelista Purkyně University, Czech Republic
- Petr Kubera, Jan Evangelista Purkyně University, Czech Republic
- Michalea Tichá, Jan Evangelista Purkyně University, Czech Republic

Program Committee Chairs:

- Hossein Moosaei, Jan Evangelista Purkyně University, Czech Republic
- Josef Jablonský, Prague University of Economics and Business, Czech Republic
- Helena Brožová, President of the Czech Society for Operations Research Czech University of Life Science Prague, Czech Republic
- Miloš Kopa, Vice-president of the Czech Society for Operations Research Charles University, Czech Republic

Program Committee:

- Paula Amaral, Universidade Nova de Lisboa, Portugal
- Adam P. Balcerzak, Universityof Warmia and Mazury in Olsztyn, Poland
- Roman Belavkin, Middlesex University, UK
- Adam Borovička, Prague University of Economics and Business, Czech Republic
- Martin Dlouhý, Prague University of Economics and Business, Czech Republic
- Neng Fan, University of Arizona, USA
- Paola Festa, University of Napoli Federico II, Italy
- Petr Fiala, Prague University of Economics and Business, Czech Republic
- Jan Fábry, ŠKODA AUTO University, Czech Republic
- Jana Hančlová, Technical University of Ostrava, Czech Republic
- Milan Hladik, Charles University, Czech Republic
- Vladimír Holý, Prague University of Economics and Business, Czech Republic
- František Koblasa, Technical University of Liberec, Czech Republic
- Jan Kronqvist, Royal Institute of Technology, Sweden
- Martina Kuncová, Prague University of Economics and Business, Czech Republic
- Elisabeth Köbis, Norwegian University of Science and Technology, Norway
- Jerry Chun-Wei Lin, Western Norway University of Applied Sciences, Norway
- Juan Enrique Martinez-Legaz, Autonomous University of Barcelona, Spain
- Antonio Mucherino, University of Rennes 1, France
- Miroslav Plevný, University of West Bohemia, Czech Republic
- Steffen Rebennack, Karlsruhe Institute of Technology, Germany

Foreword

It is with great pleasure that we present the Proceedings of the 42nd International Conference on Mathematical Methods in Economics (MME 2024), held in Ústí nad Labem from September 11 to 13, 2024. This collection showcases a selection of the finest contributions from researchers and practitioners who participated in this year's conference, reflecting the ongoing advancement and innovation in the field of mathematical economics.

MME 2024 continues the tradition of bringing together a diverse community of scholars, professionals, and students, all driven by a shared interest in the application of mathematical methods to economic theory and practice. The conference provided a vibrant forum for the exchange of ideas, discussion of new research trends, and the exploration of complex economic problems through rigorous mathematical modeling.

This year's conference saw the participation of many researchers from different countries, who contributed to a wide array of topics, from econometric theory to optimization techniques in economics. The proceedings include papers that not only push the boundaries of theoretical research but also offer practical solutions to current economic challenges.

We were honored to host two distinguished plenary speakers: Prof. Panos M. Pardalos from the University of Florida, USA, and Prof. Michal Černý from the Prague University of Economics and Business, Czech Republic. Their keynote addresses provided valuable insights into contemporary issues in mathematical economics and sparked lively discussions among attendees.

Our sincere thanks go to all the authors, reviewers, and members of the scientific committee, whose dedication and hard work have made MME 2024 a resounding success. We are also grateful to the Department of Computer Science, Faculty of Science, Jan Evangelista Purkyně University in Ústí nad Labem, for hosting this event and providing an inspiring environment for academic exchange. We also take this opportunity to acknowledge the outstanding contributions of the PhD students who participated in the competition organized by the Czech Society for Operations Research.

As you delve into these proceedings, we hope you find the research presented here both stimulating and informative, and that it contributes to your work in the field. We look forward to the continued growth and development of mathematical methods in economics and the ongoing collaboration among our community.

Program Committee Chairs of MME 2024

Hossein Moosaei, Jan Evangelista Purkyně University, Czech Republic
Josef Jablonský, Prague University of Economics and Business, Czech Republic
Helena Brožová (President of the Czech Society for Operations Research), Czech University of Life Science Prague, Czech Republic
Miloš Kopa (Vice-president of the Czech Society for Operations Research), Charles University, Czech Republic

Ethical guidelines

ETHICAL GUIDELINES

The Czech Society for Operations Research as a publisher of the Proceedings of the International Conference on Mathematical Methods in Economics is committed to the highest ethical standards. All authors, reviewers, and editors must comply with the following ethical principles. In case of any doubts regarding the Ethical guidelines, do not hesitate to contact the editors of the proceedings.

Ethical expectations

Editors' responsibilities:

- To acknowledge receipt of submitted manuscripts within two working days and to ensure an efficient, fair, and timely review process.
- To identify manuscripts that are fully within the scope and aim of the conference. To treat all submissions fairly without any favour of prejudice regarding authors' gender, religious or political beliefs, ethnic or geographical origin.
- To only authorise the review and publication content of the highest quality.
- To recuse himself or herself from processing any manuscript if he or she has any conflict of interest with any of the authors or institutions related to the manuscripts.
- To provide advice to the authors during the submission process when necessary.
- To be transparent with regards to the review and publication process with an appropriate care that individuals will not be identified when it is inappropriate to do so.
- To not use any parts or data of the submitted manuscript for his or her own future research as the submitted manuscript is not published yet.
- To respond immediately and take reasonable action when ethical problems occur concerning a submitted or published manuscript.

Authors' responsibilities:

- To carefully read the Instructions for authors published on the website of the conference.
- To claim that the submitted manuscript is not under consideration or accepted for publication elsewhere. To acknowledge and cite accordingly all content overlaps with already published content.
- To ensure that the submitted manuscript is original, prepared to a high scholarly standard and fully referenced using the prescribed referencing convention.
- To obtain permission to reproduce any content from other published sources. To clarify that all data used in the manuscript has been acquired following ethical research standards.
- To accurately acknowledge funding sources related to the submitted manuscript.
- To carefully read all the conditions included in the copyright form and to accept the copyright form during the submission process.
- To declare any potential conflicts of interest (real or apparent) at any stage during the publication process.
- To recognise that the editor has the final decision to publish the submitted, reviewed and accepted manuscript.
- To immediately inform the editors whenever any obvious error in his or her published manuscript is identified. To cooperate with the editors in the retraction or correction of the manuscript if necessary.

Reviewers' responsibilities:

- To immediately notify the editors whenever the reviewer feels unqualified to review the assigned manuscript or sees difficulties in meeting the deadline for completion of the review.
- To agree to review a reasonable number of manuscripts at the same time.
- To inform the editors if there is any possible conflict of interest related to the assigned manuscript. Specifically, to avoid reviewing any manuscript authored or co-authored by a person with whom the reviewer has an obvious personal or academic relationship.
- To treat the manuscript in a confidential manner and to not use any parts or data of the reviewed manuscript for his or her own future research as the reviewing manuscript is not published yet.
- To assist in improving the quality of the manuscript by reading the manuscript with appropriate care and attention, reviewing the manuscript objectively and being constructively critical.
- To immediately notify the editors of any similarities between the reviewing manuscript and another article either published or under consideration by another journal.

Procedures for dealing with unethical behaviour

Dealing with possible misconduct:

- Editors have a duty to act if they suspect any misconduct or if a claim of misconduct has been reported by anyone. This duty applies to both published and unpublished articles.
- Editors should not simply reject an article that raises concerns about possible misconduct. Editors are ethically obliged to pursue alleged cases.
- Whoever informs the editors of such conduct should provide sufficient information and evidence to initiate an investigation.
- Editors must take all allegations seriously and treat them similarly until a successful decision or conclusion is reached.
- Editors should first seek a response from those suspected of misconduct. If the editors are not satisfied with the response, they should ask the relevant employer, institution, or some appropriate body to investigate.
- Editors should make all reasonable efforts to ensure that a proper investigation into alleged misconduct is conducted. If this does not happen, editors should make all reasonable attempts to persist in obtaining a resolution to the problem.

Retraction of an article if:

- Editors have clear evidence that the findings are unreliable, either as a result of misconduct or honest error.
- The findings have previously been published elsewhere without proper cross-referencing, permission or justification (i.e. redundant publication). The journal that first published the article may issue a notice of redundant publication, but should not retract the article unless the findings are unreliable. Any journals that subsequently publish a redundant article should retract it and state the reason for the retraction.
- The article constitutes plagiarism and reports unethical research.
- The article should be retracted as soon as possible after the editors are convinced that the article is seriously flawed and misleading.

Outcomes of unethical behaviour in increasing order of severity:

- Informing or educating the author or reviewer where appears to be a misunderstanding or misapplication of acceptable standards.
- A more strongly worded letter to the author or reviewer covering the misconduct and as a warning to future behaviour.
- Retraction or withdrawal of a publication from the proceedings, in conjunction with informing the head of the author or reviewer's department, as well as Abstracting & Indexing services.
- Imposition of a formal embargo on contributions from an individual in the proceedings for a defined period.

Papers

Extension of Planning Poker by Work Contour Models in Project Management

Jan Bartoška[1], Josef Kunhart[2], Jiří Pilný[3]

Abstract. The paper describes the extension of Scrum Planning Poker with work contour Model. Planning Poker is used by an agile team to determine the difficulty of User Stories without determining the work effort in tasks for every User Story. Work effort variability affects the team: As it increases, team members' cooperation and proactivity decreases, and agile principles may be compromised in the team. The authors propose adding a new characteristic to the Planning Poker based on work effort models. The proposal builds on previous research by one of the authors on quantifying Student's syndrome and work contours of work effort. The previous research uses these modifications for Earned Value Management. Labeling User Stories on the Scrum Board with new characteristics added to the Planning Poker may affect cooperation on the tasks of the User Story during Scrum. User Stories marked as "last-minute work" should have set higher priority for team collaboration. The paper includes a case study for practical use and proposes a new mathematical model and concept for work contours and work effort in agile teams. The proposed concept extends the Planning Poker and enhances adherence to the agile principles in the team.

Keywords: Project management, Agile approach, Planning Poker, Work Contour, Work Effort, Resource Allocation, Scrum

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Project management is an effective tool to manage projects and changes within organizations. Project is a temporary activity undertaken to develop a unique product, service, or result with limited resources. Project teams accomplish the outcomes using diverse techniques, such as traditional and agile project management approaches and methods (PMI, 2021). Human resource management is a strategic approach to managing employees that work in the organization and constitute its most valuable resource. The purpose of human resource management is to ensure that organization achieves success via people that collectively contribute to achieve its goals (Armstrong, 2014). As stated by Smith (2010), human resources may cause problems and negatively affect project development, productivity, and outcomes due to people's behavior. Human resources thus make up an important part of projects and should be approached individually during the project planning. The importance of human resource management in project management is pointed out by Dwivedula (2019), who examines relations between project management and human resource management. The author also mentions the necessity to observe individual behavior of people and their work effort. The described human behavior causes project delays that negatively affect project planning and work scheduling.

Project management utilizes methods that derive from network analysis, mathematical methods, and Theory of Constraints. The most used methods for time analysis of projects are Critical Path Method and Critical Chain Method, as described and systematically compared by Lechler et al (2005). On one hand, Critical Path Method cannot fully implement project delays. On the other hand, the more advanced Critical Chain Method permits creating time reserves for elimination of delays as described in the Theory of Constraints (Kalender et al, 2014). Leach (1999) introduces presuppositions and starting points derived from the Theory of Constraints, Parkinson's law, and Student Syndrome. In this paper, we focus on the Student Syndrome.

Parkinson's law deals with assumption that the complexity of a task increases with amount of time allotted for the task completion. The increased task complexity then leads to the need for more resources and time than originally required (Parkinson, 1991). Student Syndrome is a tendency to finish the assignments until the last possible moment before deadline. For software projects, this means that developers work at a relaxed pace with a lot of slack

[1] Department of Systems Engineering FEM CULS, Kamýcká 129, Prague, bartoska@pef.czu.cz.
[2] Department of Systems Engineering FEM CULS, Kamýcká 129, Prague, kunhart@pef.czu.cz.
[3] Department of Systems Engineering FEM CULS, Kamýcká 129, Prague, pilnyj@pef.czu.cz.

during the project and postpone their tasks right until the deadline. As result, they waste time required to solve uncertainties that may impact the project work (Aljaž, 2023). Despite its name, Student Syndrome affects not only students, but any work environment that involves planning, tasks, and deadlines (Smith, 2010). Student Syndrome may lead team members to procrastinate, delaying the tasks until the deadline closes in. For people, it is natural to allocate work effort required to complete the task unevenly during the time interval determined by the task start and its deadline (König and Kleinmann, 2005). Task delays also lead to increased anxiety and stress among team members. Analyzing the impact of Student Syndrome on work schedules and project planning can provide insights for more effective human resource management in projects.

Agile software development is a modern approach to managing software projects. The agile approach focuses on delivery of business value in short iterations, close cooperation with the customer, and fast feedback loop. The most utilized agile method is Scrum that by a large margin dominates actual agile software development practices (Kadenic et al, 2023). Scrum defines fixed-time iterations called sprints, team roles, Scrum events, and artifacts, such as Product Backlog and Sprint Backlog (Schwaber and Sutherland, 2020). Scrum is flexible and productive software development method that emerged to implement complex software projects in dynamically changing environments. Scrum is considered a highly productive method to manage and improve projects in software companies (Guerrero-Calvache and Hernández, 2023).

Usman et al (2014) states that Planning Poker, expert judgment-based methods, and use case points are the most frequently applied techniques for estimating work effort in Scrum and agile software development. Story points and use case points are the most used size metrics for estimating task difficulty. Planning Poker is a work effort estimation technique that empowers Scrum team developers to decide on the difficulty of user stories in the Sprint. Planning Poker generally utilizes story points for the estimates, even though other techniques are possible. The results of the Planning Poker are subjective and reflect the consensus of all team members on the work effort for estimated user stories (Sudarmaningtyas and Mohamed, 2020). In this paper, we define an extension of Planning Poker with work contours model to deal with the Student Syndrome in the Scrum method. We propose a mathematical model for work contours in Scrum-based projects and perform a case study to verify the model on an array of user stories from a real project.

2 Materials and methods

2.1 Planning Poker in Scrum method

Scrum method is a widely used agile work approach that the authors define as a “lightweight framework”. The core event of Scrum is an iteration called a sprint. Sprints are the “heartbeat” of Scrum, where ideas turn into value. At the beginning of a sprint is Sprint Planning, which initiates laying out the work to be done. This resulting plan is created by the collaborative work of the entire Scrum Team. The quality of planning and its estimation is then directly reflected in the value delivered at the end of the sprint. Various techniques are used for estimation in Sprint planning event (Schwaber and Sutherland, 2020).

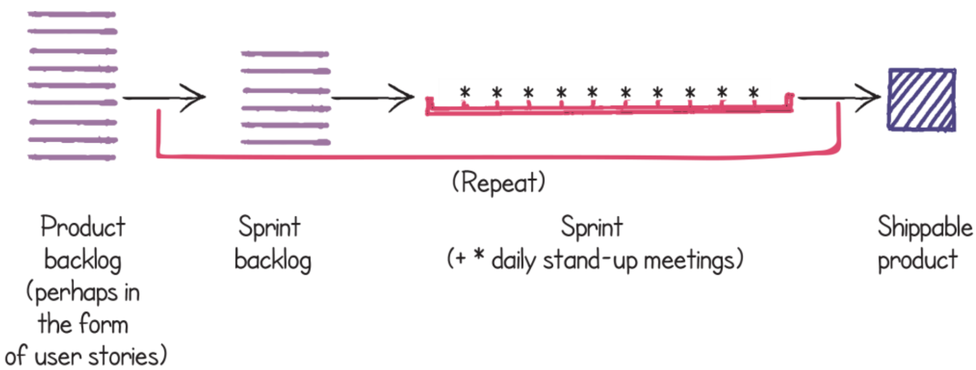


Figure 1 Scrum method – Sprint artifacts (AXELOS, 2018).

According to Grenning (2002), the best way for agile teams to estimate the work complexity is by playing Planning Poker. Planning Poker is used as a collaborative planning technique. Planning Poker combines expert opinion, analogy, and disaggregation into an enjoyable approach to estimating that results in quick but reliable estimates. Participants in Planning Poker include all the developers on the team (Cohn, 2006).

2.2 Work contour models in Projects

In the papers by Kučera et al. (2014) and Šubrt et al. (2021), the authors defined models for the Student Syndrome and the work contour (in various forms). The authors designed the Student Syndrome as a polynomial function (1) (Kučera et al., 2014).

$$p_1 = -120t^4 + 240t^3 - 150t^2 + 30t \quad (1)$$

The User Stories represent groups of tasks in a project. The work contours for User Stories can be designed as primitive linear functions. For example, back loaded (2) and front loaded (3) work contours can be defined as simple expressions (Kučera et al., 2014).

$$p_2 = 2t \quad (2)$$

$$p_2 = 2 - 2t \quad (3)$$

Both models can be combined for a simple and functional unified work contour of a User Story, as displayed in formula (4) (Kučera et al., 2014).

$$p = rp_1 + (1 - r)p_2 \quad (4)$$

In (4), parameter r represented the influence of the Student Syndrome (from 0 to 1). Either the influence of the Student Syndrome ($r = 1$) or the work contour ($r = 0$) prevails. The points made and mentioned here can be used for the purposes of this article, i.e. to extend the Planning Poker with work contour models in Project Management.

3 Results and Discussion

If we want to find an effective solution to express the work effort in a project (Smith, 2010; Kučera et al., 2014; Šubrt et al., 2021), in the case of an agile team (Scrum team), we need to define the relationship between the tasks and the User Story (Schwaber and Sutherland, 2020). A User Story represents a complex group of the tasks. The Scrum Team defines a User Story as a subset of tasks to be worked on. Therefore, a work effort with a specific work contour can be applied to the User Story.

The Planning Poker is obviously used by an agile team to determine the difficulty of User Stories when the Sprint Backlog is extended. But the team is not concerned with the determination of the work effort in User Stories. The problem is that the work effort of User Stories leads to variability that affects the team. The Agile Principles may be compromised: cooperation among members become changeable and proactivity of members increases or decreases very fast.

3.1 Extension of Planning Poker

The Planning Poker can be extended with the work effort of User Stories. The value of story points (p_i) that the team quantified for the User Story importance in the backlog can be used as the value of the work effort (e_i) of this subset of tasks. Therefore, we can write (5):

$$e = \frac{p}{\sum p} \quad (5)$$

Subsequently, the team can set the type of the work contour of the User Story. For simple use in Scrum, we may consider the work contour model from the proposal by Kučera et al. (2014). The formula for the work contour can be extended to (6).

$$c = e(rp_1 + (1 - r)p_2) \quad (6)$$

We can demonstrate the proposed solution on an explanatory case study, as presented in Table 1.

	Story Points (p)	Work effort (e)	Type of contour (p_2)
User Story 1	5	0,082	Bell
User Story 2	3	0,049	Back Loaded
User Story 3	1	0,016	Front Loaded
User Story 4	8	0,131	Turtle
User Story 5	5	0,082	Back Loaded
User Story 6	13	0,213	Back Loaded
User Story 7	20	0,328	Back Loaded
User Story 8	1	0,016	Turtle
User Story 9	5	0,082	Front Loaded

Table 1 Case Study: User Stories from Sprint Backlog with the extension of Planning Poker.

We can calculate the final work contours of the workflow using formulas (1), (5) and (6), where p_2 is given by the selected work contour (Kučera et al., 2014). For each User Story, we get a different work effort flow. Varying User Story difficulty leads to varying difficulty of the work contour in the Sprint. The User Stories in the case study can be viewed in Figure 2.

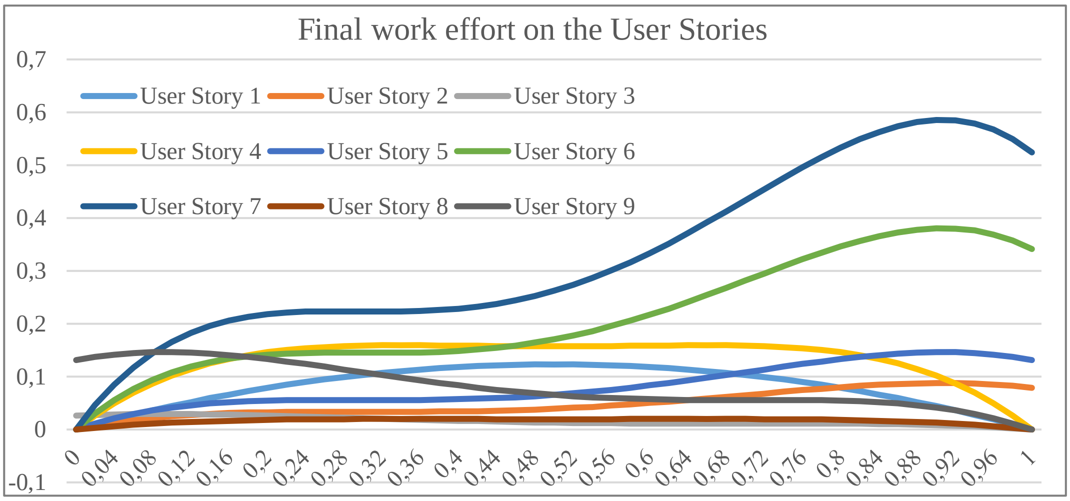


Figure 2 Case Study: Various contours of final work effort in the User Stories.

The User Stories with Back Loaded workflow have significant increase in work effort on both ends (User Story 6, User Story 7). In contrast with that, remaining User Stories have slowly growing contour throughout the workflow. The resulting workflows in the Case Study are determined by the work effort (e) and the type of the work contour (p_2).

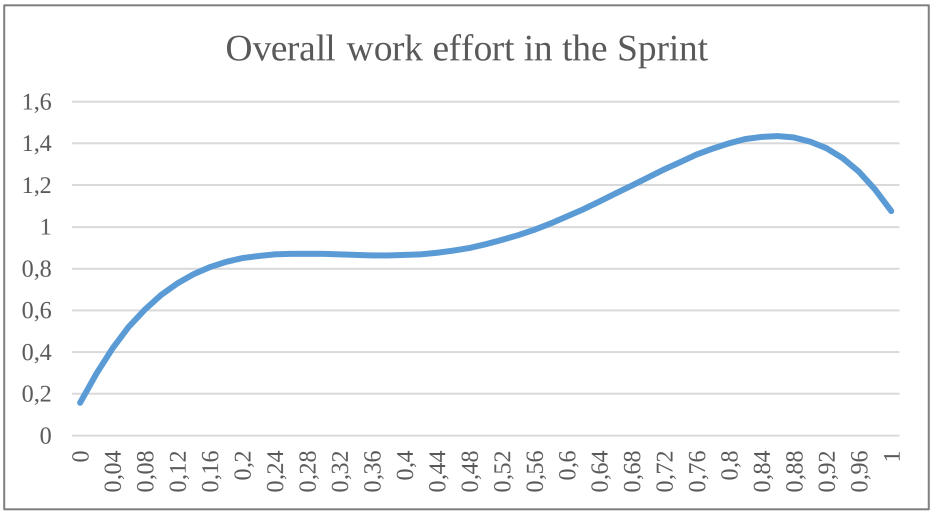


Figure 3 Case Study: The overall work effort in the team during the Sprint.

The total work effort in the team during the Sprint is given by the summation of the partial workflow (in partial User Stories). In the case study, we showed that the workflow in the Sprint can be variable, it can increase or decrease at the beginning or at the end. These changes in the teamwork workflow may not be known to the Scrum Master or Team Leader and may lead to lower efficiency in the project.

4 Conclusion

The paper deals with the possibility of extending Planning Poker with work contour models in agile Project Management (Scrum Team). The research aimed at formulating a combination of work effort and work contour for practical use in the team to estimate User Stories that represent subgroups of tasks in the project. Individual User Stories can be estimated with Story Points as well as marked with expected work contours. The Story Points thus determine the work effort (relatively to Sprint Backlog). The overall effort in the Sprint (both individual User Stories and the whole Sprint Backlog) can indicate an increase or decrease in the team's workload during the Sprint. This information about the team is new and significantly useful to the Scrum Master or Project Manager. For further research, it is important to verify the applicability of each team's work contour models within the Scrum method and to identify the relationship of observed teamwork effort (workload) during the Sprint as a function of work efficiency and team maturity.

References

Aljaž T (2023) Impact of the Student Syndrome on the project completion time. *Elektrotehniski Vestnik/Electrotechnical Review, 90*(4), 191–198

Armstrong M, Taylor S (2014) *Armstrong's handbook of human resource management practice* (13. ed.). Kogan Page

AXELOS (2018) *PRINCE2 Agile.* The Stationery Office, London. ISBN 9780113314676

Cohn M (2006) *Agile estimating and planning.* Person Education, Inc., Stoughton. ISBN 0-13-147941-5

Dwivedula R (2019) Human Resource Management in Project Management: Ideas at the Cusp. *European Project Management Journal 9*(1), 34-41. https://doi.org/10.18485/epmj.2019.9.1.5

Grenning J (2002) Planning Poker or How to avoid Analysis Paralysis while Release Planning. *Hawthorn Woods: Renaissance Software Consulting* 3, 22–23

Guerrero-Calvache M, Hernández G (2023) Team Productivity Factors in Agile Software Development: An Exploratory Survey with Practitioners. In *Applied Informatics* (pp. 261–276). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-46813-1_18

Kadenic M D, Koumaditis K, Junker-Jensen L (2023) Mastering Scrum with a focus on team maturity and key components of Scrum. *Information and Software Technology, 153,* 107079. https://doi.org/10.1016/j.infsof.2022.107079

Kalender Z, Günay N, Vayvay Ö (2014) Theory of Constraints: A Literature Review. *Social and Behavioral Sciences 150*(10), 930-936. https://doi.org/10.1016/j.sbspro.2014.09.104

König C J, Kleinmann M (2005) Deadline rush: A time management phenomenon and its mathematical description. *The Journal of Psychology 139*(1):33–45 https://doi.org/10.3200/jrlp.139.1.33-45

Kučera P, Bartoška J, Šubrt T (2014) Mathematical models of the work contour in project management. In: *32nd International Conference on Mathematical Methods in Economics 2014*. Conference Proceedings. Palacký University, Olomouc, pp 518-523

Leach L P (1999) Critical Chain Project Management Improves Project Performance. *Project Management Journal 30* (2), 39-51. https://doi.org/10.1177/875697289903000207

Lechler T G, Ronen B, Stohr E A (2005) Critical Chain: A New Project Management Paradigm or Old Wine in New Bottles? *Engineering Management Journal 17*(4), 45-58. https://doi.org/10.1080/10429247.2005.11431672

Parkinson C N (1991) *Parkinson's Law and Other Selected Writings on Management.* 1st ed., Singapure, Federal Publications (S)

PMI (2021) *A Guide to the Project Management Body of Knowledge (PMBOK® Guide) – Seventh Edition and The Standard for Project Management* (7th edition), Project Management Institute

Schwaber K, Sutherland J (2020) *The Scrum Guide*. https://scrumguides.org

Smith D (2010) The Effects of Student Syndrome, Stress, and Slack on Information Systems Development Projects. *Issues in Informing Science and Information Technology 7*(2010), 489-494. https://doi.org/10.28945/1222

Šubrt T, Bartoška J, Chamrada D (2021) Work Contour Models in Projects. In: *39th International Conference on Mathematical Methods in Economics*. Conference Proceedings. Czech University of Life Sciences Prague, pp 475-480

Sudarmaningtyas P, Mohamed R B (2020) Extended Planning Poker: A Proposed Model. *2020 7th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE).* https://doi.org/10.1109/icitacee50144.2020.9239165

Usman M, Mendes E, Weidt F, Britto R (2014) Effort estimation in agile software development: A systematic literature review. *Proceedings of the 10th International Conference on Predictive Models in Software Engineering.* https://doi.org/10.1145/2639490.2639503

Analysis of traffic accidents and weather in the Czech Republic

Blanka Bazsova , Lucie Chytilova

Abstract. Many factors affect a traffic accident. It can be people's moods, overwork, inattention, alcohol, excessive speed, and the weather. Since the latter factor has been changing a lot recently in connection with climate change, it is necessary to determine whether these general assumptions exist and affect the accident rate positively or negatively. The theoretical assumption is that these climate changes are related to temperature fluctuations in the Czech Republic in summer and winter. We are increasingly experiencing more tropical nights and, conversely, very icy days and freezing spring, sweltering summer and virtually snowless winter. These factors associated with temperature fluctuations and the number of rainfall events are gaining importance. They are worth looking at in terms of their monthly evolution over the past seven years. This article examines the weather and accident rate in the Czech Republic. A model uses standard accident rate variables (death, serious and minor injuries or material damage) and weather-specific variables. More precisely, the investigated econometric model thus includes the mentioned standard variables and average temperature and precipitation. The relevant model is examined and tested using correlation and regression analysis and their assumptions. Moreover, based on a detailed analysis, the dependence between the accident rate and the weather is proven, and it can be seen that the change in weather generally affects the accident rate positively.

Keywords: Corellation analysis, regression analysis, traffic accidents, weather.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Medina et al. (2022) argue that motorists are vulnerable to extreme weather events. Climate change around the world is likely to increase accident rates further. The impact of adverse weather on the development of accidents is at the forefront of interest in all countries. The influence of cold weather on accident rates in the Czech Republic was studied in their article by Brázdil et al. (2022). He focused his research on days when fog, rain, snow and ice were observed. Brázdil et al. (2022) analysed possible relationships between traffic accidents and weather in the Czech Republic from 1979-2020. The authors of the above-mentioned article looked at the consequences of these adverse effects on the number of traffic accidents in a given year and on the number of people who were slightly and seriously injured. They classified the above "vagaries" of weather into seven categories. They found that the effect of the relative annual proportion of all weather categories on the number of all accident characteristics fluctuated and further found that the annual numbers of accidents, fatalities and injuries attributed to individual weather categories were statistically significant in correlation with the series annual number of days with fog, rain, snow, ice and gusty winds. The closest relationship appeared for snowfall, with correlation coefficients between 0.76 and 0.94. Klaic (2001) analysed traffic accident data for the Zagreb area from 1981-1982. Klaic (2001) looked at the possible relationship between the daily number of accidents and the weather conditions that occurred after 5 consecutive days, two days before the specific day of the accident. Edwards (1999) studied the effect of weather on accident rates in the British Isles, and Amin et al. (2014) studied the impact of weather on accident rates in Canada.

Kim et al. (2021) studied traffic accidents caused by highway weather conditions. The authors concluded that bad weather in winter leads to higher mortality than in other seasons. The main causes of highway traffic accidents include human carelessness, vehicle defects, road conditions and weather factors. Gorzelanczyk (2021) developed a forecast for the number of traffic accidents in Poland depending on weather conditions. From the above-mentioned research, it is evident that many authors studied the influence of especially weather conditions on the accident rate

VŠB-TUO, Department of Systems Engineering, Sokolská třída 2416/33, 702 00 Ostrava - Moravská Ostrava, Czech Republic, blanka.bazsova@vsb.cz

VŠB-TUO, Department of Systems Engineering, Sokolská třída 2416/33, 702 00 Ostrava - Moravská Ostrava, Czech Republic, lucie.chytilova@vsb.cz

in their country. The weather is different for every country. It is determined by its location, intensity of solar radiation, wind speed, cyclonic flow and the season. This article investigates the relationship between weather and road accidents in the Czech Republic. Average temperatures and average amounts of rainfall will be investigated. They were obtained from the database of the Hydrometeorological Institute and covered the period from 2018 -2023.

The rest of the paper has the following structure: Section two provides information about the used data and the environment for which the analysis is given. Section three gives basic info about regression analysis and its assumptions.

The Results of Analysis section briefly describes this model's results and future use. The conclusion provides some conclusions and remarks.

2 Environment of the analysis

2.1 Analysis of variables

In 2023, 94,855 traffic accidents were recorded in the Czech Republic. 455 people were killed, 1,751 people were seriously injured, and 23,936 people were slightly injured. Total material damage in 2023 reached CZK 7,688 million. The number of accidents has increased slightly over the past ten years, while the number of dead and seriously injured has decreased slightly. The development of accidents stopped during the Covid era and fell to 94,920. However, the number of dead and injured in 2020 remains the same on average. If we follow the development of the number of accidents from 2016 to 2023, then the number of accidents in 2016 was 98,874, which increased until 2019, then fell below 95,000 in 2020, increased again in 2021 and was just below the threshold of 100,000. Specifically, 99,144 traffic accidents were registered. As seen in Figure 1, an increasing trend in the number of accidents was reported from 2016 to 2019, which was subsequently halted by COVID-19, and a decrease was observed from 2021. Traffic accidents can be influenced by several factors. One of the factors monitored by the Police of the Czech Republic is the day the traffic accident occurred. The most common accident day in 2023 was Thursday, followed by Monday as the second most common day. The fewest accidents happened on Sunday. Most of the accidents were investigated on 1 December 2023. The most tragic day was Friday, 21 April, followed by 27 May, Sunday, 28 May, Friday, 2 June, Saturday, 3 June, Thursday, 24 August and 13 November, when 5 people died. The highest increase in the number of fatalities was recorded among motorcycle and truck drivers (12 fatalities and more). Regarding the causes of traffic accidents caused by drivers of motor vehicles, the most significant factor in the number of traffic accidents in 2023 was incorrect driving (68.8% of accidents); 183 people were killed in these accidents. In the 1st place, regarding the causes of accidents, I was not paying attention to the vehicle's driving. In 5th place was the failure to adapt the speed to the road condition. Excessive speed and incorrect overtaking. The most tragic accidents were driving in the opposite direction. In 4,766 (an increase of 12 accidents), the culprit of the traffic accident was found to have consumed alcohol. (Police of the Czech Republic).

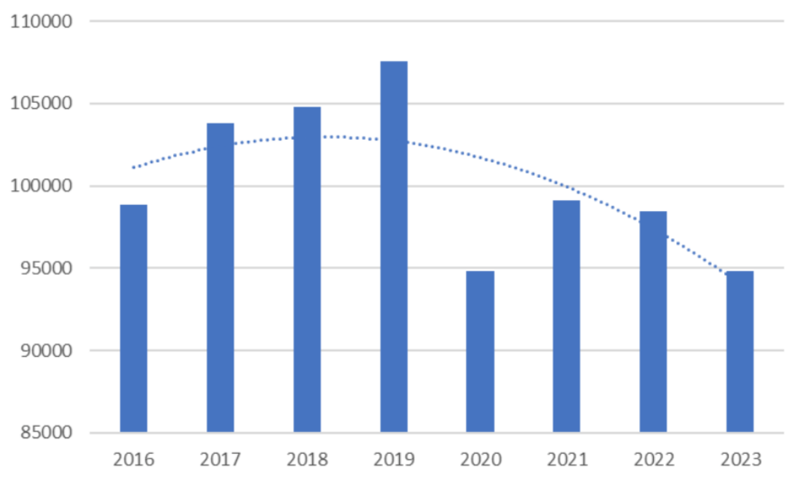


Figure 1 Development of the accident rate in 2016-2023. Source: Police of the Czech Republic

https://www.policie.cz/

2.2 The influence of the weather factor on the accident rate

The number of road accidents may be influenced by several factors. The Czech Republic, with an area of 78 871 km^2, is a landlocked country located in the middle of the temperate zone of the northern hemisphere in the central part of Europe. (CÚZK, 2023) Its climate is characterised by mutual penetration and mixed oceanic and continental influences. It is characterised by a westerly flow with a predominance of westerly winds, intense cyclonic activity causing frequent air mass changes and relatively abundant precipitation. The heaviest average rainfall between 2016-2023 was recorded in the months of June to August. The highest average temperatures were recorded in 2019, in June, when an average of 20.7 degrees Celsius was measured; on the contrary, the lowest was also measured in 2017, in March, and it was measured at - 5.6 degrees Celsius. When we compare the average maximum temperatures from 2016-2023 with the long-term average measured from 1990-2020, we find that the average temperature in the Czech Republic is rising slightly (Czech Hydrometeorological Institute, 2023).

3 Regression analysis

Regression analysis is a statistical tool for the investigation of relationships between variables. The usual goals of regression are prediction or control. Regression analysis gives information about the relationship between the chosen explained variable (denoted as y) by filling in values for the explanatory variables (X) in an equation estimated from data. Regression analysis also offers a measure of the probable accuracy of its predictions. Regression analysis for control implies that it could be manipulated with one or more explanatory variables to change the value of the explained variable.

The multiple regression model combines an equation relating the explained variable y (a.k.a. the dependant variable) to a set of predictors (a.k.a. independent variables) x_1, x_2, ..., x_k with a collection of supporting assumptions. The equation of the model describes:

$$y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + u_t, \tag{1}$$

where β_i are regression coefficients of independent variables (β_0 is a constant) and u_t expresses the error of the equation (residuals), model, noise, measurement error, or the sum of the factors not included. Multiple regression models could also be in the form of logarithms or differences, depending The optimal multiple regression model describes a data-generating process. The more actual data resemble observations from such an optimal process, the more reliable statistical results are obtained (such as confidence or prediction intervals).

In addition to the assumed equation (1), the assumptions that complete this multiple regression model are the following:

1. the independent variables x_1, x_2, ..., x_k are non stochastic;
2. the random errors are normally distributed, i.e. u_t $N(0, \sigma^2)$, that is the normality of residues;
3. the mean of random errors is zero, i.e. $E(u_t) = 0$;
4. the variance of random errors is constant, i.e. $\mathrm{Var}(u_t) = \sigma^2$, that is the homoscedasticity;
5. the random errors are uncorrelated, i.e. $\mathrm{Cov}(u_i, u_j) = 0$ for $i \neq j$, that means that model is not autocorrelated;
6. independent variables x_1, x_2, ..., x_k are collinear, i.e. there is no multicollinearity.

4 Problem solution

The authors of this paper follow the aim to predict the number of traffic accidents. The relationship is examined by a correlation matrix and, after confirming the right assumptions, is involved in the regression model, which shows the development of the accidents for the periods claimed. Formulation of the stochastic regression model according to the above-mentioned topic:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + \beta_5 x_{5t} + \beta_6 x_{6t} + u_t, \tag{2}$$

where y_t the dependent variable is the number of accidents and independent variables x_{1t}-x_{6t} are the number of killed, heavily injured and slightly injured, material damage, the average temperature in the month and the average amount of rainfall in the month, in that order. The assumption for all regression coefficients had been set up as $\beta_i > 0$ where $i = 1, \ldots, 6$. The introduced model (2) is the main analysis in the following part of the paper.

It must be said that the original data had not been stationary, so the data had changed into different data. After this change, the Dick-Fuller test tested the data, and all data was stationary. The notation of the variable had been changing from x_{it} onto dx_{it}.

https://www.chmi.cz/

Based on statistical testing, correlation analysis, and econometric testing, three models were selected that are worth mentioning during this analysis. The specific models, their designations and specifications are as follows:

1. Model A - data not seasonally adjusted and working with independent variables - dx_{3t}, dx_{4t} and dx_{5t};
2. Model B - data seasonally adjusted and working with independent variables - dx_{3t}, dx_{4t} and dx_{5t};
3. Model C - data seasonally adjusted and working with independent variables - dx_{2t}, dx_{3t} and dx_{4t}.

The specific results of all these models are presented below.

4.1 Model A - Seasonally unadjusted data

A regression model with seasonally unadjusted data is in the following form:

$$dy_t = -0.680 + 1.409dx_{3t} + 4.34 \cdot 10^{-6}dx_{4t} - 0.115dx_{5t} + u_t, \tag{3}$$

The general info about the Model A is given in Table 1.

The coefficient of regression is 72.4% for Model A, and the adjusted coefficient of regression (better for the model with more independent variables) is 71.5%. So, the model is very good. The model considers three independent variables at a 5% level of significance: the number of lightly injured, material damage, and the average monthly temperature. The assumption for β_i had been a positive value. However, it is seen that if temperature increases, the number of accidents decreases. So, it looks like winter weather, snow, and ice are worse than high temperatures and heat exhaustion, which can be compensated by air conditioning. This model is good, without heteroscedasticity, well specified, but there is a small problem with negative autocorrelation.

	value	
R^2	0.724	
R^2_{adj}	0.715	
DW	2.368	
	value	significance
β_0	-0.680	0.980
β_3	1.409	< 0.001
β_4	$4.34 \cdot 10^{-6}$	< 0.001
β_5	-0.115	0.023

Table 1 Information about Model A

4.2 Model B - Seasonally adjusted data

A regression model with seasonally adjusted data is of the form:

$$dy_t = -0.691 + 1.380dx_{3t} + 5.587 \cdot 10^{-6}dx_{4t} - 31.609dx_{5t} + u_t, \tag{4}$$

The general info about the Model B is given in Table 2.

The coefficient of regression is 77.2%, and the adjusted regression coefficient is 76.5%. So, Model B is better than Model A. Again, this model considers the same three independent variables at a 5% significance level. Generally, the results were similar. In this model, the problem is the assumption about temperature. Overall, even Model B is specified correctly, it is worse than Model A. This model has quadratic heteroscedasticity cost by dx_4 and has a small problem with negative autocorrelation.

	value	
R^2	0.772	
R^2_{adj}	0.765	
DW	2.423	
	value	significance
β_0	−0.691	0.860
β_3	1.380	< 0.001
β_4	$5.587 \cdot 10^{-6}$	< 0.001
β_5	−31.609	0.025

Table 2 Information about Model B

4.3 Model C - Seasonally adjusted data

A regression model with seasonally adjusted data is of the form:

$$dy_t = 16.250 - 4.142dx_{2t} + 1.943dx_{3t} + 3.985dx_{4t} + u_t, \quad (5)$$

Table 3 gives the general info about Model C.

R^2	0.899	
R^2_{adj}	0.896	
DW	1.999	
	value	significance
β_0	16.250	0.651
β_2	−4.142	< 0.001
β_3	1.943	< 0.001
β_4	−3.985	< 0.001

Table 3 Information about Model B

The coefficient of regression is 89.9%, and the adjusted regression coefficient is 89.6%. So, Model C is the best one. The model considers three independent variables at a 5% level of significance: slightly injured, material damage, and the average temperature in the month. Again, this model considers the same three independent variables at a 1% significance level. This model has two β_i negative, so it should be considered if it is in analogy with all theory. The temperature seems logical to be changed into the negative, but the slightly injured, not really. This model also has the worst problem with autocorrelation even if it is specified well, and the problem with heteroscedasticity was not found.

5 Conclusion

The investigation's results using three regression models confirmed that the accident rate is indeed related to the number of seriously injured and lightly injured, the amount of material damage, and extreme weather conditions.

This article set out to solve the problem of car accidents and the weather. It is the first article that should lead to further research by the authors. So the analyses are only basic. However, even here, the results are already visible. It can be seen that it is probably better to use data with seasonally unadjusted data because the value of the β_i coefficients seems more realistic for this model - Model A, and also, it is more logical to use data with season in the Czech Republic if there is a season and the weather changes. Model A also had the least problems with econometric verification. It is also necessary to focus on a better analysis of the environment - at least, the temperature would not positively affect car accidents.

For future research, however, it is necessary to focus on improving both the model (trying other types of regression models or methods) and the database (more variables, a longer timeline, etc.). Despite this, a deeper analysis of the environment should also be carried out with transport and weather experts.

References

[1] Amin, S. Md. R. A. Zareie, Amador-Jiménez, L. E. (2014). Climate change modeling and the weather-related road accidents in Canada. Transportation and Research part D: Transport and Environment, 32

[2] Brázdil, R., K. Chromá, P. Zahradníček (2020). Weather and traffic accidents in the Czech Republic, 1979-2020. Theoretical and applied climatology 149(1-2).

[3] Edwards, J. B. (1998). The relationship between road accident severity and recorded weather. Journal of Safety Research 29(4).

[4] Gorzelanczyk, P.(2023) Using neural networks to forecast the number of road accidents in Poland taking into account weather conditions. Results Engineering, 17(3/2023)

[5] Klaic, Z. B. (2001). Weather types and traffic accidents. Collegium Antropologicum 25(1):245-54.

[6] Kim, D., S. Jung, S. Yoon (2021). Risk Prediction for Winter Road Accidents on Expressways. Applied Science Basel, 11(20).

[7] Medina, R. M., G. Cerrvone and N. M. Maters (2017). Characterizing and Predicting Traffic Accidents in Extreme Weather Environments. Proffessional Geographer 69(1).

Welfare or Poverty of Czech Pensioners during the Energy Crisis: A Linear Regression Model

Diana Bílková[1]

Abstract. The paper is focused on the total net monthly income of the pensioner's household at the time of the energy crisis, which represents the explained quantitative variable in the linear regression model. This income represents one of the main characteristics of the quantitative aspect of living standards, and not only for households of senior citizens. The initial explanatory quantitative variables entering into the model are number of household members, property of the pensioner's family, number of living rooms, pensioner's age, municipality size, number of pensioner's children, age of the pensioner's partner and length of current partnership.

The results of the sample survey are for the year 2022 and include only pensioners aged 65 and over. The data was provided by the Czech Statistical Office. The length of current partnership variable was removed from the model due to harmful multicollinearity. The sequential F-test showed that the most important explanatory variable is the number of household members in terms of influence on the explained variable.

Keywords: net monthly income of the pensioner's household, welfare and poverty of pensioners, negative effect of the energy crisis, multiple linear regression, stepwise regression, sequential F-test, multicollinearity, B-coefficient, homoscedasticity and heteroscedasticity

JEL Classification: I31, I32, D12, C51

AMS Classification: 62H12, 62H15

1 Introduction

Linear regression analysis is one of the most commonly procedures used to reveal relationship between various variables practically in all areas of science and research. For example in archeology [1], agriculture and horticulturae [2]–[4], hydrology [5], chemosphere [9], radiology [10], computer science and machine learning [7] and [6]. Usage within the theory of statistics is presented in [8]. This paper is aimed at the standard of living of Czech pensioners at the time of the onset of the energy crisis in 2022. The data set for this research comes from a sample survey of the Chzech Statistical Office.

The main quantitative variable characterizing the standard of living of a pensioner's household is the total net monthly income of the household, as pensioner's households do not necessarily have only income from old-age or survivor's pensions. This is the explained variable in the linear regression model. The initial set of explanatory variables is made up of a group of eight quantitative variables: number of household members, property of the pensioner's family (in CZK), number of living rooms, pensioner's age (completed years), municipality size (number of inhabitants), number of pensioner's children, age of the pensioner's partner (completed years) and length of current partnership (completed years). SPSS statistical software was used for data processing.

In this work, the method of least squares is used to estimate the parameters of the linear regression function. The method of least squares is commonly used in the construction of regression models. The method of least squares provides sufficient parameter estimates only when all assumptions about the data and about the regression model are simultaneously met. If these assumptions are not met, the least squares results lose their properties.

2 Construction and Verification of Assumptions

Table 1 presents a matrix of sample simple correlation coefficients between explanatory variables. This matrix is symmetric, only the correlation coefficient between length of current partnership and age of the pensioner's partner variables indicates harmful multicollinearity between these variables. We therefore remove the length of current

[1] Prague University of Economics and Business, Faculty of Informatics and Statistics, Department of Statistics and Probability, Sq. W. Churchill 1938/4, 130 67 Prague 3, Czechia, bilkova@vse.cz.

partnership variable from the model. After that, no absolute value of sample simple correlation coefficient exceeds the value of 0.5. Thus, harmful multicollinearity is not present in the model now. The residuals corresponding to the model with seven explanatory variables after removing the variable the length of current partnership are shown in Figure 1.

	Variable							
Variable	**length of current partnership**	**property of the pensioner's family**	**number of household members**	**number of pensioner's children**	**number of living rooms**	**age of the pensioner's partner**	**pensioner's age**	**municipality size**
length of current partnership	1.0000	-0.0091	0.0096	-0.0211	0.1369	-0.8479	-0.2573	-0.0009
property of the pensioner's family	-0.0091	1.0000	0.0638	0.0462	-0.2590	-0.0533	0.0500	-0.1978
number of household members	0.0096	0.0638	1.0000	-0.0728	-0.2485	-0.3291	0.1321	-0.0079
number of pensioner's children	-0.0211	0.0462	-0.0728	1.0000	-0.0122	0.0052	-0.0011	0.0684
number of living rooms	0.1369	-0.2590	-0.2485	-0.0122	1.0000	0.0792	0.0251	0.0014
age of the pensioner's partner	-0.8479	-0.0533	-0.3291	0.0052	0.0792	1.0000	0.1755	0.0014
pensioner's age	-0.2573	0.0500	0.1321	-0.0011	0.0251	0.1755	1.0000	-0.1005
municipality size	-0.0009	-0.1978	-0.0079	0.0684	0.0014	0.0014	-0.1005	1.0000

Table 1 Matrix of sample simple correlation coefficients between explanatory variables (output from SPSS)

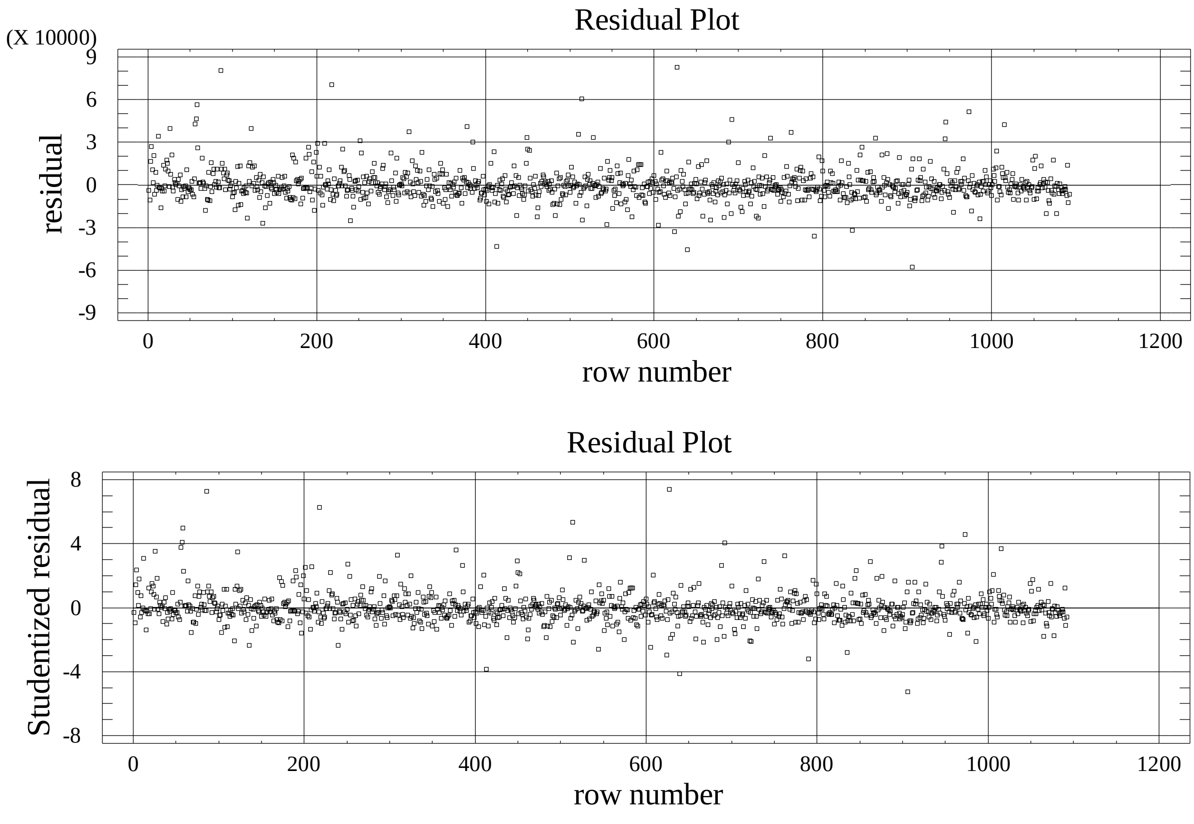


Figure 1 Residual plot (output from SPSS)

The stepwise regression process and the forward selection method, which gradually includes the explanatory variables into the model through the sequential F-test, are presented in Table 2. We can see from this table that no variable is included in the model at the beginning of the procedure. Both the coefficient of determination (R-squared) and the adjusted coefficient of determination (Adjusted R-squared) show zero percent and the mean squared error is 2.78534E8.

We can see from this table that no variables were included in the model at the beginning of the procedure. Both the coefficient of determination (R-squared) and the adjusted coefficient of determination (Adjusted R-squared) show zero percent and the mean squared error is 2.78534E8.

Stepwise regression

Method: forward selection
F-to-enter: 4.0
F-to-remove: 4.0

Step 0:
0 variables in the model. 1.091 d. f. for error
R-squared = 0.00 % Adjusted R-squared = 0.00 % MSE = 2.78534E8

Step 1:
Additing variable: **number of household members** with F-to-enter = 854.518
1 variable in the model. 1.090 d. f. for error.
R-squared = 53.94 % Adjusted R-squared = 53.90 % MSE = 1.56275E8

Step 2:
Additing variable: **property of the pensioner's family** with F-to-enter = 102.757
2 variables in the model. 1.089 d. f. for error.
R-squared = 58.78 % Adjusted R-squared = 58.70 % MSE = 1.42932E8

Step 3:
Additing variable: **number of living rooms** with F-to-enter = 45.5871
3 variables in the model. 1.088 d. f. for error.
R-squared = 60.84 % Adjusted R-squared = 60.73 % MSE = 1.3731E8

Step 4:
Additing variable: **pensioner's age** with F-to-enter = 18.0235
4 variables in the model. 1.087 d. f. for error.
R-squared = 61.57 % Adjusted R-squared = 61.43 % MSE = 1.3538E8

Step 5:
Additing variable: **age of the pensioner's partner** with F-to-enter = 16.5065
5 variables in the model. 1.086 d. f. for error.
R-squared = 62.36 % Adjusted R-squared = 62.19 % MSE = 1.33293E8

Step 6:
Additing variable: **number of pensioner's children** with F-to-enter = 6.55425
6 variables in the model. 1.085 d. f. for error.
R-squared = 62.65 % Adjusted R-squared = 62.44 % MSE = 1.32615E8

Step 7:
Additing variable: **municipality size** with F-to-enter = 5.37691
7 variables in the model. 1.084 d. f. for error.
R-squared = 62.88 % Adjusted R-squared = 62.64 % MSE = 1.32082E8

Final model selected.

Table 2 Stepwise regression, forward method and sequential F-test (output from SPSS)

The explanatory variable number of household members was included in the model in the first step, which increased the coefficient of determination to 53.94 percent, i.e. 53.94 percent of the variability of the observed values of the explained variable total net monthly income of the pensioner's household was now explained by a linear regression function (at this moment straight line) and the explanatory variable number of household members. The adjusted coefficient of determination is corrected by the number of parameters of the given regression function, and its value is always slightly lower than the value of the classic coefficient of determination. Mean squared error fell in the first step to 1.56275E8. The explanatory variable property of the pensioner's family was included in the model in the second step, which increased the coefficient of determination from 53.94 percent to 58.78 percent. This means that 58.78 percent of the variability of the observed values of the explained variable total net monthly income of the pensioner's household was now explained by a linear regression function (at this moment plane) and the explanatory variable property of the pensioner's family. Mean squared error fell in the first step to 1.42932E8. The method continues step by step, individual explanatory variables are gradually included in the model, which gradually increases the value of the coefficient of determination and decreases the value of the mean squared error. In the last the seventh step, the last explanatory variable having a statistically significant effect on the value of the explained variable was inserted into the model. This is municipality size variable. The multiple coefficient of determination reached the value 62.88 percent. This means that 62.88 percent of the variability of the observed values of the explained variable total net monthly income of the pensioner's household was explained by the constructed regression hyperline and a group of seven explanatory variables (the variable length of current partnership was not indeed included into the model) and the remaining 37.12 percent indicate the influence of unspecified factors. Mean squared error fell in the seventh step to 1.32082E8.

Multiple Regression Analysis

Dependent variable: Total net monthly income of the pensioner's household

Variable	Estimate	Standard Error	T Statistic	P-value
CONSTANT	20751.5	4135.97	5.01732	0.0000
property of the pensioner's family	0.00062969	0.0000875646	7.19115	0.0000
number of household members	11788.7	684.266	17.2283	0.0000
number of pensioner's children	-757.815	315.671	-2.4006	0.0164
number of living rooms	2192.82	315.268	6.95543	0.0000
pensioner's age	-253.085	54.1545	-4.67339	0.0000
age of the pensioner's partner	58.6524	13.5507	4.32836	0.0000
municipality size	0.00170211	0.000734045	2.31882	0.0204

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio	P-value
Model	1.91091E11	7	2.72987E10	206.68	0.0000
Residual	1.43177E11	1084	1.32082E8		
Total	3.0388E11	1091			

R-squared = 62.8837 percent
R-squared (adjusted for d.f.) = 62.6440 percent
Standard Error of Est. = 11492.7
Mean absolute error = 7417.05
Durbin-Watson statistic = 1.88144

Table 3 Results of regression analysis (output from SPSS)

Table 3 shows that all individual t-tests as well as the overall F-test are statistically significant at the five percent significance level. So, the model is fine as a whole. The value of the Durbin-Watson statistic is 1.88144, so it is in the interval (1.8; 2.2), we probably have not autocorrelation problem in the residuals. Glaser's test was used to verify homoscedasticity.

In the column labeled Estimate of Table 3, we find the sample partial regression coefficients linked to the explanatory variables in the first column of the same table. We obtain the B-coefficients by multiplying the respective sample partial regression coefficient by the ratio of the sample standard deviation of the explanatory variable, for which the respective sample partial regression coefficient stands and the sample standard deviation of the explained variable. B-coefficients are dimensionless therefore, they allow a better comparison of the influence of individual explanatory variables on the explained variable.

3 Results

Table 3 determinate the form of linear regression function using the sample partial regression coefficients, see Table 4. Given the different measurement units of explanatory variables, the values of the sample partial regression coefficients are given in the form in which they were in the output from the SPSS program, i.e. to a different number of decimal places.

Total net monthly income of the pensioner's household = 20751.5	+ 0.00062969 · property of the pensioner's family +
	+ 11788.7 · number of household members –
	– 757.815 · number of pensioner's children +
	+ 2192.82 · number of living rooms –
	– 253.085 · pensioner's age +
	+ 58.6524 · age of the pensioner's partner +
	+ 0.00170211 · municipality size.

Table 4 Final linear regression function

In addition to the constant 20751.5, Table 4 shows sample partial regression coefficients that indicate how many units the value of the explained variable in average changes (increase in the case of direct linear dependence when the sample partial regression coefficient is positive or decrease in the case of indirect linear dependence when the sample partial regression coefficient is negative), if the value of the explanatory variable for which the relevant sample partial regression coefficient is located increases by its one unit, provided that the values of the other explanatory variables do not change. For example, sample partial regression coefficient +0.00062969 shows that if the explanatory variable property of the pensioner's family increases by one CZK, the explained variable total net monthly income of the pensioner's household increases by 0.00062969 CZK on average, provided that all other explanatory variables remain unchanged. Sample partial regression coefficient –757.815 means that each additional child of a pensioner means a decrease in the total net monthly income of his/her household by 757.815 CZK on average, assuming that all other explanatory variables are constant.

Variable	Standard deviation	B-coefficients
total net monthly income of the pensioner's household	16,681.68014 CZK	–
number of household members	0.754536 members	0.533220
property of the pensioner's family	3,811,941.928290 CZK	0.143891
number of living rooms	1.302788 rooms	0.171253
pensioner's age	6.638859 years	–0.100721
municipality size	498,330.4896 inhabitants	0.050847
number of pensioner's children	1.320602 children	–0.059992
age of the pensioner's partner	7.179774 years	0.025244

Table 5 Standard deviations of individual explained and explanatory variables and conversion of partial regression coefficients to dimensionless B-coefficients

The disadvantage of these partial regression coefficients is their dependence on different measurement units of explanatory variables, which makes it impossible to compare the intensity of the influence of individual explanatory variables on the explained variable. The so-called B-coefficients overcome this disadvantage, see Table 5.

If we compare the absolute values of the B-coefficients from Table 5, then we get the order of the explanatory variables which the most influence the explained variable down to the explanatory variables that influence the least the explained variable (either directly or indirectly), see Table 6.

Order	Variable	The absolute value of the B-coefficient	Direction of dependence
1st	number of household members	0.533220	direct
2nd	number of living rooms	0.171253	direct
3rd	property of the pensioner's family	0.143891	direct
4th	pensioner's age	0.100721	indirect
5th	number of pensioner's children	0.059992	indirect
6th	municipality size	0.050847	direct
7th	age of the pensioner's partner	0.025244	direct

Table 6 The order of the explanatory variables the most influencing the explained variable

4 Conclusions

Considering the group of explanatory variables analyzed, the explained variable total net monthly income of the pensioner's household is the most directly linearly dependent on the explanatory variable number of household members, further on the explanatory variable number of living rooms and on the explanatory variable property of the pensioner's family. The explained variable total net monthly income of the pensioner's household is the least directly linearly dependent on the explanatory variable age of the pensioner's partner and on the explanatory variable municipality size. The linear dependence of the explained variable total net monthly income of the pensioner's household on the explanatory variables pensioner's age and number of pensioner's children is indirect.

Acknowledgements

This paper was subsidized by the funds of institutional support of a long-term conceptual advancement of science and research number IP400040 at the Faculty of Informatics and Statistics, University of Economics and Business, Czech Republic.

References

[1] Ali, T. H. & Saleh, D. M. (2021). Comparison between Wavelet Bayesian and Bayesian Estimators Toremedy Contamination in Linear Regression Model. *Palarch's Journal of Archaeology of Egypt/Egyptology*, *18*(10), 3 388–3 409.

[2] Al-Saif, A. M., Abdel-Sattar, M. Eshra, D. H., Sas-Paszt, L. & Mattar, M. A. (2022). Predicting the Chemical Attributes of Fresh Citrus Fruits Using Artificial Neural Network and Linear Regression Models. *Horticulturae*, *8*, 1–25.

[3] Amoozad-Khalili, M. Rostamian, R., Esmaeilpour-Troujeni, M. & Kosari-Moghaddam, A. (2020). Economic Modeling of Mechanized and Semi-Mechanized Rainfed Wheat Production Systems Using Multiple Linear Regresssion Model. *Information Processing in Agriculture*, *7*(1), 30–40.

[4] Cankaya, S., Kayaalp, G. T., Sangun, L., Tahtali, Y. & Akar, M. (2006). A Comparative Study of Estimation Methods for Parameters in Multiple Linear Regression Model. *Journal of Applied Animal Research, 29*(1), 43–47.

[5] Herrig, I. M., Böer, S. I., Brennholt, N. & Manz, W. (2015). *Water Research, 85*, 148–157.

[6] Maulud, D. & Abdulazeez, A. M. (2020). A Review on Linear Regression Comprehensive in Machine Learning. *Journal of Applied Science and Technology*, *1*(4), 140–147.

[7] Ottaviani, F. M. & De Marco, A. (2022). Multiple Linear Regression Model for Improved Project Cost Forecasting. *Procedia Computer Science*, *196*, 808–815.

[8] Ugwuowo, F. I., Oranye, H. E. & Arum, K. C. (2023). On the Jackknife Kibria-Lukman Estimator for the Linear Regression Model. *Communications in Statistics – Simulation and Computation, 52*(12), 6 116–6 128.

[9] Zhang, G., Liu, X., Lu, S., Zhang, J. & Wang, W. (2020). Occurrence of Typical Antibiotics in Nansi Lake's Inflowing Rivers and Antibiotic Source Contribution to Nansi Lake Based on Principal Component Analysis-Multiple Linear Regression Model. *Chemosphere, 242*, 125–269.

[10] Zou, K. H., Tuncali, K., Silverman, S. G. (2023). Correlation and Simple Linear Regression. *Radiology*, *227*(3), 617–622.

Asset return as a vague element in investment portfolio selection: fuzzy mathematical modelling

Adam Borovička[1]

Abstract. The return, or its level, is often an unstable factor of the intended investment. It can be expressed deterministically (e.g. by a mean) or stochastically (as a random variable with a particular probability distribution). The first option loses some valuable information. The second option can complicate subsequent, particularly computational, tasks when expressing a random process explicitly. Another view is represented by use of the apparatus of fuzzy set theory. Return as a (triangular) fuzzy number – fuzzy return – can adequately quantify the approximation, or vagueness, associated with its expected value. The triangular form offers several particular computational advantages. Their eventual comparison also proceeds more easily. Thus, processing such fuzzy information with a mathematical model for the selection of an investment portfolio need not be difficult. However, a crucial issue is the determination of the three parameters of the fuzzy number, which is sometimes neglected in papers on this topic, although it can logically significantly affect the result. The application power of the fuzzy return concept integrated into the mathematical programming model is demonstrated through a case study of ESG mutual fund portfolio selecting.

Keywords: fuzzy return, mutual fund, portfolio selection, triangular fuzzy number

JEL Classification: C44, C61, G11
AMS Classification: 90B50, 90C30

1 Introduction

Investing is an increasingly powerful phenomenon. On the one hand, people try to value their spare funds. On the other hand, they are understandably worried about losing them. After all, as the report of the Ministry of Finance of the Czech Republic (2024) shows, the desire for profit, in the best sense of the word, is getting stronger, the number of investors, and consequently the assets in investments, is getting larger. However, the fear of loss, or of "something unknown", is absolutely justified, not only in Czech society, it is still strongly rooted. The world of investments is unstable, volatile, full of sometimes difficult to predict situations and processes. Nevertheless, an effort is made to capture all these aspects in order to streamline the investment decision-making process.

Such probably the most debated aspect of investment is the return. Not surprisingly, appreciation is the main purpose of the investment. Forming a realistic expectation is a daunting task. The return can be derived using a simple deterministic characteristic, usually the mean. This characteristic represents only a sort of 'most common' scenario. The part of the information reflecting alternative developments is lost. The return can then be seen as a random variable, which by its regular distribution of its values, gives a more comprehensive picture of its form. The presence of a stochastic element, however, rather complicates the mathematical and technical processes of investment portfolio selection in particular, and investment decision-making in general. Stochasticity is not always easy to describe or quantify, especially because of the existence of potentially influencing non-random factors such as political, social or cultural. Such a vagueness can be expressed using fuzzy numbers. Specifically, a fuzzy number can effectively quantify the instability of the return. Moreover, the fuzzy element, unlike the stochastic one, can be relatively easy to integrate into mathematical model for portfolio selection.

The concept of fuzzy return is elaborated in a number of scientific publications. Most commonly, fuzzy return is expressed using triangular (Duan and Stahlecker, 2011; Mandal et al., 2024; Li et al., 2009; Qin et al., 2011; Li et al., 2009); or trapezoidal (Mansour, 2019, Mashayekhi, 2016) fuzzy number. The reason is the piecewise linear membership function, ease of computational operations or known technical comparison. Objects with such properties are then easier to work with in the mathematical description of the real, in our case, investment world. The correct specification of the fuzzy return depends on the setting of the parameters of the particular fuzzy number.

[1] Prague University of Economics and Business, Department of Econometrics, W. Churchill Sq. 4, Prague, Czech Republic, adam.borovicka@vse.cz.

Since this task may not be easy, I propose to use a fuzzy structure with fewer parametric inputs, i.e. a triangular fuzzy number. Even in reputed journals, such as Applied Soft Computing and similar, articles operating with fuzzy return, mostly do not explain much just the determination of parameters, neither theoretically nor practically. Empirical analyses rather operate with illustrative data, which, moreover, have been "somehow determined". Therefore, the genesis of the parameters of fuzzy returns will be given due attention. The main goal of the paper is then to present a comprehensive multi-objective fuzzy mathematical programming procedure based on the model accepting the strict and fuzzy data, and other necessary requirements and characteristics shaping the portfolio composition. Special emphasis will be placed on the specification of fuzzy elements, specifically on the determination of parameters of triangular fuzzy numbers describing the return.

The practical part of the paper aims at demonstrating of the application power of the proposed concept based on fuzzy mathematical programming in real decision making in the Czech capital market. The closeness to reality is enhanced not only by a proper reflection of the instability of the most important characteristic of the investment, but also by taking into account the increasingly strong role of environmental or socio-legal impact. Today very popular ESG[2] mutual funds, provided by the largest bank operating on the Czech market, are then considered for investment. The output of the fuzzy optimization process in the form of a portfolio composition appears to be a robust decision support methodology in a sometimes-uncluttered financial environment.

The structure of the article is as follows. After the introduction, in Section 2, the return as a vague element, or fuzzy mathematical model is described. Subsequently, in Section 3, the proposed methodological approach is applied to real portfolio selecting in the capital market with ESG funds. The results are analyzed and summarized from a practical-theoretical perspective. Finally, challenges and ideas for further research are outlined.

2 Return as a vague element

Return as a vague, approximate, element can be effectively expressed, or quantified, using a triangular fuzzy number. This section describes the genesis of the fuzzy expression with subsequent integration into a complex mathematical model reflecting a multi-objective view in investment portfolio making.

2.1 Fuzzy return

Let $\tilde{r}_i, i = 1, 2, .., p,$ denote the fuzzy return of the *i*-th asset expressing by fuzzy number, formalized by the triplet $(r_i^{min}, \overline{r}_i, r_i^{max})$, with the following triangular membership function

$$\mu_{\tilde{r}_i}(r_i) = \begin{cases} 0 & r_i \le r_i^{min} \vee r_i \ge r_i^{max} \\ \dfrac{r_i - r_i^{\min}}{\overline{r}_i - r_i^{\min}} & r_i^{min} < r_i < \overline{r} \\ \dfrac{r_i^{max} - r}{r_i^{max} - \overline{r}_i} & \overline{r} < r_i < r_i^{max} \\ 1 & r = \overline{r}_i \end{cases}, \tag{1}$$

where $r_i, i = 1, 2, ..., p,$ is a possible strict value of return of the *i*-th asset. As mentioned in the introduction, the scientific publications do not deal much with the genesis of the parameters of fuzzy numbers. A notable exception is Vercher (2008), who describes the computation of four parameters of a nonlinear fuzzy number based on statistical parameters (percentiles). This approach is inspiring because it captures the stratification, and hence the position, of the observed characteristic. Deciding on the percentiles used can be a minor challenge. Thus, I propose to simplify the process by using descriptive statistics of the (geometric) mean, and associated outliers. We are thus given essential, complex information about the return at its average (expected) growth rate and the range in which it can move. Then the average $\overline{r}_i$, minimum r_i^{min} and maximum r_i^{max} return of a historical period divided into T subperiods can be formalized as follows

$$\overline{r}_i = \sqrt[T]{\prod_{t=1}^{T} r_{it}}, \quad r_i^{min} = \min_{1 \le t \le T}(r_{it}), \quad r_i^{max} = \max_{1 \le t \le T}(r_{it}) \qquad i = 1, 2, ..., p, \tag{2}$$

where $r_{it}, i = 1, 2, ..., p, t = 1, 2, ..., T,$ is the return of the *i*-th asset in *t*-th time subperiod.

[2] Abbreviation from **E**nvironmental-**S**ocial-**G**overnance

2.2 Fuzzy mathematical modelling procedure

There are many models of fuzzy multi-objective optimization. Some methods, or models, are based on the principle of utility maximization (Marler and Arora, 2010). Others try to minimize the distance from the ideal solution. These techniques can also include fuzzy goal programming (Razmi et al., 2016). Finally, a lexicographic approach (Ojha and Biswal, 2009), or a technique working with aspiration levels in the vague form (Zheng and Brikaa, 2022) could be useful.

Even considering the specificities of a rather more passive investment strategy, I suggest to choose a procedure requiring rather less information from a usually inexperienced investor (decision maker). Techniques based on aspiration or goal levels are very challenging in this respect. Fuzzy approaches, across the aforementioned (e.g. Ghanbari et al., 2020), using α-cut to optimize, are also information intensive, as determining the cut value may not be a simple, but crucial, task. The lexicographic approach is not a suitable evaluation technique for the practical purposes announced due to inadequate discrimination of some criteria. The possible combination of strict and fuzzy input data does not technically favor the application of the principle of utility maximization, or weighted averaging, hence distance measurement, where, moreover, the choice of the appropriate metric is a difficult issue.

Based on the above, I propose an optimization methodology that does not require hard-to-find information, inconvenient decisions, and accepts a specific data base. Furthermore, it will reflect the investor's essential requirements or characteristics of the investment, which may be perceived by the investor with different importance. From a technical-algorithmic point of view, the procedure should be understandable, user-applicable in investment practice without major obstacles. Such an approach is the hitherto unmentioned minimum component maximization approach, which is an essential conceptual aspect of the optimization model below. Let us therefore design the methodological process in the following few steps.

Step 1 Let $f_j(\mathbf{x}) = \mathbf{c}_\mathbf{j}^\mathbf{T}\mathbf{x},\ j \in J_S^{min} \cup J_S^{max}$, or $\tilde{f}_j(\mathbf{x}) = \tilde{\mathbf{c}}_\mathbf{j}^\mathbf{T}\mathbf{x} = \left(f_j^l(\mathbf{x}), f_j^m(\mathbf{x}), f_j^u(\mathbf{x})\right),\ j \in J_F^{min} \cup J_F^{max}$, denote the objective function with strict coefficients $\mathbf{c}_\mathbf{j}^\mathbf{T} = (c_{1j}, c_{2j}, ..., c_{pj})$, or fuzzy coefficients $\tilde{\mathbf{c}}_\mathbf{j}^\mathbf{T} = (\tilde{c}_{1j}, \tilde{c}_{2j}, ..., \tilde{c}_{pj})$, where the partial coefficients are represented by the triangular fuzzy number $\tilde{c}_{ij} = (c_{ij}^l, c_{ij}^m, c_{ij}^u), i = 1, 2, ..., p,\ j \in J_F^{min} \cup J_F^{max}$. $\mathbf{x} = (x_1, x_2, ..., x_p)^T$ represents the vector of *p* variables. The set J_S^{min}, or J_S^{max} includes the indices of minimizing, or maximizing objective functions with the strict coefficients. Similarly, the sets J_F^{min}, or J_F^{max} contains the indices of minimizing, or maximizing objective functions with the fuzzy coefficients. The weight of the *j*-th objective function is labeled as w_j.

Step 2 The basal and ideal values of each objective function are determined. Let the set *X* give all relevant requirements and condition of the problem. Then the strict ideal value $f_j^I,\ j \in J_S^{min} \cup J_S^{max}$, is determined through the one objective model minimizing, or maximizing strict objective function on the set *X*. Similarly, three partial models are solved in the case of fuzzy objective functions in order to gain ideal values of three parameters $\tilde{f}_j^I = (f_j^{Il}, f_j^{Im}, f_j^{Iu}),\ j \in J_F^{min} \cup J_F^{max}$. The basal value of the objective functions, marked as $f_j^B,\ j \in J_S^{min} \cup J_S^{max}$, or $\tilde{f}_j^B = (f_j^{Bl}, f_j^{Bm}, f_j^{Bu}),\ j \in J_F^{min} \cup J_F^{max}$, is not determined by a solitary finding the opposite extreme, but is determined with respect to the optimal values of the other objective functions. The extreme values of the objectives serve not only for the standardization of their values and thus the subsequent simultaneous optimization, but also as valuable information for a decision-maker (investor).

Step 3 The model attempts to simultaneously optimize the values of the considered characteristics (objective functions) under strict conditions forming the set *X*. The fuzzy model using the declared maximin principle is then formulated in the following form.

$$
\begin{aligned}
&\max \quad \alpha \\
&j \in J_S^{min} \quad (1-w_j)\frac{f_j^B - f_j(\mathbf{x})}{f_j^B - f_j^I} \geq \alpha; \quad j \in J_S^{max} \quad (1-w_j)\frac{f_j(\mathbf{x}) - f_j^B}{f_j^I - f_j^B} \geq \alpha, \\
&j \in J_F^{min} \quad (1-w_j)\frac{\tilde{f}_j^B - \tilde{f}_j(\mathbf{x})}{\tilde{f}_j^B - \tilde{f}_j^I} \geq \alpha; \quad j \in J_F^{max} \quad (1-w_j)\frac{\tilde{f}_j(\mathbf{x}) - \tilde{f}_j^B}{\tilde{f}_j^I - \tilde{f}_j^B} \geq \alpha, \\
&\mathbf{x} \in X, \\
&0 \leq \alpha \leq 1.
\end{aligned}
\tag{3}
$$

The positive weights are standardized. Their sum over all criteria is one. They can be determined through different approaches, such as the scoring method. The indicator α represents the weighted minimum characteristic component or the weighted grade of membership of the fuzzy solution. The interval of its values is thus understandable. In order to find the solution of a fuzzy model (3), its transformation must be used. This process can be based on the simple principle of comparing fuzzy numbers using strict partial relations of their parameters. The final strict mathematical model then is specified as

$$\begin{aligned}
&\max \quad \alpha \\
&j \in J_S^{min} \quad (1-w_j)\frac{f_j^B - f_j(\mathbf{x})}{f_j^B - f_j^I} \geq \alpha; \quad j \in J_S^{max} \quad (1-w_j)\frac{f_j(\mathbf{x}) - f_j^B}{f_j^I - f_j^B} \geq \alpha, \\
&j \in J_F^{min} \quad (1-w_j)\frac{f_j^{Bl} - f_j^l(\mathbf{x})}{f_j^{Bl} - f_j^{Il}} \geq \alpha, \quad (1-w_j)\frac{f_j^{Bm} - f_j^m(\mathbf{x})}{f_j^{Bm} - f_j^{Im}} \geq \alpha, \quad (1-w_j)\frac{f_j^{Bu} - f_j^u(\mathbf{x})}{f_j^{Bu} - f_j^{Iu}} \geq \alpha, \\
&j \in J_F^{max} \quad (1-w_j)\frac{f_j^l(\mathbf{x}) - f_j^{Bl}}{f_j^{Il} - f_j^{Bl}} \geq \alpha, \quad (1-w_j)\frac{f_j^m(\mathbf{x}) - f_j^{Bm}}{f_j^{Im} - f_j^{Bm}} \geq \alpha, \quad (1-w_j)\frac{f_j^u(\mathbf{x}) - f_j^{Bu}}{f_j^{Iu} - f_j^{Bu}} \geq \alpha, \\
&\mathbf{x} \in X, \\
&0 \leq \alpha \leq 1.
\end{aligned} \tag{4}$$

In investment practice, the model will most often be in a linear or quadratic form, which determines the choice of a suitable solution finding method, usually through some software. The solution of the model provides, through the values of variables, the resulting shares of assets in the portfolio exhibiting specific characteristics quantified by the values of the objective functions.

3 ESG mutual fund portfolio selection

The empirical part is devoted to demonstrating the applicability of flexibly quantified returns in investment practice. The intention is to cover a broader range of investment cases, hence the choice of a longer-term investment policy reflecting today's increasingly intense drive towards financial security in retirement.

3.1 "Green" longer-time investment strategy

A recent survey conducted by the Czech Capital Market Association shows that approximately three quarters of respondents are looking for a longer-term appreciation of their available financial resources in order to save for retirement age, or maintain an active lifestyle in old age (AKAT, 2024). A long-term strategy with this purpose is usually conceived more conservatively, with a rather passive approach to portfolio management. Therefore, let us imagine one (not necessarily) of the 4.5 million clients of Česká spořitelna – bank with the largest client share on the Czech market. Apart from the key characteristics (return, risk, etc.), it also takes into account an increasingly important criterion, namely the environmental and social impact of the investment. Therefore, the increasingly popular ESG mutual funds, managed by Česká spořitelna, or Erste Asset Management, with sufficient history reflecting a long-term investment view, are preselected. They are listed with all characteristics in Table 1.

Fund name	**Return** [%]			**Risk** [score]	**Cost** [%]
ESG Mix 10	-3.53	0.20	2.51	2	1.13
ESG Mix 30	-4.65	0.22	4.89	2	1.52
ESG Mix 50	-7.50	0.23	5.51	3	2.26
Erste Bond Europe-High Yield	-9.33	0.05	5.52	3	1.17
Erste Bond Euro Corporate	-3.99	-0.10	4.85	3	0.71
Erste Stock Comodities	-14.05	0.77	16.74	4	2.08
Erste Stock Biotec	-12.69	0.34	17.96	5	2.07
Erste Stock Global	-11.56	0.98	10.26	4	2.07
Erste Responsible Stock Global	-11.73	1.04	11.39	4	1.79
Erste Stock EM Global	-20.84	-0.16	11.51	4	2.09
Erste Responsible Stock America	-9.12	1.23	14.50	5	2.04
Erste WWF Stock Environment	-19.29	0.91	18.17	5	1.77

Table 1 ESG funds and their characteristics, Source: Česká spořitelna Investment center (2024), self-calculation

Investing in mutual funds is undeniably shaped by return, risk and fee characteristics. In order to reflect longer-term trend, return is calculated from historical data over a five-year period from January 2019 to January 2024. This period demonstrates both larger falls or rises (e.g. in 2020) and subperiods of calmer development. The return as a fuzzy characteristic is measured through the formulas (2). The risk is to a large extent already expressed by the fuzzy return specification itself. However, in order to make a partial assessment according to this characteristic, the official scale of riskiness $\langle 1,7 \rangle$ can be used. Costs include all fees paid on a continuous basis (audit, licensing, management, etc.). The entry fee is eliminated because of its one-time nature and minimal variability across the selected funds.

The purpose of a longer-term investment, in particular financial security for retirement, can accelerate risk aversion, which is ultimately confirmed by the AKAT (2024) research mentioned above. Investors fear loss, and even fraud, from the issuer or intermediary. This concern should be more fundamentally allayed by selecting mutual funds from a reputable investment company. A user-friendly scoring method is used to quantify the investor's preferences or estimate the weights of the criteria. Although this is a more conservative strategy, capital appreciation is an essential purpose of investing. Return is therefore the most important characteristic, followed by a not inconsiderable aversion to risk, with a relatively conciliatory view of fees. The subjective preference decision on an integer scoring scale $\langle 1,10 \rangle$, with objective weighting, is as follows: return (score 9; weight 0.429), risk (8, 0.381), cost (4; 0.19). Finally, to avoid one-sided or conversely fragmented portfolios, a limitation on the share of one asset in the portfolio is added. In the spirit of diversifying the number of assets, hence investment practice, a minimum of four, and a maximum of six funds will make up the portfolio. The share of one fund will thus be in the range of $\langle 15,25 \rangle$ %.

3.2 Portfolio selecting through the customized model

To select the most appropriate mutual fund portfolio, the following strict model (following (4)), obtained from the original fuzzy version (3), is solved.

$$\begin{aligned}
&\max \quad \alpha \\
&(1-0.429)\frac{\mathbf{r}^{\mathbf{l}^{\mathrm{T}}}\mathbf{x}-(-13.79)}{-4.92-(-13.79)} \geq \alpha, \quad (1-0.429)\frac{\mathbf{r}^{\mathbf{m}^{\mathrm{T}}}\mathbf{x}-0.09}{1.04-0.09} \geq \alpha, \quad (1-0.429)\frac{\mathbf{r}^{\mathbf{u}^{\mathrm{T}}}\mathbf{x}-4.44}{16.84-4.44} \geq \alpha, \\
&(1-0.381)\frac{4.75-\mathbf{d}^{\mathrm{T}}\mathbf{x}}{4.75-2.5} \geq \alpha, \quad (1-0.19)\frac{1.99-\mathbf{c}^{\mathrm{T}}\mathbf{x}}{1.99-1.13} \geq \alpha, \\
&\mathbf{e}^{\mathrm{T}}\mathbf{x} = 1, \\
&0.15\mathbf{y} \leq \mathbf{x} \leq 0.25\mathbf{y}, \\
&0 \leq \alpha \leq 1, \\
&\mathbf{x} \geq \mathbf{0} \\
&\mathbf{y} \in \{0,1\}
\end{aligned} \tag{5}$$

where $\mathbf{x} = (x_1, x_2, ..., x_{12})^T$ is a vector of the variables expressing the share of the mutual funds (in the order listed in Table 1), $\mathbf{y} = (y_1, y_2, ..., y_{12})^T$ is then a vector of binary variables involved in the limitation for the fund share, $\mathbf{e}^{\mathrm{T}} = (1,1,...,1)$. $\mathbf{d}^{\mathrm{T}}\mathbf{x}$, or $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ is a strict function representing risk, or cost of the investment portfolio. Then vector $\mathbf{d}^{\mathrm{T}} = (d_1, d_2, ..., d_{12})$, or $\mathbf{c}^{\mathrm{T}} = (c_1, c_2, ..., c_{12})$ contains the risk, or cost connected with the particular mutual fund. The portfolio return is described by fuzzy objective function $\tilde{r} = (\mathbf{r}^{\mathbf{l}^{\mathrm{T}}}\mathbf{x}, \mathbf{r}^{\mathbf{m}^{\mathrm{T}}}\mathbf{x}, \mathbf{r}^{\mathbf{u}^{\mathrm{T}}}\mathbf{x})$, whose parameters reflect the lower, medium and upper investment return, where $\mathbf{r}^{\mathbf{l}^{\mathrm{T}}} = (r_1^l, r_2^l, ..., r_{12}^l)$, $\mathbf{r}^{\mathbf{m}^{\mathrm{T}}} = (r_1^m, r_2^m, ..., r_{12}^m)$ and $\mathbf{r}^{\mathbf{u}^{\mathrm{T}}} = (r_1^u, r_2^u, ..., r_{12}^u)$ are the vectors including lower, medium and upper returns of particular funds. The ideal and basal values are determined through a declared one-objective optimization process. Model (5) of partially binary linear programming can be easily solved, for example, in Lingo software.

3.3 Result discussion

The optimal solution of model (5) forms the following portfolio composition: 18.35% ESG Mix 30, 17.87% Erste Bond Euro Corporate, 25% Erste Stock Comodities, 18.79% Erste Responsible Stock Global and 19.99% Erste Responsible Stock America. The (fuzzy) return of the portfolio is $(-9.11, 0.66; 10.99)\%$, risk 3.65 points and cost

1.68%. In order to eliminate any misunderstanding of the fuzzy return level, an average characteristic (medium parameter), or a more complex indicator (weighted) averaging the three parameters, may not be presented. Not surprisingly, the portfolio tends to be closer to the basal value in the case of the less important characteristics.

The fund with the largest share is one of the few that has a positive difference between maximum−mean and mean−minimum values of return. This positive asymmetry keeps the fund in the portfolio even with more important return over risk preferences. Only when a relative return importance is substantial, this or other funds are replaced by the Erste Stock Biotec fund, which is dominated by a positive asymmetry with a positive mean return. However, its high riskiness does not fit well within a more conservative investment strategy. Given the considerable significance of the risk, the presence of funds with the lowest score of riskiness was expected. However, ESG Mix 10 does not participate in the portfolio because its negatively asymmetric return is not strong even in its average form. It would only become part of the portfolio assuming a significantly lower return weighting with strong risk aversion. Far better in this respect is ESG Mix 30, which shows solid, an almost symmetric, triangular return. In contrast, the Erste Stock EM Global fund, with its significantly negatively skewed (extreme) return, can hardly make its way into a portfolio of various preferences. On the other hand, another stock fund with the highest level of risk participates in the portfolio, namely Erste Responsible Stock America, which, while having a slightly weaker positive triangular return asymmetry, exhibits a much larger, even the largest, average characteristic value.

It certainly begs the question of whether the fuzzy measure of return fundamentally affects the outcome. So, let's measure the return strictly, using only the average value (medium parameter of fuzzy number). Maintaining the idea of strict evaluative optimization, the portfolio will be as follows: 18.88% ESG Mix 10, 16.12% ESG Mix 30, 15% Erste Bond Europe-High Yield, 25% Erste Responsible Stock Global and 25% Erste Responsible Stock America. At first glance, the composition of the portfolio is quite different. Not surprisingly, with the importance of return and risk criteria almost equal, the least risky funds ESG Mix 10 and 30 are more prominent. When outliers are eliminated, funds with negatively asymmetric return funds with its decent mean, such as Erste Responsible Stock Global, can play a significant role in the portfolio. In contrast, Erste Bond Euro Corporate, with its negative average return, has no chance to participate in the investment. Only the extreme values show a positive skewness of its return, which secures it a place in the former portfolio.

4 Conclusion

This paper highlights the role of return as a vague element in the portfolio selection process. The return is conceived as a triangular fuzzy number with uniquely calculated parameters. It is then integrated into the proposed fuzzy multi-objective mathematical model, as the basis of a comprehensive investment portfolio selecting procedure. Investment practice with mutual funds then demonstrates the power and usefulness of integrating fuzzy information, especially when possibly combined with strict input data. Considering not only the average return, but especially its outliers, allows for a more convenient description of the characteristic over historical time, especially its dynamic structure potentially effected by factors of different nature. Such an apparatus can more successfully describe investment reality, potentially leading to more representative decisions. The various composition of the ESG funds' portfolios, reflecting only strict, or in combination with fuzzy, returns, provides supporting evidence.

Since many publications do not deal much with the actual determination of fuzzy return parameters, this issue deserves an even deeper theoretical-application study. The fuzzy model could be extended with the concept of fuzzy relations, which would broaden its application potential. It could then reflect the investor's approximate preferences, hence the interactive process of portfolio revision leading to investor satisfaction. Fuzzy relations are actually related to the technique of comparing fuzzy numbers, which would also deserve a deeper analysis.

Acknowledgements

The research project was supported by Grant No. F4/18/2024 of the Internal Grant Agency, Faculty of Informatics and Statistics, Prague University of Economics and Business.

References

AKAT (2024) Největší chyba v investování? Neinvestovat! Available at: https://www.akatcr.cz/Dokumenty/Aktuality/tiskov225-zpr225va-akat-nejvets205-chyba-v-investov193n205-neinvestovat [cited 2024-05-16]

Česká spořitelna Investment center (2024) Funds. Available at: https://cz.products.erstegroup.com/Retail/en/index.phtml [cited 2024-05-16]

Duan L, Stahlecker P (2011) A Portfolio Selection Model Using Fuzzy Returns. Fuzzy Optimization and Decision Making 10: 167–191. https://doi.org/10.1007/s10700-011-9101-x

Ghanbari R, Ghorbani-Moghadam K, Mahdavi-Amiri N, De Baets B (2020) Fuzzy Linear Programming Problems: Models and Solutions. *Soft Computing* 24: 10043–10073. https://doi.org/10.1007/s00500-019-04519-w

Mandal P K, Thakur M, Mittal G (2024) Credibilistic Portfolio Optimization with Higher-Order Moments Using Coherent Triangular Fuzzy Numbers. Applied Soft Computing 151: 111155. https://doi.org/10.1016/j.eswa.2019.07.027

Marler R T, Arora J S (2010) The Weighted Sum Method for Multi-Objective Optimization: New Insights. Structural and Multidisciplinary Optimization 41: 853–862. https://doi.org/10.1007/s00158-009-0460-7

Mashayekhi Z, Omrani H (2016) An Integrated Multi-Objective Markowitz–DEA Cross-Efficiency Model with Fuzzy Returns for Portfolio Selection Problem. Applied Soft Computing 38: 1–9. https://doi.org/10.1016/j.asoc.2015.09.018

Ministry of Finance of the Czech Republic (2024) Report on Financial Market Developments in 2022. Available at: https://www.mfcr.cz/en/regulation-and-taxes/financial-market-analyses/report-on-financial-market-developments-in-2022-52232 [cited 2024-05-16]

Li X, Zhang Y, Wong H-S, Qin Z (2009) A Hybrid Intelligent Algorithm for Portfolio Selection Problem with Fuzzy Returns. Journal of Computational and Applied Mathematics 233: 264–278. https://doi.org/10.1016/j.cam.2009.07.019

Ojha A K, Biswal K K (2009) Lexicographic Multi-Objective Geometric Programming Problems. International Journal of Computer Science Issues 6: 20–24. https://doi.org/10.48550/arXiv.0912.1832

Razmi J, Jafarian E, Amin S H (2016) An Intuitionistic Fuzzy Goal Programming Approach for Finding Pareto-Optimal Solutions to Multi-Objective Programming Problems. Expert Systems with Applications 65: 181–193. https://doi.org/10.1016/j.eswa.2016.08.048

Qin Z, Wen M., Gu C (2011) Mean-Absolute Deviation Portfolio Selection Model with Fuzzy Returns. Iranian Journal of Fuzzy Systems 8: 61-75. https://doi.org/10.22111/IJFS.2011.308

Vercher E (2008) Portfolios with Fuzzy Returns: Selection Strategies Based on Semi-Infinite Programming. Journal of Computational and Applied Mathematics 217: 381–393. https://doi.org/10.1016/j.cam.2007.02.017

Review on assortment selection problem integrated with capacitated multi-product order optimization

Tomasz Brzęczek[1]

Abstract. Assortment planning belongs to operations research classical problems. The optimal order for a random demand product is solved with the newsvendor model application. However, the optimal order of assortment maximizes expected profit subject to total budget or other common constraint for products. This research provides literature review on model extensions and solving techniques for the multi-product problem. Economic aspects prevailing in research are specified. General assortment planning research statistics checks integration with other techniques.

Keywords: assortment planning, capacity constraint, newsvendor, stochastic demand

JEL Classification: C44
AMS Classification: 90C15

1 Literature review (Automatically Numbered)

Classically newsvendor problem called hereafter NV was inventory single-period problem (SPP). (Khouja, 1999) presented review of 92 research and classified them into 11 extension categories. The multi-product problem is one category that was concerned in 5 references. These were focused on budget constraint optimal solving procedures especially Lagrangian multiplier iterative solving or alternatively dual problem with much smaller number of variables and shorter running time. Other extension category contained 7 references concerning the multi-product problem with substitution. NV's total expected profit increases due to substitution because stock shortage is compensated with overstocked substitute. Khouja has been cited in 270 following papers for 25 years. More and more visibility was gained by multi-product problem (MPNP). To check extensions subsequent development and application popularity we perform statistics of 90 papers citing (Lau & Hing-Ling Lau, 1995) that were among first to cover multiple resources constraints. To cross-validate whether industry application studies introduce theoretical model extensions we extended citing references by 59 papers referring to NV research and industrial application (Rajaram, 2001) as of May 2024 was indicated. No duplicate were found in those two groups of references. Declared key words and titles were checked to classify type of application context into categories like: assortment planning, product-mix inventory management, demand analysis covering for sales forecasting.

Recent reviews of the literature treated about price-dependent demand and supply chain relations aspects (Qin et al., 2011), and multi-product extensions (Mu et al., 2019; Turken et al., 2012). We try to identify potential of MPNP extensions to constraint versus to objective function. We found also papers extending both simultaneously like (Amaruchkul, 2021) that solves labor supplydemand transportation constraints to minimize cost of dynamic random yield of sugarcane.. We refer to models concerning some typical management applications like the case of company budget constraint with multi-product marketing effort of a company. To assess the integration of theoretical extensions research into economics theories and applications we did literature statistics.

Reminder of this paper is following. Section 1 starts with literature review subsections: NV formulation, typical constraints of MPNP, objective functions adapted from single-product to MPNP by aggregation, demand parametrization in MPNP. Section 2 provides statistical analysis in both literature groups that are theoretical versus applied research references. We conclude with findings and some points for future research.

1.1 NV problem

The following notation is used (Khouja, 1999):

- x quantity demanded, a random variable with probability density function $f(x)$ and cumulative distribution function $F(x)$;
- Q is order quantity for which optimality is denoted with Q^*;
- P is regular selling price;

[1] Poznan University of Technology/Logistics Department, J. Rychlewskiego street 2, tomasz.brzeczek@put.poznan.pl

- C is unit cost;
- S is underage unit penalty cost, if it doesn't apply $S = 0$;
- V is salvage value for overage unit;
- $i = 1,2, \dots, m$ is product index.

The classical formula for expected profit found in the literature is (Khouja, 1999):

$$E(\pi) = (P + S - C)\int_{Q}^{\infty} Qf(x)dx \ - S\int_{Q}^{\infty} xf(x)dx + (P - V)\int_{0}^{Q} xf(x)dx \\ -(C - V)\int_{0}^{Q} Qf(x)dx \tag{1}$$

where $E(\cdot)$ is the expectation operator. Let's compare it with formulation from (Qin et al., 2011) that we denote consistently with notation of formula (1):

$$E(\pi) = \int_{0}^{Q} [Px - CQ + S(Q - x)]f(x)dx + \int_{Q}^{\infty} [PQ - CQ - S(x - Q)]f(x)dx \\ = (P - V)\mu - (C - V)Q - (P - V + S)ES(Q) \tag{2}$$

where μ is the mean demand and $ES(Q)$ represents the expected shortage units assuming Q units are stocked and can be determined as $\int_{Q}^{\infty}[(x - Q)]f(x)dx$.

Subject to budget capacity and other constraints that have many variants concerning applications discussed further.

1.2 Constraints

The simplest MPNP just contains the optimal orders of all products calculated independently. Each product order quantity has to meet a constraint by itself. The best example is service level constraint used by (Urban, 2000) for each product. However, his model that has no common all-product constraint covers for each product dynamic multi-period demand. Details are presented in Table 1.

Constraint type	Formulation	Where	Reference
Replenishment cost budget	$\sum_{i=1}^{m} C_i Q'_i + \sum_{i=1}^{m} \theta_i K'_i \le B$	Q'_i – replenishment K'_i – fixed replenishment cost B – budget limit θ'_i – replenishment dummy	(Moon & Silver, 2000)
Service level	$\sum_{t=t_1}^{t_2-1} F_t(Q_t) = \tau\varphi$ $\varphi = \frac{P_i - C_i + S_i}{P_i + V_i + S_i}$	φ – static service level $\tau = t_2 - t_1$ – number of periods under consideration Negative C_i follows from two options of defining V: $C_i < V_i < 0$ salvage value $V_i > 0$ – disposal cost	(Urban, 2000) considers single product model that is easy extendable to multi-product problem
Disposable budget	$\sum_{i=1}^{m} C_i Q_i + \sum_{i=1}^{m} \beta_i K_i \le \propto B_A$ $B_A = \sum_{i=1}^{m} C_i Q_i^*$	B_A – Budget that maximizes all products expected profit $\propto \epsilon < 0, 1)$ – disposable $\beta_i = 1,\ for\ Q_i > 0$ K_i – fixed cost of product i	(Rajaram, 2001)
Supplier quantity discount	$\sum_{i=1}^{m} \sum_{j=1}^{k_i} C_{ij}\, Q_{ij}$ $= order\ cost$	$d_{ij}^{L} < Q_i < d_{ij}^{U}$ – discount segment j in which is solved product i order quantity $y_{ij} = 1$ if product is purchased at price segment j	(G. Zhang, 2010) for income integrals calculus stay Q_i while costs cover for Q_{ij} at cost j
Transportation of workers under dynamic yields	$\sum_{i=1}^{I} x_{ij} = w_j$ $\sum_{j=1}^{J} x_{ij} \le s_i$	x_{ij} – workers to go from source i (its supply is s_i) to destination j (demand is w_j)	(Amaruchkul, 2021) solves labor supply to 46 provinces farms for two products

forecast and demand for workers			expected profit maximization

Table 1 Constraint categories in MPNP extensions

Constraint formulated to cover jointly for products' order quantities satisfies limit of common resource.For example all-product order spendings are limited with budget amount. Also limited space is allocated for products. Weight in kilograms is another typical limitto assortment ordered Main method used in the literature was Lagrangian based approach with standard expected profit function being summed for multi-product analysis under one common constraint. (Lau & Hing-Ling Lau, 1995) were first to report an issue of such approach. Namely, negative solution for at least one order under tight constraint of resource. (L. L. Abdel-Malek & Montanari, 2005a); (L. L. Abdel-Malek & Montanari, 2005b) developed dual solution space procedure for two tight constraints problem with large number of products with order quantity lower bound of 0 and shows the lower bounds of the minimum amount of the resources required to include each of the contending products in the solution. This iterative solving technique workhas been developed for other probability distributions in subsequent work of (L. Abdel-Malek et al., 2020).
Challenging is also extension of budget constraint as well as expected profit function with fixed costs and dummies of product selection (Amaruchkul, 2021);(Moon & Silver, 2000);(Rajaram, 2001). All of them proposed heuristic procedures to receive a sub-optimal solution. Performance of the solution is assessed in percentage of upperbound of the maximum profit solution solved from Lagrangian dual problem for dummies set to 0 according to product profitability heuristic results or negative order quantity for first iteration solution.

1.3 Objective

If constraints are complex to solve then usually expected profit standard formulations (1-2) are used. These and other formulations of objective function are set in Table 2.

Objective[1]	**Reference**	**Multi-product orders constraint**	**Dynamic (multi-period)**
Target ROI probability	(Thakkar R.B. et al., 1983)	0	0
$E(\pi)$ – expected profit/mismatch cost	(Moon & Silver, 2000)	1	1
The cost of periodical adjustment of order quantity $w_t = \begin{cases} w, & Q_t \neq Q_{t-1} \\ 0, & Q_t = Q_{t-1} \end{cases}$	(Urban, 2000) solves single product supply to dynamic demand with network programming. Introduces multiproduct	1	1
Minimize risk of profit	(Vaagen et al., 2008)	1	1
Utility function	(Wang & Webster, 2009)	0	0
The probability of a target profit	(Shi et al., 2010)	1	0
Profit = Revenue – Labor cost – Placement cost	(Amaruchkul, 2021) plans labor for harvesting random yield. $\psi -$ minimum placement cost of supply source i to province j	1	1
Cost-volume-profit	(Kamrad et al., 2021)	1	0

Table 2 Objective function formulations

[1] For more extensions to single product NV objective proposed in 80's and 90's see (Khouja, 1999). We focus on main extension types which objective functions were adapted to multi-product problem or dynamic multi-period demand. Moreover, "1" identifies models and solving procedures that cover for common constraint to order quantities like capacity or budget constraint.

1.4 Parametrization

The most popular as well as practical issue is the demand distribution parametrization in MPNP. Sales are often modelled with joint distribution of demands or with substituteleftover aggregation with other substitute shortage. Therefore, we differ between multi-product models with:

- theoretical demand distribution for example: normal, Poisson, exponential, binomial,
- so called distribution-free that is the worst-case demand,
- data driven forecast of demand,

- demands substitution relation model.

Reference	Distribution	Single/multiple product
(Alfares & Elmorra, 2005)	Mean-variance, the worst case demand	Single. Resource constrained multi-product
(Vaagen et al., 2008)	2-point equally likely with substitution	Sport apparel product variety
(Huber et al., 2019)	Empirical with machine learning	Bakery assortment
(Amaruchkul, 2021)	Yield distribution mean and standard deviation predictions from historical data smoothing models updated with image-based yield prediction technology.	Green or burnt sugar cane
(Urban, 2021)	Location scale demand distribution	Single product profit values range for order quantity decision

Table 3 Parametrization. Particularly demand parametrization issue

More and more visibility receive data mining or machine learning approaches to demand forecasting based on sales data. Surprisingly, (Huber et al., 2019) find the profit advantage of this approaches applied to big data sales set in more precise forecasting demand for each product separately and subsequently solving classical or percentile NV problem. For a benchmark they examined supply order optimization integrated with profit model derived from empirical sales data mining. (Hrabec et al., 2023) thoroughly analyze demand, pricing and marketing effort in NV.

2 Statistics of NV citing assortment planning research

We checked source titles frequency for 90 research publications that referred to (Lau & Hing-Ling Lau, 1995) so introducing rather theoretical work on SPP extension to multi-product formulations and compared with 59 research publications referred to (Rajaram, 2001) applied research study. No duplicated publication found in sets. The objective was to cross-reference theoretical and applied research studies. Listing of source titles publication frequency is provided in Figure 1. Big variety of 104 source titles confirms that theoretical and practical aspects of the problem are widespread across economic research literature. This number of publications goes beyond the group of operations research focused journals. The highest bar at Figure 1 stands for 95 papers publicated in 70 source titles each frequencie being at most two publications from the analyzed sample. Among source titles were hybrid journals, electronic journals, conference proceedings and monographs.

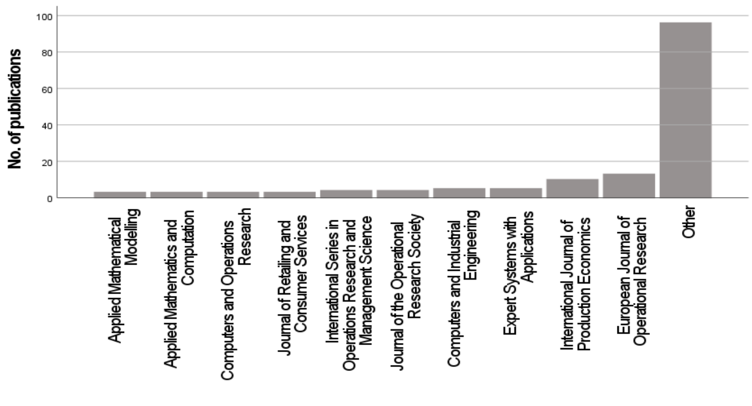


Figure 1 Source title distribution among 149 research population that referred to NV

Following statistics analysis we classified analyzed research population into main focus on theoretical model extensions, applied studies or alternative methodology with results comparison to NV.

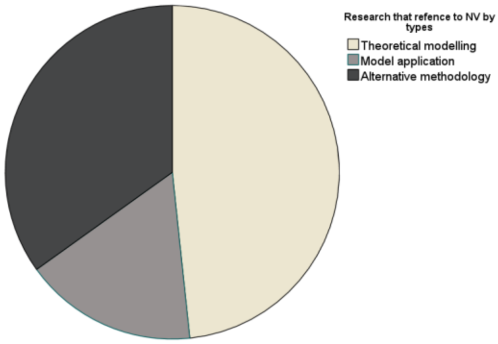


Figure 2 Classification of the main focus in the research 1997-2024 population that referred to NV theoretical research and applied studies

We classified multi-product extension of NV with not standard theoretical distribution into alternative research. To present variety of alternative research approaches Table 4 is provided with a number of publications from the population.

Asortment selection. Pricing	2	Behavioral aspects of supply decision making	3
Assortment selection. Column generation	1	Control theory. Inventory	1
Assortment selection. Data mining	3	Demand analysis	14
Assortment selection. Marketing theory	2	Digital technologies	6
Assortment selection. Portfolio theory	1	EOQ for inventory with stochastic demand	2
Assortment selection. Substitution	4	Fuzzy numbers demand or supply	3
Assortment selection.Data mining	7	Revenue management	1
Assortment selection.Location choice model	1	Supply chain coordination	2

Table 4 Detailed classification of 53 publications presenting alternative methods for NV similar business or operations management problems

Some sub-classes from Table 4 looks to be rare and they really are. We mean modern portfolio theory application on which we found outside the population a few papers only that are reviewed in(B. Zhang & Hua, 2010).

Other like behavioral aspects of supply decision making seems to be rare while we found to be quite popular (Benzion et al., 2010). Behavioral aspects back also target profit literature (G. Zhang & Shi, 2010) and alternative methods of order quantity decision simplification (Urban, 2021).

Important methodology being increasingly incorporated into NV with SPP formulation is game theory and collaborative decision-making in supply chain due to new digital technologies popularization like blockchain or omni-channel. These are difficult to introduce into MPNP as price and cost become variables constrained with assumed demand relation under centralization or contract based mechanism of supply chain coordination

Acknowledgements

Poznan University of Technology, Poland: scientific grant 0812/SBAD/4232

References

Abdel-Malek, L. L., & Montanari, R. (2005a). An analysis of the multi-product newsboy problem with a budget constraint. *International Journal of Production Economics*, *97*(3), 296–307. https://doi.org/10.1016/j.ijpe.2004.08.008

Abdel-Malek, L. L., & Montanari, R. (2005b). On the multi-product newsboy problem with two constraints. *Computers and Operations Research*, *32*(8), 2095–2116. https://doi.org/10.1016/j.cor.2004.02.002

Abdel-Malek, L., Shan, P., & Montanari, R. (2020). A Constructive Methodology to Solving the Capacitated Newsvendor Problem: an Approximate Approach. *Operations Research Forum*, *1*(2). https://doi.org/10.1007/s43069-020-0008-7

Alfares, H. K., & Elmorra, H. H. (2005). The distribution-free newsboy problem: Extensions to the shortage

penalty case. *International Journal of Production Economics*, *93–94*, 465–477. https://doi.org/https://doi.org/10.1016/j.ijpe.2004.06.043
Amaruchkul, K. (2021). Planning migrant labor for green sugarcane harvest: A stochastic logistics model with dynamic yield prediction. *Computers & Industrial Engineering*, *154*, 107016. https://doi.org/10.1016/j.cie.2020.107016
Benzion, U., Cohen, Y., & Shavit, T. (2010). The newsvendor problem with unknown distribution. *The Journal of the Operational Research Society*, *61*(6), 1022–1031. http://www.jstor.org/stable/40608274
Hrabec, D., Kučera, J., & Martinek, P. (2023). Marketing effort within the newsvendor problem framework: A systematic review and extensions of demand-effort and cost-effort formulations. *International Journal of Production Economics*, *257*, 108754. https://doi.org/https://doi.org/10.1016/j.ijpe.2022.108754
Huber, J., Müller, S., Fleischmann, M., & Stuckenschmidt, H. (2019). A data-driven newsvendor problem: From data to decision. *European Journal of Operational Research*, *278*(3), 904–915. https://doi.org/https://doi.org/10.1016/j.ejor.2019.04.043
Kamrad, B., Ord, K., & Schmidt, G. M. (2021). Maximizing the probability of realizing profit targets versus maximizing expected profits: A reconciliation to resolve an agency problem. *International Journal of Production Economics*, *238*, 108154. https://doi.org/https://doi.org/10.1016/j.ijpe.2021.108154
Khouja, M. (1999). The single-period (news-vendor) problem: Literature review and suggestions for future research. *Omega*, *27*(5), 537–553. https://doi.org/10.1016/S0305-0483(99)00017-1
Lau, H.-S., & Hing-Ling Lau, A. (1995). The multi-product multi-constraint newsboy problem: Applications, formulation and solution. *Journal of Operations Management*, *13*(2), 153–162. https://doi.org/https://doi.org/10.1016/0272-6963(95)00019-O
Moon, I., & Silver, E. A. (2000). The multi-item newsvendor problem with a budget constraint and fixed ordering costs. *Journal of the Operational Research Society*, *51*(5), 602–608. https://doi.org/10.1057/palgrave.jors.2600938
Mu, M., Chen, J., Yang, Y., & Guo, J. (2019). The Multi-product Newsvendor Problem: Review and Extensions. *BESC 2019 - 6th International Conference on Behavioral, Economic and Socio-Cultural Computing, Proceedings*. https://doi.org/10.1109/BESC48373.2019.8962977
Qin, Y., Wang, R., Vakharia, A. J., Chen, Y., & Seref, M. M. H. (2011). The newsvendor problem: Review and directions for future research. *European Journal of Operational Research*, *213*(2), 361–374. https://doi.org/https://doi.org/10.1016/j.ejor.2010.11.024
Rajaram, K. (2001). Assortment planning in fashion retailing: Methodology, application and analysis. *European Journal of Operational Research*, *129*(1). https://doi.org/10.1016/S0377-2217(99)00406-3
Shi, C. V., Zhao, X., & Xia, Y. (2010). The setting of profit targets for target oriented divisions. *European Journal of Operational Research*, *206*(1), 86–92. https://doi.org/https://doi.org/10.1016/j.ejor.2010.01.047
Thakkar R.B., Finley D.R., & Liao W.M. (1983). A stochastic demand CVP model with return on investment criterion. *Contemporary Account Research*, *1*, 77–86.
Turken, N., Tan, Y., Vakharia, A. J., Wang, L., Wang, R., & Yenipazarli, A. (2012). The multi-product newsvendor problem: Review, extensions, and directions for future research. *International Series in Operations Research and Management Science*, *176*, 3–39. https://doi.org/10.1007/978-1-4614-3600-3_1
Urban, T. L. (2000). Supply contracts with periodic, stationary commitment. *Production and Operations Management*, *9*(4), 400–413. https://doi.org/10.1111/j.1937-5956.2000.tb00466.x
Urban, T. L. (2021). Visualising newsvendor profits: the single-period problem with location-scale demand distributions. *Journal of the Operational Research Society*, *72*(1), 23–34. https://econpapers.repec.org/RePEc:taf:tjorxx:v:72:y:2021:i:1:p:23-34
Vaagen, H., Wallace, S. W., & Kaut, M. (2008). Assortment Planning in Fashion Supply Chains - Substitutable Newsvendor Problems. *86Th Textile Institute World Conference, Vol 1, Conference Proceedings*.
Wang, C. X., & Webster, S. (2009). The loss-averse newsvendor problem. *Omega*, *37*(1), 93–105. https://doi.org/https://doi.org/10.1016/j.omega.2006.08.003
Zhang, B., & Hua, Z. (2010). A portfolio approach to multi-product newsboy problem with budget constraint. *Computers & Industrial Engineering*, *58*(4), 759–765. https://doi.org/https://doi.org/10.1016/j.cie.2010.02.007
Zhang, G. (2010). The multi-product newsboy problem with supplier quantity discounts and a budget constraint. *European Journal of Operational Research*, *206*(2), 350–360. https://doi.org/10.1016/j.ejor.2010.02.038
Zhang, G., & Shi, J. (2010). The newsvendor pricing problem with supplier discounts. *2010 7th International Conference on Service Systems and Service Management, Proceedings of ICSSSM' 10*, 771–774. https://doi.org/10.1109/ICSSSM.2010.5530140

Analysis of strategy for the Secretary problem with cardinal function based on job specifications

Petr Chládek[1], Marika Hrubešová[2], Štěpán Mudra[3], Martin Polívka[4], Tomáš G. Roskovec[5]

Abstract. The classical Secretary problem is an applied mathematical model for choosing the best applicant from a sample of a given size; the decider can not return to an already rejected candidate and does not a priori know the scale of quality of the candidates. In this well-known problem with a given number of data with unknown distributions, we should stop the search once we consider a number to be the highest one. We opt for the original motivation, but rather than attempting to choose the best one and fail in most cases, we suggest a strategy to control the mean value of the chosen one. The utility of the candidates in different experiments is designed based on the theoretic distribution for specific job titles; we use a cardinal function with the argument being the candidate's percentile among the population. Based on the experiment, we present strategies for hiring for different roles and reveal that the strategy has to be specific for a job; otherwise, it could be suboptimal.
Our strategies are based on two parameters: the size of the sample part we examine and reject at the beginning of the search and the percentile from the examined sample we demand the candidate to outperform to be accepted. We analysed all possibilities for ordering 12 candidates and evaluated the strategies based on the experiment.

Keywords: Secretary problem, hiring strategy, cardinal function, decision making in management

JEL Classification: C44, M51
AMS Classification: 90B50, 62C05

1 Introduction

This paper aims to design and recommend an algorithm or advice for the problem of the choice of hire among several candidates; at the same time, we can not return to a candidate we refused before, and we know the number of candidates. Such a model is both realistic and very close to a mathematical problem called the Secretary problem. In the original question, our goal is to hire the best candidate, and any other is considered an unsuccessful hire. We aim for a more realistic approach as we do not require only the best candidate. In our analysis, we wish to pick somebody of high quality in general, but not needlessly the highest, where the evaluation of the quality is based on the modeling cardinal function, and its argument is the percentile of the candidate among the population (or sample). Such a setting with the cardinal value of each candidate was already partially covered in (Angelovski and Güth, 2020), but we wish to set our model even closer to the application.

In our paper, we use a computer program to answer the following problem: Considering the value of the candidate being given by the percentile and given cardinal function, what is the best size of the sample we reject, and what is the percentile from the sample we should require among the rest of candidates to reach the best mean value of picked candidate? We run all possible permutations on the size of the population $n = 12$ and give the answer for six different admissible cardinal functions simulating hiring for different jobs. The results reveal that several strategies are suitable for particular distribution but can be improved for others.

1.1 Secretary problem

The classical Secretary problem can be formulated as follows: We get a random sample of n numbers in a random order. We do not know anything about the distribution or range of the numbers, but we wish to pick the highest.

[1] University of South Bohemia in České Budějovice/Faculty of Agriculture and Technology, Studentská 1668, 370 05 České Budějovice, Czech Republic, chladek@jcu.cz
[2] University of South Bohemia in České Budějovice/Faculty of Education, Department of Mathematics, Jeronýmova 10, CZ - 371 15, České Budějovice, Czech Republic, mhrubesova@jcu.cz
[3] University of South Bohemia in České Budějovice/Centre of Information Technologies/ Branišovská 31a, CZ - 371 15, České Budějovice, smudra@jcu.cz
[4] University of West Bohemia/Faculty of Economics, Department of Finance and Accounting, Univerzitní 22, 30614 Plzeň, Czech Republic, polivkam@fek.zcu.cz
[5] University of South Bohemia in České Budějovice/Faculty of Education, Department of Mathematics, Jeronýmova 10, CZ - 371 15, České Budějovice, Czech Republic, troskovec@jcu.cz

Once we see a number, we must accept it as our final pick or reject it and move to the next one. Such a problem can be described as hiring a new secretary. Once the candidate is tried and we know his/her quality, we can either accept him/her and end the hiring process or reject him/her and try the next candidate. Even in a naive way, it is clear that the best idea is to try several candidates and then, based on experience, estimate the threshold value we would expect and try to find one satisfying the threshold among the rest.

The solution is well known, and we would not explain it in detail; for n the number of candidates and e the Euler constant, the best way is to find the values of the first n/e candidates and then pick the first value above the numbers we have seen. The result of the method is getting the best one in 37% of cases, regardless of the n. This is both shocking and highly interesting.

The classical Secretary problem, including the historical context, can be found in (Christian and Griffiths, 2016). Furthermore, there are countless variations of the Secretary problem, repeated selection in a fixed group of objects (Goldstein et al, 2019), selection of k candidates from an n element set (Albers and Ladewig, 2021), analysis of optimal algorithms for the Secretary problem using linear programming (Dütting et al, 2021), and others.

1.2 Value of candidates – literature review

Although the widely accepted definition of "talent" in the world of work is still missing, this word is, in general, connected with the high performance of the employee, either present or potential (compare Gallardo-Gallardo et al, 2013). In 1998, the renowned global consultancy group McKinsey & Company introduced the term "War for Talent," declaring the superior employees to be tomorrow's prime source of competitive advantage of a company, warning at the same time that the supply of such employees is going to shrink (Chambers et al, 1998). This assumption quickly spread throughout both the business practice and the academic discussion (O'Mahoney and Sturdy, 2016), contributing to the emergence and rapid rise of a discipline called "talent management" (Swailes, 2016; Ansar and Baloch, 2018). As the "War for Talent" is still raging, hiring the best talents remains a crucial issue for companies. (Piva and Stroe, 2023; Obukhova and Tian, 2024).

The aforementioned statement does not imply that the unconditional effort to select the best candidate from all possible applicants must always be a wise strategy, though, as it is worth bearing in mind the following points. Firstly, the selection process consumes a considerable amount of precious time not only for the recruiters but also for the candidates' future supervisors and department leaders (Navarra, 2022). Secondly, despite the frequently-occurring tendency to understand talent as an innate characteristic, experience from the business practice shows that the majority of performance is rather a result of learning and training (Ericsson et al, 2007) – according to the professionals of Toyota Group only 10% or even less of employees' talent can be attributed to her native-born gifts, while the rest is based on the effort and practice (Liker and Meier, 2007). Last but not least, the original McKinsey & Co. paper referred to the senior executives, who have a crucial influence on the company performance, not to every particular position in a company (compare Chambers et al, 1998).

Based on these observations, we dare to conclude that it is not always necessary to insist on selecting the best applicant for every particular job. For the less exposed job, it can be perfectly reasonable to hire an applicant who fits the fundamental requirements of the position and subsequently focus on her training and development. Such an approach will still result in hiring an employee appropriate for such a job, with concurrent savings in recruitment costs. Unconditional pursuit of selection of the best applicant can, therefore, be reserved only for the key expert and executive positions, where even the innate 10% of talent can play a decisive role in the performance of the organization.

We should obviously evaluate what it means to be talented for a given job. In general, two prominent values for white-collar jobs are analytical thinking and social intelligence. There are contradictory stances on what should be the priority. A study by (Lathesh and Avadhani, 2018) claims that people with high social intelligence (SI) are faster to learn a new skill and perform better in general. Although there is no universal scale for SI, there are professional tests to be used there (Frankovský and Birknerová, 2014). On the other hand, IQ as a standard measurement of analytical thinking is still prominent in the hiring process; in the literature (Armstrong and Taylor, 2020), the IQ test is highly recommended as a part of the evaluation of the candidate, increasing the probability of picking a highly skilled worker.

As our work does not focus on measuring talent but on decision-making based on given measurements, we would assume the hiring subject is using some one-dimensional scale and value the candidate's potential based on the percentile position on the scale among the population. But the function mapping the percentile to value can vary based on the field and the position, as we explain and describe in the following Section 2.1.

2 Methods and model design

2.1 Utility function design

We design a function to represent the situation in the workforce market, where we attempt to model the value of possible candidates based on their skill/talent among the population. Let us note that all models are one-dimensional to order the candidates. In application, picking the best candidate is naturally a multi-criteria problem, but let us assume that the committee is able to order the candidates by performance.

Our models are based on strictly increasing continuous functions but are discretized into the model with 12 candidates. Note that all models are chosen in a way that the best has a value of 1 and the worst 0. We add an insignificant increment to the locally constant parts to fulfill the assumption of monotonicity but not to influence the outcome. Let us compare the continuous and discrete values in Figure 1.

Classical Secretary problem

The first model is the classical version of the Secretary problem, where the utility is binary, 1 for the best candidate and 0 for all others; we denote such a function f_{cs}. Such a model can associated with the search for an exceptional person, where all others are not useful for such a position at all. This can be seen as a limiting case for models f_{pa}, f_{no} or f_{un} explained later. Such a model can be useful for searching for a crisis manager, the lead of a scientific lab, or a key member of a professional sports team where the requirements are extreme. However, let us note that even in this position, ignoring the value of even the second-best candidate is probably a rough estimate. In reality, other models would be more reasonable.

Order distribution

In the second model, let us give the candidates points based on their order from 0 to 11 and norm them. We denote such a function f_{or}. Such a model is primitive and should be used only in cases where we can not estimate the performance, for example, hiring for a new position that has not existed before, so there is no reasoning for using other, more specialized models.

Normal distribution

The third model is based on the normal distribution, where 12 candidates are placed by the normal distribution of IQ among the population; the value is then normed, and we denote such a function f_{no}. Such a model represents a specialist using one key skill. This may be a handyman, repairman, accountant, or IT guy. One essential skill is needed, and other aspects are marginal.

Double-normal distribution

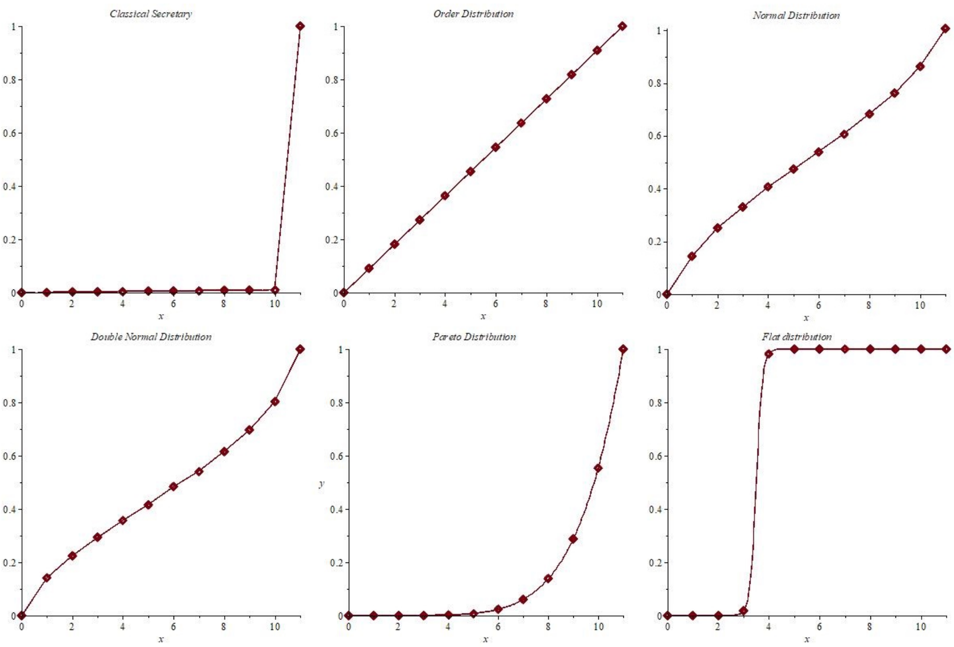
The fourth model is based on the product of two normal distributions, where 12 candidates are placed by the product of two noncorrelated normal distributions similar to IQ distribution among the population (for simplicity, we assume SI has the same distribution as IQ); values are then normed, and we denote such a function f_{un}. Such a model represents possible team leaders and difference makers, as illustrated by the Venn diagram, with rather a small intersection of great intelligence and social intelligence (often combined with another skill specific to a given industry). Such a candidate is sometimes called a "unicorn" for its rarity and is sometimes considered unreal. Note that the product between values of IQ and SI represents that weakness in either aspect makes a candidate not very beneficial, and only people strong in both aspects belong to such a position. This model is somehow an interpolation between f_{cs} and f_{no}, as the distribution is similar to normal, but we strongly prefer the best candidate.

Pareto distribution

The fifth model is based on the Pareto rule, where 80% of the population makes 20% of value and 20% of the population makes 80% of value; we simulate such growth by $f_{pa} = 0.000000339x^{6.2126}$, where $x \in [0, 12]$ represents the position of the candidate in population. Such a model represents creative positions where such a distribution can be observed, such as professional sportsmen, artists, or scientists. Note that this model is very close to classical Secretary f_{cs}, see Figure 1.

Flat distribution

The sixth model is based on the binary distribution, where the stronger 70% of the population makes the value one and 30% of the population makes the value 0; we denote such a function f_{fl} and model by $f_{fl} = (1 + e^{28-8x})^{-1}$, where $x \in [0, 12]$ represents the position of the candidate in population. Such a model represents very basic positions such as gatekeeper, laborer, or unqualified factory worker. In this position, it is hard to create extra value by being good at it, and if somebody is skilled, they could probably be recognized and transferred to the more demanding position.



function\ candidate	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Classical Secretary f_{cs}	0	0	0	0	0.01	0.01	0.01	0.01	0.01	0.01	0.01	1
Order dist. f_{or}	0	0.09	0.18	0.27	0.36	0.46	0.55	0.64	0.73	0.81	0.91	1
Normal dist. f_{no}	0	0.14	0.25	0.33	0.41	0.47	0.54	0.61	0.68	0.76	0.86	1
Double-normal dist. f_{un}	0	0.14	0.22	0.29	0.36	0.42	0.48	0.54	0.62	0.70	0.80	1
Pareto dist. f_{pa}	0	0	0	0	0	0.01	0.02	0.06	0.13	0.29	0.55	1
Flat dist. f_{fl}	0	0	0	0.01	0.98	1	1	1	1	1	1	1

Figure 1 Graphs and values of utility function for candidates

2.2 Strategies

We define a strategy as a generalization of the classical solution; the decider examines a sample from a population, the first k candidates. Then he considers a number of them he wishes the preferred candidate to overcome m; for example, $k = 5, m = 3$ would mean we see the first five candidates (not pick them) and pick the first candidate better than the worst three from 5 tried. The best strategy for the classical problem should be $k = m \approx n/e$. If the condition of choice is not fulfilled by any candidate, we pick the last one. Note that plausible values for the size of the sample are $k \in \{1, \ldots n-2\}$ and for the number of candidates from the sample under threshold $m \in \{1, \ldots k\}$. $k = 0$ would mean pick the first one and $k = n-1$ would mean pick the last one; such strategies are clearly inferior.

Evaluation of strategies

Let us assume that we are given a list of candidates with their value for the position given by function from the previous section. We use a strategy, but the winner is influenced by the random order of candidates, so the result is a random variable. As we wish to provide full results, we analyse all possible ordering by going through all permutations. Each strategy is marked by the frequency of all possible outcomes. From these frequencies, we could evaluate the mean value and the standard deviation of the process. Obviously, the preferred variant is the high mean value and low deviation. We simplify the evaluation quality of the strategy to the mean value of the picked candidate. Therefore, the result for the given size of the sample k and the given threshold m from the sample gives the exact number, and the highest number indicates the best strategy, i.e., the strategy with the highest mean value of the chosen candidate. This evaluation method is compatible with the original problem; given 1 to the best candidate and 0 to all others, the mean value is equal to the probability of picking the best one.

3 Results

Our analysis provides the frequencies of the choice of any candidate based on the parameters of the strategy. Then, we multiply the frequency by the utility value of the candidate and sum it up into the mean value. We calculate the mean values for the admissible combinations of k, the size of the sample, and m, the number of candidates from the sample to set the threshold. We get a table of mean values, unique for each distribution; as an example, we present values for normal distribution in Figure 2.

Mean value		The number of candidates from sample under the treshold m									
		1	2	3	4	5	6	7	8	9	10
The size of sample k	1	0,73									
	2	0,674	0,778								
	3	0,637	0,745	0,781							
	4	0,613	0,705	0,775	0,765						
	5	0,596	0,675	0,742	0,78	0,739					
	6	0,584	0,653	0,712	0,756	0,767	0,708				
	7	0,575	0,635	0,687	0,728	0,75	0,739	0,672			
	8	0,568	0,62	0,664	0,699	0,72	0,722	0,698	0,633		
	9	0,561	0,605	0,64	0,666	0,681	0,685	0,674	0,644	0,592	
	10	0,55	0,581	0,603	0,616	0,623	0,622	0,616	0,602	0,58	0,549

Figure 2 Mean values for strategies in case of normal distribution

The best strategy is the one with the highest mean value, in our case $k = 3, m = 3$, but surprisingly almost the same value is $k = 5, m = 4$. It seems that there are at least two local maximums.

3.1 The best strategies

Similarly to what is revealed in the case of a normal distribution, we find out that for the classical Secretary problem, the best strategy is $k = m = 4$, which corresponds to established results ($k = m = 4 \approx 4.4 \approx 12/e$). Allowing m to be different from k does not improve our chances. In the case of order distribution, the best strategy is $k = 5, m = 4$, with the second best strategy being less than 1 percent lower, $k = 4, m = 3$. The best strategy for the normal distribution is $k = 3, m = 3$ with $k = 5, m = 4$ less than the percent below. For the Double-normal distribution $k = 3, m = 3$ is the best. For the Pareto distribution $k = 4, m = 4$ is the best while $k = 3, m = 3$ is less than one percent below. Finally, for the Flat distribution $k = 5, m = 3$ is the best, with $k = 6, m = 3$ being closely second.

3.2 Applying nonspecific strategies

Let us consider an experiment when we use the optimal strategy for a given distribution to another distribution. This may represent misuse or a wrong estimate of how the population of candidates may perform at the job in question; we provide a table of what percentage of optimal choice we reach by this misjudgment of optimal strategy.

	f_{cs}	f_{or}	f_{no}	f_{un}	f_{pa}	f_{fl}
Normal dist. and Double-normal dist. strategy: ($k = 3, m = 3$)	0.96	0.99	1	1	1	0.93
Classical Secretary and Pareto dist. strategy: ($k = 4, m = 4$)	1	0.96	0.98	0.98	1	0.90
Order dist. strategy: ($k = 5, m = 4$)	0.67	1	1	0.99	0.89	0.99
Flat dist. strategy: ($k = 5, m = 3$)	0.47	0.96	0.97	0.93	0.69	1

Table 1 Nonspecific strategies success rate

Note that the failure of a strategy with $k > m$ for unsuitable distributions is more significant than using strategies based on $k = m$ in cases where it is not optimal.

4 Conclusions and discussion

As we have revealed in section 3.2, we can split using the optimal strategies and the distributions into three cases. The first case is the one with high increments close to the best candidate, which the classical Secretary problem or Pareto model can cover. The second case is if any above-average candidate choice is good, and the best choice does not bring much extra value; this corresponds to the Flat and Order distributions. In the third case, in between, where above-average is reasonable but opting for the best is beneficial, corresponds to the Normal and Double-normal distribution. In the first case, it is reasonable to pick a small sample (around one-third of the population) and set the threshold to be the best from the sample. This is risky, but the bigger chance for the best candidate is worth it.

In the second case, picking a bigger sample (between one-third and one-half) and then setting the threshold based on the better but not the best from the sample is preferred. This lowers the chance of getting the best candidate but highly increases the ratio of picking above average to below average and provides value by this. In the third case, both views are valid and can be highly effective. The most interesting outcome of our model is the double local maximum of the mean value for the Normal and Double-normal distribution cases. This means that both strategies (risking but aiming for the best or safely studying more populations to avoid the risk) are valid. Then, it is up to the decision maker to prefer either the higher risk with a bigger mean value of the candidate or a slightly lower mean value with a lower risk of the bad choice. This inspires the incorporation of the standard deviation of the choice and ideas similar to the Markowitz models of a portfolio combining deviation and mean value based on the affection to risk, which would be the topic of future research.

References

Albers S, Ladewig L (2021) New results for the k-secretary problem. Theoretical Computer Science 863:102–119. https://doi.org/10.1016/j.tcs.2021.02.022

Angelovski A, Güth W (2020) When to stop — a cardinal secretary search experiment. Journal of Mathematical Psychology 98:102425. https://doi.org/10.1016/j.jmp.2020.102425

Ansar N, Baloch A (2018) Talent and talent management: Definition and issues. IBT Journal of Business Studies 14(2):174–186. https://doi.org/10.46745/ilma.jbs.2018.14.02.14

Armstrong M, Taylor S (2020) Armstrong's handbook of human resource management practice. Kogan Page Publishers, London

Chambers EG, Foulon M, Handfield-Jones H, et al (1998) The war for talent. The McKinsey Quarterly 3:44–57

Christian B, Griffiths T (2016) Algorithms to Live By: The Computer Science of Human Decisions. Henry Holt and Company, New York

Dütting P, Lattanzi S, Paes Leme R, et al (2021) Secretaries with advice. In: 22nd ACM Conference on Economics and Computation, EC 2021. Association for Computing Machinery, New York, pp 409–429

Ericsson KA, Prietula MJ, Cokely ET (2007) The making of an expert. Harvard business review 85(7-8):114—21, 193. http://europepmc.org/abstract/MED/17642130

Frankovský M, Birknerová Z (2014) Measuring social intelligence-the mesi methodology. Asian Social Science 10:90. https://api.semanticscholar.org/CorpusID:144457708

Gallardo-Gallardo E, Dries N, González-Cruz TF (2013) What is the meaning of 'talent' in the world of work? Human Resource Management Review 23(4):290–300. https://doi.org/https://doi.org/10.1016/j.hrmr.2013.05.002, https://www.sciencedirect.com/science/article/pii/S1053482213000302

Goldstein DG, McAfee RP, Suri S, et al (2019) Learning when to stop searching. Management Science 66(3):1375–1394. https://doi.org/10.1287/mnsc.2018.3245

Lathesh K, Avadhani V (2018) A study on social intelligence and its impact on employee performance of insurance sectors in mysuru city. International Journal of Mechanical Engineering and Technology 9:530–537

Liker JK, Meier D (2007) Toyota Talent: Developing Your People – The Toyota Way. McGraw-Hill, New York

Navarra K (2022) The real costs of recruitment. https://www.shrm.org/topics-tools/news/talent-acquisition/real-costs-recruitment

Obukhova E, Tian F (2024) Referral bonuses in global talent acquisition: the role of social networks in china and the us. Journal of International Business Studies https://doi.org/10.1057/s41267-024-00682-x

O'Mahoney J, Sturdy A (2016) Power and the diffusion of management ideas: The case of McKinsey & Co. Management Learning 47(3):247–265. https://doi.org/10.1177/1350507615591756, https://doi.org/10.1177/1350507615591756, https://doi.org/10.1177/1350507615591756

Piva E, Stroe S (2023) New ventures fighting the war for talents: the impact of product innovativeness and entrepreneurs' passion on applicant attraction. Small Business Economics 61:1133–1159. https://doi.org/10.1007/s11187-022-00714-6

Swailes S (2016) The cultural evolution of talent management: A memetic analysis. Human Resource Development Review 15(3):340–358. https://doi.org/10.1177/1534484316664812

Sub-optimizing Routes in MS Excel: An Approach of Solving the Traveling Salesman and Vehicle Routing Problems

Tereza Chmelová[1]

Abstract. This paper introduces an approach for solving Traveling Salesman Problem (TSP) and Vehicle Routing Problem (VRP) for a heterogeneous fleet of vehicles in MS Excel. MS Excel, along with its optimization solver add-in application with implemented evolutionary algorithm, was selected as the platform for development of the procedure. The evolutionary algorithm enables MS Excel to identify sub-optimal routes of the considered problems within a short timeframe. The second part of the paper presents two examples that illustrate the application of the proposed procedure for a TSP with 23 locations, and a VRP with 26 locations. All locations are situated in Prague, the capital city of the Czech Republic, and its vicinity. The results prove the effectiveness of the algorithm in optimizing route planning for urban environments, highlighting its practicality in real-world scenarios.

Keywords: traveling salesman problem, vehicle routing problem, MS Excel solver, heterogeneous fleet

JEL Classification: C44, C61
AMS Classification: 90C08

1 Introduction

MS Excel and its Solver add-in can be used as an optimization tool. Its main advantages are its (financial) accessibility, user-friendliness, intuitive operation and, most importantly, its widespread use by companies and individuals. Although numerous programs support optimization, they are usually very expensive, making the investment too high for many companies. Since many companies and individuals already use MS Excel, so the method described in this paper does not require any further investment.

The aim of this article is to present the method of solving TSP and VRP in MS Excel. First, the algorithms will be described in detail and then their application will be demonstrated through real examples. This paper focuses on solving TSP and VRP with a heterogeneous fleet. The mathematical formulations of these problems will not be described here as they are well-known. Their introduction and detailed description can be found e.g. in (Pelikán, 2001) and (Fábry, 2014).

TSP and VRP are roles we have known for many years. Many articles have been devoted to their description and solution. One example for all may be (Applegate et al, 2006). There are also many ways to generalize or extend these tasks. This is the focus of e.g., (Toth and Vigo, 2002), (Berger et al, 2018) or (Gutin and Punner, 2002). Many authors have also already outlined tasks of solution TSP or VRP in MS Excel. Here we can mention e.g. (Gusron, 2023), (Rasmussen, 2011) and (Stoilova and Stoilov, 2020).

The rest of the paper contains 3 main sections. Section 2 describes the algorithms and approaches for solving TSP and VRP. Section 3 illustrates using the proposed approach in solving real-world TSP and VRP examples. The last Section 4 concludes the research and summarizes the findings.

2 Solving optimization problems in MS Excel

There are many optimization programs on the market. Among them are Lingo, MPL for Windows, MS Excel Solver, AIMMS, XPRESS-MP, GAMS, OPL Studio, AMPL. Of these, I would like to describe MS Excel Solver in this section.

Although MS Excel is not primarily optimization software, Solver helps fulfill this purpose. The unquestionable advantage of using Solver is, of course, its price. In the case that a user has purchased a license for MS Excel, Solver is already included. Another reason why Solver seems advantageous is the widespread use of MS Excel. It is almost a given to be able to use at least basic functions of MS Excel in many larger companies today. For most people, the MS Excel environment is not entirely unfamiliar, making the use of Solver relatively intuitive.

[1] University of Economics in Prague, Department of Econometrics,Winston Churchill Square 4, 13067 Prague, CzechRepublic, chmt02@vse.cz

All functions mentioned in following parts of this paper can be found in the basic version of Excel, and their detailed descriptions can be found directly on the company's website (Microsoft, 2021).

Solver offers the use of three different algorithms – simplex, gradient, and evolutionary methods. The simplex method is used for linear functions, the gradient method is used for nonlinear smooth functions, and the evolutionary algorithm can be used for any function.

2.1 Evolutionary algorithm

The evolutionary algorithm is a heuristic method that can solve non-smooth functions. This makes its use suitable for the algorithm for solving TSP and VRP, which is the subject of this paper

As Pezer (2016) writes, the use of evolutionary methods or genetic algorithms for finding the shortest route in TSP is quite common. The reason is that although the solution found may not always be optimal, it is at least very close to the optimal one, which is a trade-off often worth making for the significantly shorter time it takes for this search.

If we wanted to use gradient methods for the calculations, it would be necessary to adjust the input so that it does not contain the INDEX function in MS Excel, meaning that the task should be formulated in the same way as in the classic mathematical model of TSP and VRP. However, this would mean a huge increase in the number of variables. For comparison, let's consider the TSP problem. In the model described by Stoilova and Stoilov (2020), variables are interpreted as indices of places visited in the specified order, and their number is only n. In contrast, in the classic TSP model, there are $n^2 + n$ variables. As the number of visited places n increases, the difference becomes even more noticeable. Therefore, it is much more advantageous to solve the problem using an evolutionary algorithm, as also discussed by Jiang (2010).

2.2 Solving the Traveling Salesman Problem in MS Excel

One of the many tasks that can be solved in the Solver for MS Excel is the Traveling Salesman Problem (TSP). As described, for example, by Stoilova and Stoilov (2020), there is a simple procedure to input the necessary values for Solver. Following part describe this algorithm of solving TSP in MS Excel.

	A	B	C	D	E	F	G	H	I	J	K
1			Distances between places								
2			Place 1	Place 2	Place 3	Place 4	Place 5		Visited	Variables	Distance travelled
3	1	Place 1	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}		1. place visited	1	=INDEX($C$3:$G$7;J3;J4)
4	2	Place 2	c_{21}	c_{22}	c_{23}	c_{24}	c_{25}		2. place visited	2	=INDEX($C$3:$G$7;J4;J5)
5	3	Place 3	c_{31}	c_{32}	c_{33}	c_{34}	c_{35}		3. place visited	3	=INDEX($C$3:$G$7;J5;J6)
6	4	Place 4	c_{41}	c_{42}	c_{43}	c_{44}	c_{45}		4. place visited	4	=INDEX($C$3:$G$7;J6;J7)
7	5	Place 5	c_{51}	c_{52}	c_{53}	c_{54}	c_{55}		5. place visited	5	=INDEX($C$3:$G$7;J7;J8)
8									Ending point	=J3	
9											
10											
11									z=	=SUMA(K3:K7)	--> min

Figure 1 Model TSP in MS Excel environment, Source: self-processed

In Figure 1, the general forms of tables for finding routes between five places are displayed, serving as the initial source for the Solver. The first table is a database of distance values between individual places, denoted generically as c_{ij}. The second table shows the sequence of the five places, one of which is listed twice because the last visited place must match the first one (place where the vehicle started, it must return to at the end of the route). The model variables, marked in yellow, are indices of places 1 to 5. E.g. the first variable means that the first visited place will be the one with index 1, the second variable labeled with 2 indicates the second visited place with index 2 etc. The last visited place (matching the starting place) is no longer a variable but is determined by a function to correspond to the value of the first visited place. This is included in the model so that the `INDEX` function can retrieve the distance of the last route of the vehicle returning to the starting point. The objective function highlighted in orange represents the sum of all traveled distances.

Figure 2 shows the Solver and how requests are input into it – the objective function is the field marked in orange. Variables, of which there are n (in the sample model given above, $n = 5$). Then, the condition states that the values of all variables must be different, and none of them can appear twice in the solution.

2.3 Solving the Vehicle Routing Problem in MS Excel

Since the Vehicle Routing Problem (VRP) is essentially just a generalization of the TSP model, the original algorithm for solving the TSP task in MS Excel – described in article (Stoilova and Stoilov, 2020) – can be adapted to solve VRP as well.

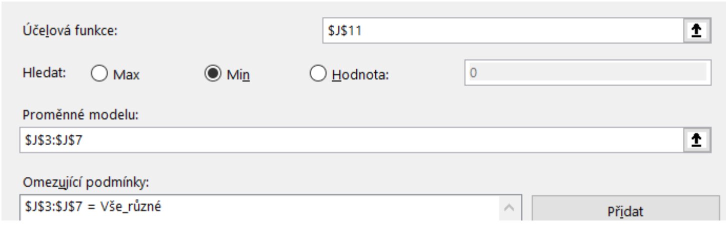
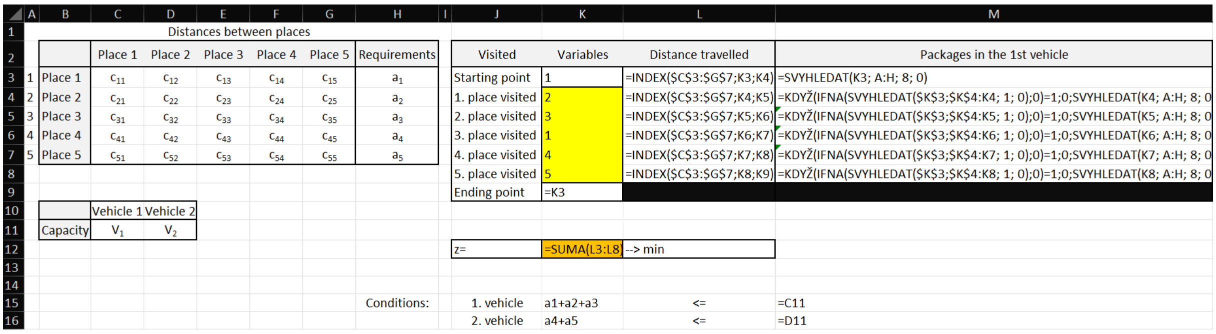


Figure 2 Solver for the TSP model, Source: self-processed

Solving Vehicle Routing Problem for Two Vehicles in MS Excel

In Figure 3, modified version of the model is displayed to solve the VRP for two vehicles.



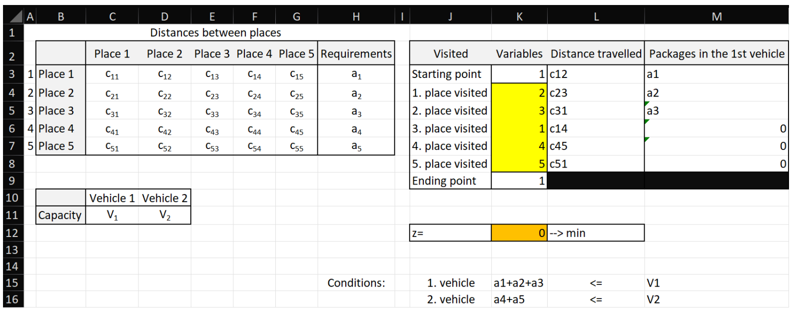
	A	B	C	D	E	F	G	H	I	J	K	L	M
1			Distances between places										
2			Place 1	Place 2	Place 3	Place 4	Place 5	Requirements		Visited	Variables	Distance travelled	Packages in the 1st vehicle
3	1	Place 1	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	a_1		Starting point	1	`=INDEX($C$3:$G$7;K3;K4)`	`=SVYHLEDAT(K3; A:H; 8; 0)`
4	2	Place 2	c_{21}	c_{22}	c_{23}	c_{24}	c_{25}	a_2		1. place visited	2	`=INDEX($C$3:$G$7;K4;K5)`	`=KDYŽ(IFNA(SVYHLEDAT($K$3;$K$4:K4; 1; 0);0)=1;0;SVYHLEDAT(K4; A:H; 8; 0`
5	3	Place 3	c_{31}	c_{32}	c_{33}	c_{34}	c_{35}	a_3		2. place visited	3	`=INDEX($C$3:$G$7;K5;K6)`	`=KDYŽ(IFNA(SVYHLEDAT($K$3;$K$4:K5; 1; 0);0)=1;0;SVYHLEDAT(K5; A:H; 8; 0`
6	4	Place 4	c_{41}	c_{42}	c_{43}	c_{44}	c_{45}	a_4		3. place visited	1	`=INDEX($C$3:$G$7;K6;K7)`	`=KDYŽ(IFNA(SVYHLEDAT($K$3;$K$4:K6; 1; 0);0)=1;0;SVYHLEDAT(K6; A:H; 8; 0`
7	5	Place 5	c_{51}	c_{52}	c_{53}	c_{54}	c_{55}	a_5		4. place visited	4	`=INDEX($C$3:$G$7;K7;K8)`	`=KDYŽ(IFNA(SVYHLEDAT($K$3;$K$4:K7; 1; 0);0)=1;0;SVYHLEDAT(K7; A:H; 8; 0`
8										5. place visited	5	`=INDEX($C$3:$G$7;K8;K9)`	`=KDYŽ(IFNA(SVYHLEDAT($K$3;$K$4:K8; 1; 0);0)=1;0;SVYHLEDAT(K8; A:H; 8; 0`
9										Ending point	=K3		
10			Vehicle 1	Vehicle 2									
11		Capacity	V_1	V_2									
12										z=	=SUMA(L3:L8)	--> min	
13													
14													
15								Conditions:		1. vehicle	a1+a2+a3	<=	=C11
16										2. vehicle	a4+a5	<=	=D11

Figure 3 VRP model in MS Excel environment, Source: self-processed

There are several differences between the TSP and VRP models. The first is the addition of customer demands a_i and vehicle capacities V_k. Customer demands can be understood, for example, as the number of packages to be delivered to a certain location. Vehicle capacities represent the upper limit of how many packages can fit in vehicle k.

Another difference is the meaning of the first visited (i.e., starting) location. In the previous model, the starting point was a variable. In contrast, here it is a fixed value. Index 1 therefore denotes a specific point, which could be, for example, a warehouse or a store from which goods are distributed to customers. Since the first location is not a variable, it can appear once more in the model. This can have two interpretations: either considering a single vehicle that delivers the cargo in parts, or considering a multi-vehicle problem where vehicles can depart simultaneously – the first circuit is then the route of the first vehicle ending at the first occurrence of variable 1, where the route of the second vehicle starts, ending with the return to the ending point (same as starting point). In the example shown in Figure 3, we are considering VRP of two vehicles with different capacities V_k.

Additionally, a column "Packages in 1st Vehicle" has been added, where packages carried by the first vehicle are looked up using MS Excel's functions. (Similarly, it would be possible to add a column for "Packages in 2nd Vehicle" instead, but it is not necessary for the model.)



	A	B	C	D	E	F	G	H	I	J	K	L	M
1			Distances between places										
2			Place 1	Place 2	Place 3	Place 4	Place 5	Requirements		Visited	Variables	Distance travelled	Packages in the 1st vehicle
3	1	Place 1	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	a_1		Starting point	1	c12	a1
4	2	Place 2	c_{21}	c_{22}	c_{23}	c_{24}	c_{25}	a_2		1. place visited	2	c23	a2
5	3	Place 3	c_{31}	c_{32}	c_{33}	c_{34}	c_{35}	a_3		2. place visited	3	c31	a3
6	4	Place 4	c_{41}	c_{42}	c_{43}	c_{44}	c_{45}	a_4		3. place visited	1	c14	0
7	5	Place 5	c_{51}	c_{52}	c_{53}	c_{54}	c_{55}	a_5		4. place visited	4	c45	0
8										5. place visited	5	c51	0
9										Ending point	1		
10			Vehicle 1	Vehicle 2									
11		Capacity	V_1	V_2									
12										z=	0	--> min	
13													
14													
15								Conditions:		1. vehicle	a1+a2+a3	<=	V1
16										2. vehicle	a4+a5	<=	V2

Figure 4 Optimization model of VRP in MS Excel with values, Source: self-processed

For a better understanding of the combined functions, Figure 4 displays the model with the actual values, which were looked up or calculated and how they are visible to the user. The function in column `M` (= Packages in 1st

Vehicle) works in such a way that as long as it finds the variable denoting the starting location (in this case, location with index 1) in the designated field, which starts with the first variable and ends with the variable in the given row (where the current formula is explained), it assumes that the first vehicle is always carrying the packages and displays the value of demand a_i. However, if it finds the index of the starting point in the variables, it means the end of the 1st vehicle's route and the departure of the 2nd vehicle. The formula does not apply to the first starting location because the index of the starting location would obviously be found there. Therefore, it immediately looks up the number of packages.

With this column containing packages for the 1st vehicle, conditions can be added, which are also shown in Figure 3. The left side of the first condition sums up all the packages carried by the first vehicle. The sum of all carried packages must be less than the capacity of the 1st vehicle. The second condition is exactly the same, only the number of packages carried by the second vehicle is calculated as the difference between the total number of packages and those already transported by the 1st vehicle.

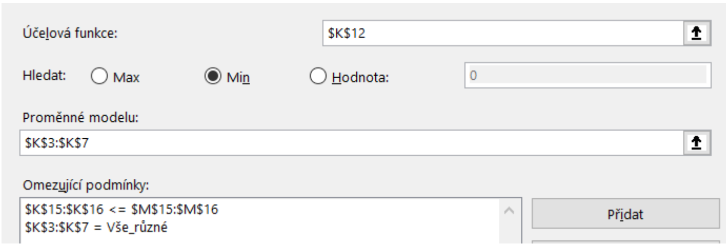


Figure 5 Solver for the optimization model of VRP, Source: self-processed

Neither the objective function nor the condition that all variables must be different has changed compared to the previous TSP model. The parameters of the Solver for the VRP in MS Excel are shown in Figure 5, but the only difference is the addition of two previously mentioned new conditions. All other parameters remain the same.

3 Illustrative examples

The algorithm was used to find two routes between locations in the capital city of the Czech Republic – Prague. Distances and travel times between locations were generated on the website Openroute Service `https://openrouteservice.org`. The input parameters for distance calculation were latitude and longitude. Coordinates of individual locations were obtained from Google Maps `https://www.google.com/maps`.

For the application of the task, it is possible to choose both minimizing time and minimizing distance. In the following two examples, distance minimization is chosen. In the results, you will also find the travel time, which is more informative.

3.1 Traveling Salesman Problem

The algorithm was first used to find the shortest path among 23 locations. The solution can be found in Table 1 and Figure 6. Locations labeled 1 and 24 represent the starting and ending points. In the first column, you will find the sequence in which the locations were visited; in the second column, the street names where the stops are located. The third and fourth columns show the distance traveled and the time spent on the journey between individual locations.

In Table 1, you can see that some streets were visited twice. These are different locations within the same street. Therefore, the distances and travel times between these locations are not high but also not zero.

Overall, the vehicle will cover a distance of 93.7 km, which will take 189 minutes (or 3 hours and 9 minutes). The results were then compared with those generated by the Lingo software (LINDO Systems, 2023), and it was found that this heuristic algorithm did indeed find the optimal solution.

In Figure 6, the route is displayed on a map, with the visualization sourced from Openroute Service `https://openrouteservice.org`. The location indices correspond to the order in which they were visited (the list is provided in the results table 1). If a particular location is not visible (for example, location *12*: Šrobárova), it is because certain locations are obscured on the map due to being in the same or similar location as the following location.

Table 1 Overview of the results of Task 1, Source: self-processed

Order	Place (street)	Dist. [km]	Time [min]	Order	Place (street)	Dist. [km]	Time [min]
1	Křivatcova	0,3	2	13	Šrobárova	2,6	6
2	Křivatcova2	12,7	14	14	Velehradská	5,2	11
3	Kněževes	8,5	13	15	Hvězdova	4,9	11
4	Evropská2	8,5	18	16	Vídeňská	7,7	18
5	Jungmanova	8,2	21	17	Na Slupi3	0,9	2
6	Mazurská2	1	3	18	U Nemocnice2	0,1	0
7	Ústavní	4,4	10	19	U Nemocnice	0,9	3
8	Budínova	3,1	6	20	Karlovo nám.	1,6	4
9	Lovosická	4,6	8	21	Bozděchova2	0	0
10	Sokolovská	3,1	6	22	Bozděchova2	5,4	10
11	Poděbradská	5,1	11	23	V Úvalu	4,9	11
12	Šrobárova	0	0	24	Křivatcova	0	

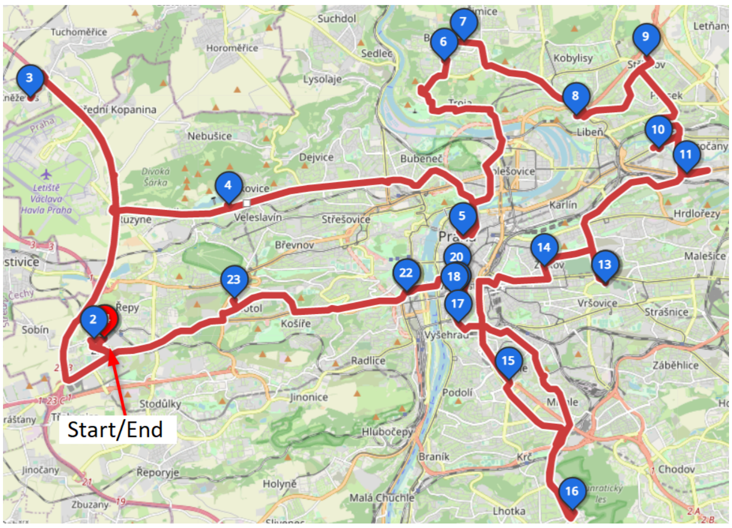


Figure 6 Route 1 shown on the map, Source: Openroute Service

3.2 Vehicle Routing Problem with Two Vehicles

The second example of utilizing the algorithm for solving VRP in MS Excel was used to find the shortest route among 27 locations. Now, additional packages have been added to the task, which need to be transported, along with the capacities of the vehicles available. In total, there are 329 packages to be transported in two vehicles with heterogeneous capacity. Vehicle 1 has a capacity set to 233 packages and Vehicle 2 to 500 packages.

Table 2 Overview of the results of Task 2, Source: self-processed

Idx.	Place (street)	Dist. [km]	Time [min]	Packs	Veh.	Idx.	Place (street)	Dist. [km]	Time [min]	Packs	Veh.
1	Křivatcova	16,4	20	0	v1	15	Šrobárova	9,1	20	50	v2
2	Pekařská	11	19	1	v1	16	Hviezdoslavova	7,1	17	5	v2
3	Čs. Exilu	5,6	11	1	v1	17	K Pérovně	0	0	93	v2
4	Vídeňská2	4,4	11	1	v1	18	K Pérovně	0,8	2	5	v2
5	Antala Staška	3,6	8	59	v1	19	Štěrboholská	6,6	16	3	v2
6	Podolské nábř.	3,1	8	4	v1	20	Vajgarská2	7	14	13	v2
7	Nádražní3	10,5	23	5	v1	21	Lovosická	4,6	12	8	v2
8	Křivatcova	9	17	0	v1/2	22	Budínova2	0,7	2	1	v2
9	Walterovo nám.	0,7	2	2	v2	23	Budínova	1,9	4	11	v2
10	Jinonická	5,7	19	2	v2	24	Zenklova	2,1	5	15	v2
11	U Nemocnice	3,2	7	2	v2	25	Thámova	2,2	7	235	v2
12	Velehradská	1,5	4	5	v2	26	Klimentská	8,2	18	9	v2
13	Olšanská	2,6	6	7	v2	27	Nádražní3	7,3	15	7	v2
14	Koněvova	3,2	7	20	v2	28	Křivatcova			0	v2

The solution found can be found in Table 2 and in Figure 7. The first four columns of the table remain the same as in the results table 1. Additionally, two new columns have been added - the number of packages to be delivered to the customer and the vehicle assigned for transportation.

Locations marked 1, 8, and 28 again represent a single location, indicating the start/end. Specifically, location 1 denotes the start for vehicle 1. Location 8 denotes the end for vehicle 1 and simultaneously the start for vehicle 2. Location 28 is then the end for vehicle 2.

Vehicles can, of course, travel concurrently, thus saving time. Therefore, the first column does not represent a sequential order in the traditional sense, as it is not entirely indicative for a multi-depot VRP. Hence, the first column can be interpreted more as an index or identifier, which will be subsequently used for graphical representation on the map.

The first, smaller vehicle travels 54.6 km and spends 100 minutes on the route (1 hour and 40 minutes). The second vehicle travels 83.5 km, and completing its entire route takes 194 minutes (3 hours and 14 minutes). The resulting value of the objective function (i.e., the sum of all distances traveled by both vehicles) is 138.1 km.

Here, we will take a closer look at the route. From Table 2, it can be inferred that the routes consist of 25 customer locations and the Schubert CZ warehouse visited three times, as explained earlier. The visual representation will aid in easier interpretation, now labeled as Figure 7. The route is again created using Openroute Service `https://openrouteservice.org`.

The image now shows two routes. The warehouse is labeled 1, 8, and 28, as in the results table 2, indicating the last stop on the route.

After comparing the results from the Lingo software (LINDO Systems, 2023), it was found that in this case, the route was not optimal. The value of the objective function for the optimal route was 137.4 km. This represents a small difference, nevertheless, demonstrating that this algorithm is indeed just a heuristic and cannot reliably provide us with the optimal solution.

Figure 7 Route 2 shown on the map, Source: Openroute Service

4 Conclusion

In this paper, a modification of the previously presented algorithm for solving TSP in MS Excel (described by Stoilova and Stoilov (2020)) was proposed. The modification consisted in extending the algorithm to solve a more general problem – namely VRP for 2 vehicles with heterogeneous capacity. Thus, a relatively simple way to solve VRP in MS Excel is possible. The disadvantage of this algorithm is that it is a sub-optimization and cannot guarantee finding the optimal solution. However, this is redeemed by the fact that the solution of the problem is relatively fast and especially by the fact that there is no need to purchase any expensive software for solving such problems.

Further consideration could, for example, be given to extending this approach for VRPs with multiple vehicles. In this case it would be necessary to create more fixed points between variables and to deal with the problems associated with this step.

Acknowledgements

The resrach is supported by the Prague University of Economics, Faculty of Informatics and Statistics, project no. F4/10/2024.

References

Applegate DL, Bixby RE, Chvátal V, Cook WJ (2006) The Traveling Salesman Problem: A Computational Study. Princeton University Press, Princeton

Berger A, Kozma L, Mnich M, Vincze R (2018) A time- and space-optimal algorithm for the many-visits tsp. Cornell University 16(3):1–22. `https://dl.acm.org/doi/abs/10.1145/3382038`

Fábry J (2014) Okružní a rozvozní úlohy. habilitační práce, Vysoká Škola Ekonomická v Praze, `https://www.janfabry.cz/Habilitace.pdf`

Gusron R (2023) Solving traveling salesman problems using excel solver evolutionary algorithm. `https://medium.com/@rihot_gusron/solving-traveling-salesman-problems-using-excel-solver-evolutionary-algorithm-e8deea89ca42`

Gutin G, Punner AP (2002) The Travelling Salesman Problem and Its Variations. Springler Science+Business Media, `https://doi.org/10.1007/b101971`

Jiang C (2010) A reliable solver of euclidean traveling salesman problems with microsoft excel add-in tools for small-size systems. Journal of Software 5(7):761–768. https://doi:10.4304/jsw.5.7.761-768

LINDO Systems I (2023) About lindo systems. https://www.lindo.com/index.php/company

Microsoft (2021) Excel functions. https://support.microsoft.com/en-us/office/excel-functions-by-category-5f91f4e9-7b42-46d2-9bd1-63f26a86c0eb

Pelikán J (2001) Diskrétní modely v operačním výzkumu. Professional Publishing, Praha

Pezer D (2016) Efficiency of tool path optimization using genetic algorithm in relation to the optimization achieved with the cam software. Procedia Engineering 149:374–379. https://doi.org/10.1016/j.proeng.2016.06.681

Rasmussen R (2011) TSP in spreadsheets – a guided tour. International Review of Economics Education 10:94–116. https://doi.org/10.1016/S1477-3880(15)30037-2

Stoilova K, Stoilov T (2020) Transportation modelling and solving travelling salesman problem. IOP Conference SeriesMaterials Science and Engineering 878(1). https://doi-org.zdroje.vse.cz/10.1088/1757-899X/878/1/012026

Toth P, Vigo D (2002) Vehicle Routing Problem. Society for Industrial and Applied Mathematics, Philadelphia

Measuring and Analyzing the Technical Efficiency of Floorball Players

Lucie Chytilova , Michal Pazak

Abstract. Floorball, a global phenomenon, demands peak player efficiency for teams to thrive. However, current evaluation methods are often subjective or based on national statistics data. This research introduces objective tools for floorball players, clubs, managers, and coaches. We are using Data Envelopment Analysis (DEA), specifically the basic CCR and BCC models with different variables according to the type of players. We assess the technical efficiency of players in the Czech Extraliga, the top men's league, across an entire season. Analysing players overall, then by forwards and defenders, we identify areas for improvement based on playing position and team affiliation. This novel approach demonstrates the potential of DEA for efficiency measurement, paving the way for a more systematic and objective methodology and coaching in floorball player assessment.

Keywords: Data Envelopment Analysis, floorball, efficiency

JEL Classification: C10, C61, C67
AMS Classification: 90C05

1 Introduction

Effectiveness in sports is an important aspect that affects the performance of athletes and sports organisations, and effectiveness can be understood in many contexts. In general, this concept is more related to the optimisation relationship between maximising results with minimal use of resources. Efficiency is often associated with social justice in sports facilities and public services. For example, a study from Charles University in Prague examined the effectiveness of sports facilities in municipalities and their impact on social justice (Popelka, 2014). This study focused on how the forms of provision of sports facilities affect the conditions for social justice in sports (Špaček, 2016). Another study focuses on social inequalities in sports activities and how sports activities can contribute to reproducing these inequalities (Špaček, 2016). This issue is often overlooked, although sports can contribute to health inequalities and thus be one of the critical mechanisms by which health differences between social classes are created Špaček, 2016).

Effectiveness in sports is complex, including performance optimisation, social justice issues, and inequality. Sports organisations and public institutions must consider these aspects when planning and implementing their sports programs. However, this work focuses primarily on evaluating effectiveness in terms of performance measurement and recommending further training. The effectiveness of players in sports, including floorball, can be measured and tracked in various ways. Here are some methods that are often used:

- Statistical tracking: Tracking statistics during matches, such as goals, assists, shots on goal, shooting success and more, can provide valuable information about player performance (Hůlka et al., 2014).
- Physiological Tests: Physiological tests such as aerobic and anaerobic fitness can provide helpful information about players' physical fitness (Hůlka et al., 2014).
- Heart rate monitoring: Heart rate monitoring is considered one of the most widely used methods for analysing internal workload in a match (Hůlka et al., 2014).
- RPE (Rating of Perceived Exertion): RPE is a valid and effective method of quantifying the load of players in a training unit (Hůlka et al., 2014).
- Video Analysis: Video analysis allows coaches and players to examine the game in detail and identify areas for improvement (Kadlec, 2022).
- DEA models: In one study, data envelopment analysis (DEA) models were used to evaluate the effectiveness of floorball players from the highest Czech competition (Denk, 2016).

It is important to note that no single tool or method is "best" for all players or teams. Performance measurement should be tailored to the individual needs and goals of the players and team in the context of the overall training

Czech University of Life Sciences Prague, Department of Systems Engineering, Kamýcká 129, 165 00 Praha-Suchdol, Czech Republic, chytilova@pef.czu.cz

VŠB-TUO, Department of Systems Engineering, Sokolská třída 2416/33, 702 00 Ostrava - Moravská Ostrava, Czech Republic, michal.pazak.st@vsb.cz

and competition plan.

As already mentioned, DEA will primarily be used in this article. This is a method that is often used in sports. In their overview article, Bhat et al. (2019) systematically analyse the application of DEA in evaluating performance and efficiency in nine popular sports (football, golf, basketball, etc.). They dealt with 102 published articles on sports and various applications of DEA (classical models, cross efficiency, two-stage DEA, super efficiency, assurance regions, DEA slack-based measure, etc.). The authors point out here that DEA is a great performance evaluation technique in sports - it provides a measure of relative effectiveness where players/teams' performance is measured against others and helps identify the strengths and weaknesses of players' play. This method appears to be very popular in football, which can also be seen in the work of Pelloneová and Tomíček (2022) in the Czech environment. The authors focus on measuring the effectiveness of players in the Czech and Danish top football competitions in the seasons 2015/16 to 2019/20. Here, the authors deal with the effectiveness of the players and their position on the field. Švátora et al. (2022) already focused on floorball, where they compared the level of floorball skills and movement abilities of children aged U-11 - U-14 in floorball clubs in the Czech Republic and Australia. The presented results indicate a different level of floorball skills between Czech and Australian floorball players in the age category we tested. Since this article considers the selection of variables so that the players get injured as little as possible, it is good to draw attention to the article by Tervo and Nordström (2014), where the authors review 75 articles. One article each was identified from sports management and sports psychology. The most thoroughly researched topic was the epidemiology of injuries among floorball players, where the result says that more attention needs to be paid to prevention in the club itself (better trainers, more physiotherapists, etc.).

The article is divided into 5 parts. The first part is this: an introduction that covers what effectiveness in sports is and how effectiveness in sports can be measured. There is also a mention of floorball. The second part, Methodology - Data Envelopment Analysis, introduces the methodology used in this article. This is followed by a presentation of the data and models used in chapter three. Chapter four then analyses all the results that were carried out and summarises the most fundamental findings in the last chapter, i.e., the conclusion.

2 Methodology of DEA

Data Envelopment Analysis (DEA) is a non-parametric approach. It is widely used for measuring the relative efficiency of decision-making units (DMUs) with multiple inputs and outputs. Assume, there is a set of T DMUs (DMU$_k$ for $k = 1, \ldots, T$), let input and output variables data be $X = \{x_{ik}, i = 1, \ldots R; k = 1, \ldots T\}$ and $Y = \{y_{jk}, j = 1, \ldots S; k = 1, \ldots T\}$, respectively. Also, u_i for $i = 1, \ldots R$ and v_j for $j = 1, \ldots, S$ be the weights of the i^{th} input variable and the j^{th} output variable, respectively. Mathematically, the relative efficiency score of DMU$_k$ can be defined as:

$$e_k = \frac{\sum_{j=1}^{S} v_j y_{jk}}{\sum_{i=1}^{R} u_i x_{ik}}, \text{ for } k = 1, \ldots, T. \tag{1}$$

Charnes et al. (1978) have proposed the following CCR model to measure the efficiency score of the under-evaluation unit, DMU$_Q$ where $Q \in \{1, \ldots, T\}$:

$$\begin{array}{lll} & \max e_Q = \frac{\sum_{j=1}^{S} v_j y_{jQ}}{\sum_{i=1}^{R} u_i x_{iQ}}, & \\ \text{s.t.} & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} \leq 0, & k = 1, \ldots, T, \\ & u_i \geq 0, & i = 1, \ldots, R, \\ & v_j \geq 0, & j = 1, \ldots, S. \end{array} \tag{2}$$

The model (2) is non-linear. It is the model of linear-fractional programming. The model (2) could be transferred by Charmes-Cooper transformation to the standard linear programming problem:

$$\begin{array}{lll} & \max e_Q = \sum_{j=1}^{S} v_j y_{jQ}, & \\ \text{s.t.} & \sum_{i=1}^{R} u_i x_{iQ} = 1, & \\ & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} \leq 0, & k = 1, \ldots, T, \\ & u_i \geq 0, & i = 1, \ldots, R, \\ & v_j \geq 0, & j = 1, \ldots, S, \end{array} \tag{3}$$

where $Q \in \{1, \ldots, T\}$. DMU$_Q$ is CCR-efficient if and only if $e^* = 1$ and if there exists at least one optimal solution $(\mathbf{u}^*, \mathbf{v}^*)$ with $\mathbf{u}^* > \mathbf{0}$ and $\mathbf{v}^* > \mathbf{0}$ for the set $Q \in \{1, \ldots, T\}$. The inefficient units have a degree of relative efficiency that belongs to interval $[0, 1)$. Note: The model must be solved for each DMU separately.

The model (3) is called a multiplier form of the input-orient-CCR model. However, for computing and data interpretation, it is preferable to work with a model that is dual associated to the model (3). The model is referred to as an envelopment form of the input-oriented CCR model, see Charnes et al. (1978). A multiplier form and envelopment form of the output-oriented CCR model exist. Both models give the same results; see **?** .

Banker et al. (1984) have extended the CCR model. The extended model is called the BCC model and considers variable returns to scale assumption. The model has a convex data envelope, leading to more efficient DMUs. The mathematical model of the dual multiplier form of the input-oriented BCC model is:

$$
\begin{array}{lll}
 & \max e_Q = \sum_{j=1}^{S} v_j y_{jQ} - v_0, & \\
\text{s.t.} & \sum_{i=1}^{R} u_i x_{iQ} = 1, & \\
 & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} - v_0 \leq 0, & k = 1, \dots, T, \\
 & u_i \geq 0, & i = 1, \dots, R, \\
 & v_j \geq 0, & j = 1, \dots, S, \\
 & v_0 \in (-\infty, \infty), &
\end{array} \tag{4}
$$

where v_0 is the dual variable assigned to the convexity condition $\mathbf{e^T}\lambda = \mathbf{1}$ of envelopment form of input-oriented BCC model. Note: The BCC model can be rewritten into the envelopment form or changed into the output orientation.

The input-oriented CCR and BCC models are used in applications.

3 Data and definition of used models

The analysis of the effectiveness of the players (defenders and forwards, goalkeepers were not included in these statistics due to lack of data) is made from the statistics of the 2023/2024 season at the end of the 24th game round of the men's super league. The statistics were taken from the website ceskyflorbal.cz (2024). These statistics are the same across all world leagues. Czech floorball collects 25 statistical information, unlike Finnish or Swedish, where there are only around 10. The Czech league is the best and most suitable for the given analysis. 120 players (60 forwards and 60 defenders) were selected for the given analysis.

In floorball, each player/post has a specific playing position, which varies depending on where the player is on the court and their role in the game system. Since the analysis takes place for two basic types of players who have different positions in the game, it is necessary to make two different models for them. To better understand these models, the basic game position of these two basic types of players is briefly explained below:

- defender - a player who tries to prevent the opponent from scoring a goal or, in general, to prevent the creation of a scoring chance. The defenders are divided into left and right defences, each responsible for their part of the field and defending the opponent's attacks. Also, the defenders are important to the game towards the attack, able to play out a potential attack.
- forward - a player who tries to score a goal or create a scoring chance for his team. Forwards are divided into left and right wings and centres. The role of the left and right wings is to put pressure on the opponent's defence, support the attack and create scoring opportunities. The centre (middle forward) is the key figure of the formation and, at the same time, the game's creator. His responsibility is to support the attack and the defence; he directs the game and tries to create shooting space for his teammates.

Since this article is the first analysis in this direction and the position in floorball often changes (in the sense of right/left wing, etc.), only the division into defenders and forwards is used in this analysis. It should be noted that the team's strategy also affects the players' positions; however, for this analysis, we will assume the basic distribution mentioned above.

Models built for forwards and defenders use the following statistical attributes:

- number of matches - indicates how many times the player has played for his team during the season.
- number of substitutions - indicates how often a player substituted or was substituted in a match or season. A substitution is a change of players on the field that can occur at any time during the game, as long as the rules about the maximum number of players on the field and the prohibition of goalkeeper substitutions are observed.
- average time on the field - indicates how many minutes and seconds the player spent on the field on average in one match. The total average time on the court is calculated as the ratio of the time on the court to the number of games played by the player.

- total penalty time - indicates how many minutes the player spent on the penalty bench in a match or a season, and it is the sum of all the penalties the player received.
- minuses - indicates how often his team got a goal while the player was on the pitch. A player's minus is counted as the number of minus situations the player has experienced.
- total Canadian scoring - indicates how many points the player scored in a match or in a season and is the sum of goals and assists.
- player's plus/minus indicates the difference between the number of goals his team scored and the number of goals his team conceded when the player was on the pitch. A player's plus/minus is calculated as the difference between plus and minus situations. A plus situation is when his team scores a goal while the player is on the pitch. A minus situation is when his team scores a goal while a player is on the pitch.
- shooting success - indicates the percentage rate at which the player converted his shots into goals in a match or season. Shooting success is calculated as the ratio of goals and shots of the player.
- throw-in success - indicates the percentage of throw-in success. Throw-in success is calculated as the ratio of face-offs won and lost.

Specifically, the model is shown in the following Figure 1.

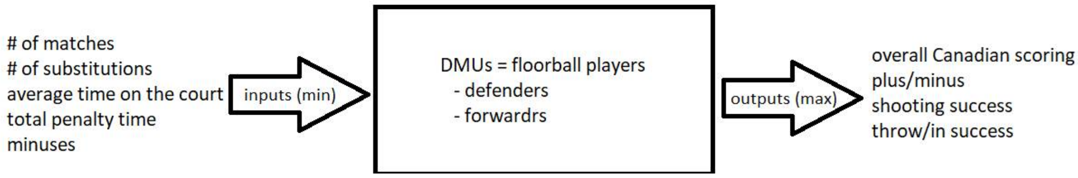


Figure 1 The model for defender and forwarder

It should be noted that there were DMUs with zero variables. Since DEA models cannot work with zero values, the data was modified so there were no zeros or negative values , and it should be noted that 60 defenders and 60 forwards were selected from the league, who hit at least 80 % of the matches.

4 Calculation and discussion

The analyses that were carried out as part of the survey are as follows:

1. analysis of all players regardless of position with the same input and output variables;
2. analysis of players concerning the position and different variables for that position.

Both of these analyses were performed for the CCR model and the BCC model. Input-oriented models were used since inputs are among the variables the player can more easily influence.

Although the calculation of the CCR and BCC models may be unnecessary, based on the agreement between the authors, both models were calculated. First, it was about determining if the model is correct and whether the number of practical units is the same or more significant in the BCC model. This is, therefore, a technical matter Secondly, it was also a confirmation that it is better to measure people using BCC. Since more effective variables are more logical for this model, the sportsman field is not a machine and cannot generate the same output based on input. Still, we need to take into account his physiology and psychology. These parameters must also be added to the model in the future.

The results of all analyses are also shown in Table 1, below.

Within the joint model, Table 1 shows the results of the effectiveness of forwarders and defenders together in one model - a total of 120 DMUs. The number of effective players for the CCR is 24, and for the BCC model, it is 31 (15 forwards and 16 defenders). The average player efficiency is 0.7254 for the CCR model and 0.7869 for the BCC model. The following analyses are already primarily for BCC models, as this model is identified as more suitable for the human effectiveness measurement environment. The ACEMA Sparta Praha team, which at the time was in first place in the competition, had 4 effective (1 attacker and 3 defenders) and 5 ineffective players (5 forwarders and 1 defender). The most effective players, i.e. 10 effective players out of 31, are from the team - Předvýběr.CZ Florbal Mladá Boleslav, which at the time was in second place in the competition table and had the most representative players. There are 7 forwards and 3 defenders. The average number among the most efficient players is 32 Canadian points, and the average number in the plus/minus rating is 39.3. Both metrics represent a really high number, based on which we can label the Předvýběr.CZ Florbal Mladá Boleslav team as an offensively

The second author was a master's student, and therefore, it was necessary to solve the underlying issues.

	A model for all players		A model for a defender		A model for a forwards	
	CCR	BCC	CCR	BCC	CCR	BCC
# of players	120	120	60	60	60	60
# of efficient players	24	31	15	22	20	22
mean value	0.7255	0.7870	0.7674	0.8377	0.8284	0.8840
error of mean value	0.0181	0.0162	0.0276	0.0223	0.0207	0.0149
median	0.7106	0.7910	0.7645	0.8951	0.8097	0.9078
mode	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
standard deviation	0.1984	0.1779	0.2137	0.1728	0.1601	0.1150
maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
minimum	0.2740	0.4519	0.2741	0.4623	0.4371	0.6332

Table 1 Descriptive statistics off all results

tuned team. Only 6 people in this team are ineffective (2 forwards and 4 defenders). 1.SC TEMPISH Vítkovice, the team in third position, has only 4 (2 forwarders and 2 defenders) effective and 10 ineffective players (6 forwarders and 4 defenders). So, it can be seen here that even a relatively good team can have players who are not quite effective, but there is some synergy that will move the team forward. Also, with these results, it is necessary to consider that the data collected is general. If we wanted to go into more detail, we would need more specific data focused on a specific post. For the attacker, it would be various losses or, on the contrary, keeping the ball in the zone, while for the defenders, it would be, for example, the number of unblocked shots or the number of fights won in the defensive zone.

BCC model for 60 defenders (60 DMUs). There are a total of 22 effective and 38 ineffective defenders in this model. The minimum value of the model is 0.462, the average value is equal to 0.837, and the average value of all ineffective players is 0.7438. At the same time, the current typology of statistics is not adequately focused on the defender post. Therefore, the evaluation of the model serves to compare the players' results across models and to recognise the correctness of the models. The Předvýběr.CZ Florbal MB team once again won first place in the number of effective players and in the model, which is primarily focused on defenders, with a total of 5 effective defenders. This fact statistically confirms the fact that the Předvýběr.The CZ Florbal MB team is the strongest on paper and has the most experienced players in the league (even if it is currently second). Among ineffective defenders, this team has only one player with a near-maximum effective BCC score of 0.9572. The team FAT PIPE FLOORBALL CHODŮ took the first-second place in the number of effective goalkeepers. At the same time, it has 4 ineffective players but with an average BCC efficiency of 0.792. ACEMA Sparta Prague has 3 effective defenders; compared to the common model, a defender with national team experience has been added, who was a reinforcement of the team before the 2023/2024 season and, at the same time, earned the captain's armband. So it can be assumed that the team will rely on this player in key moments and should play an important role in establishing and supporting the attack. Compared to his teammates from the defenders of the ACEMA Sparta Prague team, he achieves much higher statistics.

The BCC model for forwards is only aimed at 60 forwards. In total, this model has 22 effective and 38 ineffective forwarders. The minimum value of the model is 0.633, and its average value is 0.884. Předvýběr.CZ Floorball Mladá Boleslav again has the most effective players, with nine effective forwards and no ineffective forwards. Compared to the BCC-I joint model, where ACEMA Sparta Prague was the second most effective team, in the BCC-I forward models, the second team is 1.SC TEMPISH Vítkovice. It has 4 effective players and 4 ineffective ones. The already mentioned ACEMA Sparta Prague team took third place with 3 effective and 2 ineffective.

The analysis was also carried out for specific player names based on intermediate knowledge (the author of the thesis actively plays the given league (Pažák, 2024)). Still, this is not presented here due to the length of the article. However, the authors are happy to provide a more detailed analysis on request. Below is a summary with suggestions and recommendations based on the aforementioned unmentioned analyses:

- Individual Training: Identifying ineffective players and analysing their input and output statistics can indicate areas players should focus on in individual training. Improving technical skills and physical fitness can lead to increased player effectiveness.
- Tactical preparation: Analyzing player effectiveness can also provide information about their game decisions and ability to contribute to overall team performance. Coaches should use this knowledge in preparing tactical strategies and team formations.

- Mental preparation: Besides the physical and technical side of the game, the players' mental preparation is also important. Identifying ineffective players and providing them with motivation, confidence and concentration support can positively affect their performance.
- Continuous monitoring: It is important to regularly monitor players' performance and update their effectiveness based on new statistics and results. This continuous process will allow coaches and team managers to adapt training plans and strategies to the team's current needs.
- Support for teamwork: The effectiveness of players is not only about individual performances but also about the ability to work as a team. Fostering teamwork, communication and trust between players can lead to an increase in overall team effectiveness.

Based on the analysis and recommendations for improving the effectiveness of players in Czech floorball, it is important to pay attention to improving the collection of statistical data and data. Specific steps that could help improve this process include:

- Standardization of data collection: It is important to have clearly defined and standardised procedures for collecting statistics from floorball matches. These procedures should include specifications for specific game posts, leading to a more accurate analysis of players.
- Automating the data collection process: Using modern technology and software to automate the data collection process can improve the accuracy and efficiency of statistics collection. For example, using sensors and tracking devices during matches can provide detailed and reliable data.
- Data quality control: It is important to have mechanisms for quality control of the data collected, including verification of its accuracy, completeness and consistency. Regular auditing and data validation can help eliminate errors and inaccuracies.
- Cooperation with experts: The involvement of experts from the fields of statistics, informatics and sports analysis in the process of data collection and processing can contribute to a better understanding of statistics and the use of optimal analytical tools.

Implementing these steps could improve the statistics and data collection for Czech floorball, supporting a more accurate analysis of player and team performance and providing valuable information for coaches, managers and other stakeholders. In summary, the analysis of the effectiveness of the players and the implementation of the mentioned recommendations can contribute to improving the performance of floorball teams and achieving better results in competitions.

5 Conclusion

The contribution of the work is primarily the demonstration of the use of DEA models in the sports industry and the overall relationship between data mining and sports. Based on the analysis of the efficiency of floorball players using the DEA model, this model provides not only a comprehensive view of the players' performance but also their contribution to the team's overall performance. The results suggest that identifying the best players in a team and their effectiveness can be key factors in team success. Further improvements and adjustments in the team's training methods and strategies can be derived from analysing the DEA model results. From the results, it can be seen that it is better to use the BCC model - more players are efficient; it is also understandable that variable returns of scale are more logical for human activities - each person is an individual, and a person is not a machine and does not have constant performances.

The overall conclusion suggests that analysing the effectiveness of floorball players using the DEA model is useful for evaluating team performance and identifying the best players. The results of this analysis serve as a basis for identifying improvements to the team's training methods and strategies, thereby contributing to the team's overall success in floorball competition. Effective training can help players develop their skills to perform optimally and increase their effectiveness. As part of further research, it would be appropriate to use other variables for inputs and outputs or to measure completely different statistical data that could be focused on a specific game post or specific game situation. For forwards, one could measure turnovers in the zone, the number of blocked shots, or the number of faceoffs won and possessions in the offensive zone. For defenders, it would be the number of blocked shots of the attacking team, the number of balls won in the defensive zone, the number of successfully established attacks that end with a shot on goal, and others.

The expansion of this work is expected in the future. And the use of some other extensions of DEA models (cross efficiency, two-stage DEA, super efficiency, assurance regions, DEA slack based measure etc.). There is also an effort to expand the data - more years for the Czech floorball league, as well as for the women's league. Because like Bhat et al. (2019) it should be noted that this article deals only with big sports and it would be worth analyzing women's floorball as well and comparing whether there are any specifications in it.

References

Banker, R. D., Charnes, A., and Cooper, W. W. (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, **30**, 1078–1092.

Bhat, Z. U. H., Sultana, D., Dar, Q. F. (2019). A comprehensive review of data envelopment analysis (DEA) in sports. *Journal of Sports Economics & Management*, 9(2), 82–109.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research*, **2**, 429–444.

Český florbal. Livesport Superliga: Statistiky. Online, 2024. Dostupné na: https://www.ceskyflorbal.cz/competition/detail/statistics/8XM1. [cit. 2024-04-13].

Denk, M. (2016) Charakteristiky vývoje sportovní kariéry hráčů výkonnostního fotbalu v kontextu rodinného a fotbalového prostředí. Západočeská univerzita v Plzni, Plzeň.

Hůlka, K., Bělka, J. and R. Weisser. (2014) Analýza herního zatížení v invazivních sportovních hrách. Univerzita Palackého v Olomouci.

Kadlec, D. (2022) Principy tréninkových cvičení a her. FAČR.

Pažák, M. (2024) Hodnocení efektivnosti florbalových hráčů pomocí DEA modelů. VŠB-TUO.

Pelloneová, N. and M. Tomíček (2022) Ranking Players by DEA: An analysis of Czech and Danish football. *Studia sportiva*, 16(1), 76–90. https://doi.org/10.5817/StS2022-1-8.

Popelka, J. (2014) Sport jako veřejná služba: Sociální spravedlnost a efektivnost sportovních zařízení v obcích.

Špaček, O. (2016) Sport pro všechny? Sociální nerovnosti a sportovní aktivity. *Sport*, **8**, 53–78. https://doi.org/10.5817/SOC2011-1-53.

Švátora, K. et al. (2022). International comparison of motor abilities and floorball skills in U-11 to U-14 teams from the Czech Republic and Australia. *International Journal of Physical Education, Fitness and Sports*, 11(2), 45–52. https://doi.org/10.34256/ijpefs2226

Tervo, T. and A, Nordström. (2014) Science of floorball: a systematic review. *Open Access J Sports Med*, doi: 10.2147/OAJSM.S60490.

A Two-Stage DEA Model for Evaluating the Efficiency of SMEs with Multi-Year Accounting Data

Lucie Chytilova , Hana Stverkova

Abstract. This study proposes extending the two-stage Data Envelope Analysis (DEA) model to assess the efficiency of small and medium enterprises (SMEs). The model leverages multi-year accounting data and incorporates a temporal dimension to capture the dynamic nature of SME operations. The primary focus is on evaluating these enterprises' stability and efficiency. The proposed model decomposes the evaluation process into two sub-stages: Stage 1 focuses on human capital efficiency, and Stage 2 assesses business efficiency. Outputs from Stage 1 serve as inputs for Stage 2, reflecting the sequential nature of these processes. These phases influence each other, even in different periods. Data from 2020 to 2022 are used. This approach allows for a more comprehensive evaluation by capturing the effectiveness of human capital utilisation (Stage 1) and its subsequent translation into business efficiency (Stage 2). The analysis will categorise SMEs into efficient and inefficient groups, further delving into efficiency levels at each stage. By examining the relationships between human capital, business skills, and overall efficiency, the study aims to identify key drivers of efficiency in SMEs. Finally, based on the findings, the research proposes practical recommendations for enhancing SME operations and developing effective business support mechanisms.

Keywords: Data Envelopment Analysis, two-stage, small and medium business.

JEL Classification: C44
AMS Classification: 90C05

1 Introduction

Small and medium-sized enterprises (SMEs) are a category of businesses characterised by their limited size and scope. They usually have a smaller workforce, lower revenue, and a smaller market share than larger corporations. SMEs represent 99% of all businesses in the EU, and they often try to access to finance and EU support programmes targeted specifically at these enterprises.

The key characteristics of SMEs include:

- Size - They typically have a limited number of employees, usually ranging from a few dozen to a few hundred.
- Revenue - Their revenue is generally lower than that of large corporations.
- Ownership - These firms are usually owned by private individuals or small groups of entrepreneurs, such as family businesses or startups.
- Flexibility - SMEs are often more flexible and can respond more quickly to market changes through innovation and the adoption of new trends.

Most countries consider SMEs to be crucial drivers of economic growth and job creation. They often form the backbone of the local economy and are a source of innovation and entrepreneurial spirit. Both governments and non-profit organisations provide programs, financial incentives, and advice to support the development of SMEs. Therefore, understanding these businesses and how they can be effective is important.

There are various ways to measure the effectiveness of SMEs. Common methods include analysis using financial indicators (such as turnover, gross margin, cost per employee, profitability, and return on investment), operational indicators (such as capacity, employee productivity, throughput, process time, and waste rate), and customer satisfaction measures (such as customer satisfaction surveys, customer feedback ratings, and recommendations). It is essential to select metrics and indicators that align with the specific objectives and characteristics of the SME. Combining different metrics provides a more comprehensive view of the business's effectiveness.

This article continues the work from last year 5 when we chose the data envelopment analysis method for efficiency

Czech University of Life Sciences Prague, Department of Systems Engineering, Kamýcká 129, 165 00 Praha-Suchdol, Czech Republic, chytilova@pef.czu.cz

VŠB-TUO, Department of Systems Engineering, Sokolská třída 2416/33, 702 00 Ostrava - Moravská Ostrava, Czech Republic, hana.stverkova@vsb.cz

measurement and financial indicators, as these are the most readily available. We have been inspired by some publications, such as the article by Zhou, Ang, and Poh 13, which have already linked these topics. One of the first connections can be seen in the work of Wu and Liang 12, who applied DEA to SMEs in China. A more specific focus on the energy industry was then introduced in 2015 by Park and Jeong 11. All of this suggests the potential for use. It is also good to remember that nowadays, many other topics can be solved at MPS, and DEA can be used for this, for example, ecology (affect the environment and society negatively) see 7 or the connection with competitiveness and sustainability see 6. This may be used in future, too.

We based our model on the complexity of the SMEs' business and as the growth of the business is seen. We concluded that the two-phase DEA method published in 2012 by Chen, Cook, and Zhu 4 is suitable. We also referred to the article by Li and Feng 10, where the authors used two phases and two different periods for analysis. However, our primary data source was accounting information, marking the start of long-term research. This single model from the article 5 was rated as very good, so we decided to use yet another timeline to describe this promisingly developing business and be able to recommend improvements based on an analysis of its effectiveness.

The rest of the paper has the following structure: Section Two-stage Data Envelopment Analysis provides brief information about the DEA models. Section 3 - Model and Data defines the concrete model and all used variables and gives information about them. The Results of Analysis section briefly describes this model's results and future use. The conclusion provides some conclusions and remarks.

2 Two-stage Data Envelopment Analysis

DEA (Data Envelopment Analysis) is a method used to evaluate the efficiency and productivity of units based on their inputs and outputs. Two classical models: the CCR (Charnes et al. 3) and the BCC (Banker et al. 1). One variant of DEA is a two-phase DEA, sometimes called network DEA. Färe and Grosskopf 8 were the first to deal with this model. It is an extension where we can combine different inputs and outputs when evaluating the efficiency of units. This makes it possible to model more complex relationships between inputs and outputs and better capture the specifics of units or industries. It helps evaluate units with a more complex structure of inputs and outputs or requires additional adjustments to measure efficiency more accurately. It can compare units' performance in multi-level systems or industries with diverse characteristics.

There are different types of approaches for the network DEA. In this paper, we use the multiplicative method introduced by Kao and Hwang 9.

Consider the basic input oriented CRS DEA models that estimate the 1st stage, 2nd stage and the overall efficiency for the evaluated unit k_0 independently:

- 1st stage:

$$\begin{aligned} \max \quad & e^1_{k_0} = \frac{\sum_{l=1}^{r} w_l z_{lk_0}}{\sum_{i=1}^{m} v_i x_{ik_0}} \\ \text{s.t.} \quad & \frac{\sum_{l=1}^{r} w_l z_{lk}}{\sum_{i=1}^{m} v_i x_{ik}} \le 1 \qquad \text{for } k = 1, \dots, K, \\ & w_l, v_i \ge \epsilon, \end{aligned} \tag{1}$$

- 2nd stage:

$$\begin{aligned} \max \quad & e^2_{k_0} = \frac{\sum_{j=1}^{n} u_j y_{jk_0}}{\sum_{l=1}^{r} w'_l z_{lk_0}} \\ \text{s.t.} \quad & \frac{\sum_{j=1}^{n} u_j y_{jk}}{\sum_{l=1}^{r} w'_l z_{lk}} \le 1 \qquad \text{for } k = 1, \dots, K, \\ & u_j, w'_i \ge \epsilon, \end{aligned} \tag{2}$$

- overall:

$$\begin{aligned} \max \quad & e_{k_0} = \frac{\sum_{j=1}^{n} u_j y_{jk_0}}{\sum_{i=1}^{m} v_i x_{ik_0}} \\ \text{s.t.} \quad & \frac{\sum_{j=1}^{n} u_j y_{jk}}{\sum_{i=1}^{m} v_i x_{ik}} \le 1 \qquad \text{for } k = 1, \dots, K, \\ & u_j, v_i \ge \epsilon, \end{aligned} \tag{3}$$

where k is set of DMUs (DMU$_k$ for $k = 1, ..., K$), let external input, intermediate measures and final output variables data be $X = \{x_{ik}, i = 1, ..., m; k = 1, ..., K\}$, $Z = \{z_{lk}, l = 1, ..., r; k = 1, ..., K\}$ and $Y = \{y_{jk}, j = 1, ..., n; k = 1, ..., K\}$, respectively. Also, v_j for $j = 1, ..., n$ and v_i for $i = 1, ..., m$ be the weights of the i^{th} input variable, w_l for $l = 1, ..., r$ be the weights of the l^{th} intermediate measures on the side of output of the 1st stage, w'_l for $l = 1, ..., r$ be the weights of the l^{th} intermediate measures on the side of input of the second stage, and the j^{th} final output variable, respectively.

To link the efficiency assessments of the two stages, it is universally accepted that the weights associated with the intermediate measures are the same (i.e. $w = w'$), no matter if these measures are considered outputs of the first stage or inputs to the second stage.

In the multiplicative method introduced by Kao and Hwang 9, the overall efficiency and the stage efficiencies of the DMU$_k$ are defined as follows:

$$e^1_{k_0} = \frac{\sum_{l=1}^{r} w_l z_{lk_0}}{\sum_{i=1}^{m} v_i x_{ik_0}}, \quad e^2_{k_0} = \frac{\sum_{j=1}^{n} u_j y_{jk_0}}{\sum_{l=1}^{r} w_l z_{lk_0}}, \quad e^o_{k_0} = \frac{\sum_{j=1}^{n} u_j y_{jk_0}}{\sum_{i=1}^{m} v_i x_{ik_0}}, \tag{4}$$

whereas the decomposition model used is

$$e^o_{k_0} = \frac{\sum_{j=1}^{n} u_j y_{jk_0}}{\sum_{i=1}^{m} v_i x_{ik_0}} = \frac{\sum_{l=1}^{r} w_l z_{lk_0}}{\sum_{i=1}^{m} v_i x_{ik_0}} \cdot \frac{\sum_{j=1}^{n} u_j y_{jk_0}}{\sum_{l=1}^{r} w_l z_{lk_0}} = e^1_{k_0} \cdot e^2_{k_0}. \tag{5}$$

i.e. the overall efficiency is the square geometric average of the stage efficiencies.

Given the above definitions, the model below assesses the overall efficiency of the evaluated unit k_o:

$$\begin{aligned} \max \quad & e^o_{k_0} = \frac{\sum_{j=1}^{n} u_j y_{jk_0}}{\sum_{i=1}^{m} v_i x_{ik_0}} \\ \text{s.t.} \quad & \frac{\sum_{l=1}^{r} w_l z_{lk}}{\sum_{i=1}^{m} v_i x_{ik}} \leq 1 \qquad \text{for } k = 1, ..., K, \\ & \frac{\sum_{j=1}^{n} u_j y_{jk}}{\sum_{l=1}^{r} w_l z_{lk}} \leq 1 \qquad \text{for } k = 1, ..., K, \\ & u_j, v_i, w_l \geq \epsilon. \end{aligned} \tag{6}$$

Notice that the constraints $\frac{\sum_{j=1}^{n} u_j y_{jk}}{\sum_{i=1}^{m} v_i x_{ik}} \leq 1$, $k = 1, ..., K$, are redundant and, thus, omitted. Model (6) is a fractional linear program that can be modelled and solved as a linear program by applying the Charnes and Cooper 2 transformation as usual.

3 Model and Data

The two-stage DEA model for the SME market was established after a long analysis with academic and business staff. The general model with notation t a $t + 1$ is in Figure 3 and is used for both periods.

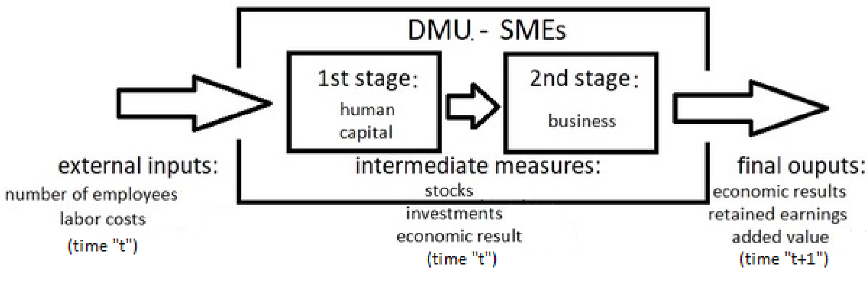


Figure 1 Complete model of the SMEs Business

The model includes the first part, which deals with the efficiency of financial management of human capital, and the second part, which deals with the efficiency of the enterprise. As written, SMEs are small regarding number of employees, turnover, independence, and flexibility. Therefore, the following variables were important to analysis:

- external inputs in time period t:
  - number of employees (number) - if the number of employees is appropriate, the business can function appropriately and save money.
  - labour costs (CZK) - appropriate employee salaries will help support businesses (people's ideas and growth) and save money for future investments.
- intermediate measures in time period t:
  - stocks (CZK) - on the output side, companies want to have as many stocks as possible to fight for the next years; on the other hand, from the input side, and accounting point of view, companies want to reduce it as much as possible, so that it does not overburden us with storage.
  - investment (CZK) - on the output side, the company wants to increase investment so that companies can grow, innovate and develop. On the input side, the company sees it as wanting to reduce investment to save money and restructure the business, but at the moment, it is because of the economic downturn.
  - economic result (CZK) - companies usually aim to increase economic results, so obviously, it is an output. Still, we have also included it as an input, and we want further economic growth to be greater than the previous period.
- final outputs in time period $t + 1$
  - economic result (CZK) - companies usually aim to increase economic results forever.
  - retained earnings - is usually the goal of companies.
  - added value (CZK) - is a goal of many companies because it can contribute to the financial stability and growth of the company (but not much used today).

For the analysis, 20 same firms from the same environment (hospitality) were selected from 2020 and 2022. The descriptive data can be seen in Table 1. Since some of the data were negative or null, there was still some adjustment with the DEA method, as usual.

A model for 2020 - 2021								
id	i1	i2	s1	s2	s3	o1	o2	o3
name of variable	number of employees	labor costs	stocks	invest-ments	economic result	economic result	retained earnings	added value
average	32.45	16260.35	13413.20	1516.50	2561.55	5366.70	25563.55	17382.40
max	105	63868	104321	11274	27471	71983	89226	63846
min	2	853	0	67	-92	-2614	-331	1992
std.dev	26.58	17327.05	25375.09	2439.00	6076.96	16167.82	26499.66	14804.06
A model for 20221 - 2022								
id	i1	i2	s1	s2	s3	o1	o2	o3
name of variable	number of employees	labor costs	stocks	invest-ments	economic result	economic result	retained earnings	added value
average	33.75	19234	11009.2	347.03	5366.70	1023.07	25563.55	17382.4
max	107	81439	69251	2308.37	71983	14396.6	89226	63846
min	3	959	34	1.10	-2614	-522.8	-331.00	1992
std.dev	28.29	22001.88	15611,22	505.26	16167.82	3159.43	25828.68	14429.21

Table 1 SMEs in the Czech Republic in the time period 2020 - 2022

According to the picture and time. Logically, variable o1 from 2020 - 2021 is similar to s3 from 2021 - 2022.

Note that if this analysis were done for other sizes of companies, the variables in the study would also need to be changed. For example, in the case of large enterprises, it could add the amount of innovation or money provided for innovation. Furthermore, there should be variable economies of scale, which MPS usually fails to do, but a healthy large firm has.

4 Results of Analysis

Table 2 shows the result of all the analyses. We can see there:

- e^1 - efficiency of the 1st stage (equation (1)) - *human capital efficiency*,
- e^2 - efficiency of the 2nd stage (equation (2)) - *business efficiency*,
- e - efficiency of the overall model (equation (3)) - *SMEs efficiency*,
- $e^1 \cdot e^2$ - efficiency of multiplication of two stages (equation (5)) - *SMEs multiplication efficiency*,
- e^o - efficiency of the overall multiplicative model (equation (6)) - *SMEs multiplicative efficiency*.

	A model for 2020 - 2021					**A model for 2021 - 2022**				
	e^1	e^2	e	$e^1 \cdot e^2$	e_o	e^1	e^2	e	$e^1 \cdot e^2$	e_o
DMU01	0.8245	0.5454	0.7122	0.4497	0.2928	1.0000	0.9797	1.0000	0.9797	0.8872
DMU02	0.8482	0.3959	0.3944	0.3358	0.2057	0.8173	0.9328	0.9698	0.7624	0.4243
DMU03	0.9176	1.0000	1.0000	0.9176	0.3881	1.0000	1.0000	1.0000	1.0000	1.0000
DMU04	0.4354	0.6606	0.4582	0.2876	0.1393	0.4059	0.5798	0.3320	0.2353	0.2197
DMU05	0.1892	1.0000	0.2467	0.1892	0.0537	0.0120	1.0000	0.0540	0.0120	0.0020
DMU06	0.3200	0.3707	0.2498	0.1186	0.0905	0.4690	0.0583	0.3927	0.0273	0.1046
DMU07	0.5399	0.5085	0.2812	0.2745	0.1717	0.1246	0.8572	0.4968	0.1068	0.1221
DMU08	1.0000	1.0000	1.0000	1.0000	0.7628	0.5441	0.3051	0.4222	0.1660	0.3187
DMU09	0.5800	0.7154	0.3866	0.4249	0.3087	0.4915	0.0098	0.2111	0.0048	0.1360
DMU10	0.2827	0.9314	0.2687	0.2633	0.0691	1.0000	0.0020	1.0000	0.0020	1.0000
DMU11	0.8534	0.4575	0.3560	0.3904	0.1998	0.3233	0.2623	0.7350	0.0848	0.1444
DMU12	0.6355	1.0000	1.0000	0.6355	0.1445	0.2359	0.4280	1.0000	0.1010	0.1843
DMU13	0.6657	0.4270	0.3242	0.2843	0.2173	1.0000	1.0000	0.2859	1.0000	0.1010
DMU14	0.5763	0.4672	0.4157	0.2692	0.2902	0.0888	0.7815	0.1834	0.0694	0.0888
DMU15	0.7456	0.6827	0.3351	0.5090	0.1792	0.3499	0.0600	1.0000	0.0210	0.1279
DMU16	0.6825	0.5407	0.3450	0.3690	0.1640	0.3015	0.0464	0.8258	0.0140	0.0890
DMU17	0.4885	1.0000	0.7363	0.4885	0.1830	0.1942	0.4315	0.3283	0.0838	0.1902
DMU18	0.8280	0.3164	0.2183	0.2620	0.1990	0.2365	0.0549	1.0000	0.0130	0.0639
DMU19	0.3284	0.3659	0.2487	0.1202	0.0927	0.4133	0.0516	0.3671	0.0213	0.1033
DMU20	1.0000	0.3006	0.4671	0.3006	0.2913	0.0077	1.0000	1.0000	0.0077	0.0013
average	0.6371	0.6343	0.4722	0.3940	0.2222	0.4508	0.4921	0.6302	0.2356	0.2655
std.dev	0.2429	0.2628	0.2659	0.3678	0.1541	0.3393	0.4114	0.3490	0.3668	0.3163
# of efficient	2	5	3	1	0	4	4	7	2	2

Table 2 Results of efficiency for each stage and overall efficiency

When it came to comparing the three different "overall" effectiveness, it can be said that the "most positive" model (the model with the highest effectiveness) in both periods is the *SMEs efficiency* model (e) compared to others. This model does not deal with the middle part - the black box. This looks nice for analysis, but this model does not give us a closer look at the matter, so it can distort the results. It certainly does not provide such immediate information and can not help the managers understand the business's problems.

When the models *SMEs multiplication efficiency* ($e^1 \cdot e^2$) and *SMEs multiplicative efficiency* (e^o) are compared, it can be seen that *SMEs multiplication efficiency* is more "positive" in the first period and similar in the second period. Generally, these results are much smaller than the *SMEs efficiency* model. It may look hard to decide if the results of *SMEs multiplication efficiency* or *SMEs multiplicative efficiency* are better - they are similar. One way how to decide may be to compare results with other analyses. But logically, multiplication is not what works in business. Business is more connected, so the *SMEs multiplication efficiency* where the connection between both parts is given more seems more reasonable.

Big differences can be seen in the efficiency of individual phases compared to the overall one. A closer analysis was carried out, which is not presented due to the article's length. But we will mention a few points of interest and a few recommendations. The problem is mostly in the first part of the first and second periods rather than the second. So, it is possible that the entrepreneurs themselves sensed that there was a problem in human capital and solved it. Still, unfortunately, their focus was so great that they forgot to solve the business. It is recommended that you try to understand the entire business comprehensively. Interestingly, the DMU03 unit appears efficient in all models in the second period. In the first period, the efficiency was quite high, except for the *SMEs multiplicative efficiency* model.

Based on the results and closer analysis, we are providing information and recommendations for the optimisation:

- Optimizing the number of employees is important and beneficial for many reasons, especially for the success and efficiency of businesses. Some of the main reasons why it is good to optimise the number of employees are to reduce costs, increase efficiency and flexibility, focus more on core competencies, and improve company culture and sustainability. However, it is important to note that the proper optimisation of the number of employees should be done regarding social responsibility and the preservation of human rights.
- Optimizing labour costs - e.g., changing additional payments, increasing standards, etc. Here it is, but you have to be careful - costs per person even after termination (2 months' notice + 3 months' severance pay), i.e. if people are fired in the second half of the year, they may not be full effect in the following year. In general, labour costs do not affect the warehouse. On the other hand, investments can affect labour costs; if we invest, we try to reduce the cost of labour cost, thanks to the investment.
- Reducing inventory means freeing up cash that can be used elsewhere. It can be invested to bring additional profit or financial gain, which can be made from inventory optimisation, an important element of successful supply chain management and warehousing. Proper inventory optimisation allows businesses to minimise costs, maximise efficiency, and achieve higher customer satisfaction. Some of the key reasons for inventory optimisation are cost reduction, risk reduction, improved cash flow, increased availability of goods, improved supply chain efficiency, and better decision-making. Companies may or may not aim to maximise their economic results. Some companies try to optimise their economic results so that the level of taxation is adequate. Profit maximisation and investment are interrelated concepts that affect a business's financial performance and long-term success.
- Overall, the optimal relationship between maximum economic output and investment lies in strategic planning and the right balance between generating a profit and investing it in the long-term growth and development of the business.

5 Conclusion

This article continues research on measuring the efficiency of small and medium enterprises using the two-stage method. This method is still only necessary to choose a suitable potential and model.

The results show that the *SMEs multiplication efficiency* model is the best for overall efficiency - it does not give the best results. Till, it best portrays the situation and business connections. In the future, it would be advisable to examine SMEs for a longer period of time because even a company can be in different stages of its existence, which will affect the results. It also revealed that most firms seem unable to handle human capital and entrepreneurship simultaneously. However, this may also be a problem of the economic cycle, which should be investigated more or include a longer period of time for investigation.

When evaluating effectiveness, there is an urgent need to focus more on the model's use and for whom and what the analysis can be used. If it is a more comprehensive analysis, models that work with the efficiency of individual phases would be appropriate. It's crucial to consider the best method and whether the process/method suits the specific industry. That is why we are considering doing three-stage/net DEA instead of two-stage DEA models over a longer period. We are also looking to use other data models, such as the additive method, the solution of negative variables, different returns to scale, etc.

References

[1] Banker R.D., Charnes A. & Cooper W.W. (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manage Science*, 30, 1078—1092.

[2] Charnes A. & Cooper W.W. (1962) Programming with linear fractional functional. *Naval Res Logist*, 9, 181–185.

[3] Charnes A, Cooper W.W. & E. Rhodes. (1978) Measuring the efficiency of decision-making units. *European journal of operational research*, 2, 429–444.

[4] Chen, Y., Cook, W.D. & Zhu, J. (2012). Deriving the DEA frontier for two-stage processes. *Omega*, 40(5), 611–618.

[5] Chytilova, L. & Stverkova, H. (2023). Application of the Two-Stage DEA Model in SMEs Business In: 41th International Conference Mathematical Methods in Economics 2023: Proceedings: Prague University of Economics and Business, Faculty of Informatics and Statistics: 13 - 15 September 2023, Prague, Czech Republic. 2023. s. 134-140. ISBN 978-80-11-04132-8

[6] De, D., Chowdhury, S., Dey, P.K. & Ghosh, S.K. (2020) Impact of Lean and Sustainability Oriented Innovation on Sustainability Performance of Small and Medium Sized Enterprises: A Data Envelopment Analysis-based framework, *International Journal of Production Economics*, 219, 416–430. https://doi.org/10.1016/j.ijpe.2018.07.003.

[7] Dey, P.K., Yang, Gl., Malesios, C. et al. (2021) Performance Management of Supply Chain Sustainability in Small and Medium-Sized Enterprises Using a Combined Structural Equation Modelling and Data Envelopment Analysis. *Comput Econ*, 58, 573–613. https://doi.org/10.1007/s10614-019-09948-1.

[8] Färe R. & Grosskopf S. (1996) Productivity and intermediate products: a frontier approach. *Econ Lett*, 50, 65–70.

[9] Kao C. & Hwang S.N. (2008) Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. *European journal of operational research*, 185, 418–429.

[10] Li, Y. & Feng, L. (2017). Efficiency evaluation of small and medium-sized enterprises (SMEs) in China based on a novel network DEA model. *Sustainability*, 9(12), 2222.

[11] Park, K. H. & Jeong, B. (2015). The application of data envelopment analysis for evaluating efficiency of small and medium-sized enterprises in the energy industry. *Energy Policy*, 87, 429–438.

[12] Wu, J. & Liang, L. (2012). Efficiency evaluation of small and medium-sized enterprises (SMEs) in China using data envelopment analysis (DEA). *African Journal of Business Management*, 6(4), 1505–1514.

[13] Zhou, P., Ang, B.W. & Poh, K.L. (2008). A survey of data envelopment analysis in energy and environmental studies. *European journal of operational research*, 189(1), 1–18.

Portfolio Analysis: Exploring Rank Length Metrics

Tereza Čapková[1]

Abstract. Portfolio analysis is a crucial aspect of financial management, with numerous specialists continually seeking to develop novel approaches that may enhance decision-making methods. This study investigates the application of Extreme Rank Length (ERL) and Continuous Rank Length (CRL) metrics as alternative approaches for assessing portfolios, differing from traditional optimization methods. The motivation for this work stems from the robustness of stochastic dominance. To determine the effectiveness of ERL and CRL in evaluating portfolios, we simulate portfolios using the eleven most active stocks by dollar volume. The performance of these portfolios is evaluated using ERL and CRL metrics. Our research opens the door for further investigation and development in financial analysis by highlighting the potential of these metrics in portfolio evaluation.

Keywords: Continuous Rank Length, CRL, ERL, Extreme Rank Length, Financial Analysis, Financial Management, GET package, Portfolio Performance

JEL Classification: G17
AMS Classification: 91G10

1 Introduction

In modern financial analysis, evaluating portfolio performance is a critical aspect that aids investors in making informed decisions. Traditional portfolio evaluation methods, such as the mean-variance optimization introduced by Markowitz (1952), focus on balancing expected return against risk, typically measured by standard deviation. However, these methods often fail to capture the nuances of portfolio behavior under varying market conditions, particularly during periods of high volatility or economic upheaval. Because of that, we are motivated to explore new methods or use metrics that were not used before in the context of finding the optimal portfolio.

There are many methods for evaluating portfolios and choosing the optimal one. This research is widely inspired by stochastic dominance (Levy, 1990). Stochastic dominance has been extensively applied in portfolio analysis, offering a robust framework for comparing the entire return distributions of different assets. Motivated by the comprehensive stochastic dominance framework, our research explores another ranking method: Extreme Rank Length (ERL) and Continuous Rank Length (CRL).

2 Methodology

Throughout the research, we utilized the package GET (Myllymäki & Mrkvička, 2020). First, let us define the variables and metrics for the portfolio analysis.

2.1 Variables and metrics

Return

The return on an asset is a critical measure in financial analysis, defined as the percentage change in the asset's value over a specified period. To calculate the return, we consider the asset's end-of-period and beginning-of-period prices. Specifically, when dealing with stocks, we utilize the adjusted closing prices from the first and last days of the period under examination. For this case study, we used simple returns, as we deemed this approach adequate for our analysis. Simple returns are straightforward to interpret, making it suitable for our purposes.

Suppose we have stock i, where $i = 1,\ 2, \dots, n$ and we have adjusted closing stock prices in period t, where $t = 1, 2, \dots m$, as initial values IV_{it} of stock i at the start of period t and final value FV_{it} of stock i at the end of period t. The simple return of one stock in one period R_{ti} is calculated using the following formula. To determine the return of a specific portfolio CR_t for period t, the individual simple returns of stocks within the portfolio need to

[1] University of South Bohemia/Faculty of Economics, Department of Applied Mathematics and Informatics, Studentská 13, 370 05 České Budějovice 5, capkot00@jcu.cz.

be weighed and summed. Assume we have n stocks and let w_i represent the weight of stock i in the portfolio. The weighted return of the portfolio in period t can be calculated using the following formula.

$$R_{it} = \frac{FV_{it}}{IV_{it}} - 1 \qquad CR_t = \sum_{i=1}^{n} w_i R_{it}$$

Risk

To measure risk, we employed the standard deviation (SD) of adjusted returns in a given quarter. We used complete daily data from that quarter to calculate the SD for each quarter. This approach provides a more granular and accurate measure of volatility within each period, capturing the daily fluctuations in returns and ensuring a comprehensive risk assessment. We also needed to use the inverse of the SD for every quarter because our software tools only allow us to maximize or minimize functions when combined. Therefore, we decided to use the inverse SD to facilitate this requirement.

Extreme rank length metric

Both metrics under investigation are primarily employed to identify outliers within a set of functions. Their prime application is in the global envelope method, where they identify significantly different functions from others. The first metric is Extreme Rank Length (ERL) (Mrkvička et al., 2020; Myllymaki et al., 2017; Narisetty & Nair, 2016).

The bounding curves define pointwise envelopes:

$$T_{low}^{(k)}(r) = \min_{i=1,\dots,s+1}{}^{k}\, T_i(r) \qquad T_{upp}^{(k)}(r) = \max_{i=1,\dots,s+1}{}^{k}\, T_i(r) \quad \text{for } r \in I$$

where min^k and max^k denote the kth smallest and largest values respectivetly, and $k = 1,2,\dots,[(s+1)/2]$ where s is fixed number of simulations and k is fixed probability of the type error I. The interval I is an interval of distances, in our case interval of quarters. Function T_i in our case is performance of portfolio i.

$$R_i \coloneqq \max\{k: T_{low}^{(k)}(r) \le T_i(r) \le T_{upp}^{(k)}(r) \text{ for all } r \in I\}$$

The value R_i, which we call the extreme rank, is a depth measure that represents the apparent ‘extremeness’ of the curve T_i in the bundle of functions $T_i \dots, T_{s+1}$. Surely, in our case we will be using only one-way ERL, because we want function that is the highest. For a better understanding see Section 4.1 of Myllymaki et al., 2017.

ERL computes pointwise ranks and evaluates the duration of the most extreme rank for each function. It orders functions from the most to least extreme based on this duration. However, a function that reaches a high rank but then drops significantly can still be evaluated as the most extreme, which in our case could be undesirable due to biasing this metric by the high fluctuation of portfolio functions. Because of this, we assume that ERL will not be the best metric to use in portfolio analysis.

Continuous rank length metric

The Continuous Rank Length (CRL) (Hahn, 2015; Mrkvička et al., 2022) metric is a tool used to evaluate function behavior continuously. See Section 2.3.2 in Mrkvička et al., 2022, for the original formula. CRL evaluates the ranks of functions over a continuous interval. Specifically, it calculates how frequently each function achieves specific ranks throughout the interval, forming a vector of rank lengths. This metric induces a natural ordering where functions with frequent low ranks are considered more extreme.

To better capture the portfolio behavior, we modified the original formula of CRL by switching minimum with a sum,

$$c_i = \frac{1}{s+1} \sum C_i(r) \quad \text{with} \quad C_i(r) = s + 1 - c_i(r)$$

Where c_j is the continuous poinwise rank that is refinement of R_i. For better, understanding see Section 2.3.2 in Mrkvička et al., 2022.

By summing the ranks, the metric accounts for the total performance across the interval rather than focusing on the minimum rank. This provides a more comprehensive view of the function's behavior. Also, this sum-based metric is less sensitive to extreme fluctuations. A single low rank (which would dominate the minimum-based metric) will not overshadow the overall performance in the sum-based metric. This is extremely helpful when evaluating portfolios because functions (portfolios) that maintain good ranks consistently will be favored. At the same time, those with significant drops will be penalized, addressing the issue of temporary extremes impacting the evaluation with ERL.

2.2 Case study

As a first step, we chose eleven stocks to include, which is explained further. Secondly, we simulated random portfolio weights for our eleven selected stocks. To showcase and compare these metrics, we first simulated 10^5 random portfolios. Then, we iterated on the method for 10^6 portfolios to show its robustness. The simulation for every iteration was performed only once because the objective was to generate random portfolios for analysis, not to optimize the portfolio composition. In future research, the use of Monte Carlo simulations is recommended.

Data

The data was collected using a quantmod package (Ryan & Ulrich, 2024) which retrieves financial data from Yahoo Finance.

Determining which stocks to include was a critical section of our research. To address potential confounding factors arising from the COVID-19 pandemic, we utilized data from 2008 to 2018. This decision guarantees that our initial implementation of the technique is founded on more foreseeable and consistent data. Analyzing the method's performance on unpredictable data, such as during the COVID-19 pandemic, will be the focus of our subsequent research. Our analysis is performed on quarterly periods. As for choosing the stocks, we focused on the most active stock by dollar volume (Nasdaq, 2024), assuming that higher trading volumes would imply lower risk and more stable returns. The assumption is based on the premise that actively traded stocks are stable and less susceptible to drastic price swings due to their liquidity. Some stocks needed to be eliminated because they began trading during our chosen timeframe and did not have a complete dataset — specifically, TSLA, META, and ARGO.

Ticker	Name	Mean	Min	Max	SD	Median
NVDA	Nvidia Corporation	7,52	-42,88	56,17	22,54	9,95
AMD	Advanced Micro Devices, Inc.	6,04	-54,43	81,63	33,34	0,32
APPL	Apple Inc.	5,63	-34,93	45,80	16,27	8,19
SMCI	Super Micro Computer, Inc.	4,58	-30,27	51,98	23,12	-1,27
MSFT	Microsoft Corporation	3,23	-26,09	25,20	12,21	3,58
MSTR	MicroStrategy Incorporated	2,60	-34,71	52,18	20,32	1,14
AMZN	Amazon.com, Inc.	7,54	-26,30	47,76	16,78	9,70
ADBE	Adobe Inc.	4,28	-44,15	28,87	14,36	6,80
GOOGL	Alphabet Inc. (Class A)	3,37	-35,72	30,35	15,45	3,84
GOOG	Alphabet Inc. (Class C)	3,37	-35,72	30,35	15,50	2,93
MU	Micron Technology, Inc.	8,08	-38,60	58,30	27,83	6,87

Table 1 Basic Characteristics of Stock Returns Calculated from Quarterly Data

3 Results

First, let us assume that our objective is to identify the optimal portfolio. Within this framework, our objective is to identify a function that optimizes the generation of profits while simultaneously reducing the potential for loss. This entails identifying a portfolio that maximizes returns while maintaining a specific level of risk or minimizing risk while achieving a specific level of returns in our simulated portfolios using ERL and CRL metrics.

Table 2 displays the stock weights in selected portfolios based on the ERL metric. The noteworthy aspect is the reduction of stocks GOOGL and GOOG, which do not have a high average return (Table 1). Stock MU is also being reduced, but from Table 1, we do not see any specific characteristic explaining why, and that is because the basic characteristics do not tell us much about the function's behavior.

ID	NVDA	AMD	APPL	SMCI	MSFT	MSTR	AMZN	ADBE	GOOGL	GOOG	MU
■ 83684	0,27	0,13	0,32	0,03	0,05	0	0,08	0,01	0,06	0,05	0
■ 66694	0,27	0,06	0,06	0,04	0,05	0,35	0,04	0,02	0	0,01	0,1
■ 76603	0,09	0,14	0,16	0,12	0,09	0,01	0,11	0,15	0,06	0,07	0

Table 2 Optimal portfolios weights according to our simulation (ERL) 10^5

Figure 1 showcases the three selected portfolios by ERL metric in our initial simulation, and the basic characteristics of these portfolios can be found in Table 3. From the graph, it is hard to understand why these portfolios were selected because the functions are not extreme on the whole interval, but the extremeness is based on the metrics.

Notice how the second portfolio, 66694 (blue), differs from the others in the graph and its basic characteristics. This portfolio showcases that ERL has a problem with portfolio analysis, which we expected. This portfolio was evaluated as a good one probably because of the extremeness of returns, even though it is one of the worst in SD. What we mean can be seen most significantly between the 13th and 16th quarters, but it can be observed many times in its behavior. Based on this observation, we deem ERL not ideal for portfolio analysis because 66694 is not the best due to its high risk and was selected just because, in some intervals in return, it is the most extreme.

ID	Mean (return)	SD (return)	Min (return)	Max (return)	Mean (risk)
■ 83684	4,62	13,52	-29,38	31,32	0,64
■ 66694	3,13	18,38	-34,36	45,83	3,07
■ 76603	4,41	14,75	-34,61	40,41	0,61

Table 3 Basic characteristics of selected portfolios (ERL) 10^5

Figure 1 Three portfolios having the most extreme curves (ERL metric) 10^5

Let us look at the other metric. Table 4 displays the portfolios selected based on the CRL metric. These portfolios differ significantly from those selected by ERL and exhibit substantial variability. This implies that employing this approach may give us a wide range of portfolios that exhibit strong performance.

ID	NVDA	AMD	APPL	SMCI	MSFT	MSTR	AMZN	ADBE	GOOGL	GOOG	MU
■ 24958	0,1	0,15	0,25	0,14	0	0,01	0,24	0,01	0	0,06	0,04
■ 86661	0,23	0,11	0,01	0,21	0,05	0,01	0,17	0,07	0	0,02	0,12
■ 25958	0,17	0,15	0,16	0,1	0,02	0,02	0,17	0,01	0,05	0,08	0,07

Table 4 Optimal portfolios weights according to our simulation (CRL) 10^5

Let us examine these portfolios. Figure 2 showcases the three portfolios with the most extreme curves determined by the CRL metric. Table 5 shows the basic characteristics of these portfolios. At first glance, it is evident that we obtain more consistent outcomes compared to the ERL case. The distributions exhibit a high degree of similarity, especially in the behavior of a return function, but it can be seen even from basic characteristics. The portfolio with the 24958, which is highlighted in dark blue, consistently ranks in the top half of the functions in both graphs. It also consistently ranks at or near the top in the risk. Portfolio 86661 (green) is also interesting because if we looked solely at the return, we would probably deem it better than the blue one (see the functions between quarters

20 and 24; overall, the function is higher many times). However, because of its higher risk, as seen by the lower function in the SD graph, it was not selected as best.

ID	Mean (return)	SD (return)	Min (return)	Max (return)	Mean (risk)
■ 24958	4,71	13,41	-32,26	37,99	0,65
■ 86661	4,54	13,74	-35,44	33,87	0,81
■ 25958	4,27	13,54	-32,11	36,62	0,69

Table 5 Basic characteristics of selected portfolios of our simulation (CRL) 10^5

This observation supports our assumptions before the research. In the future, we will only utilize the CRL metric for analysis. However, it is premature to determine its effectiveness at this research stage.

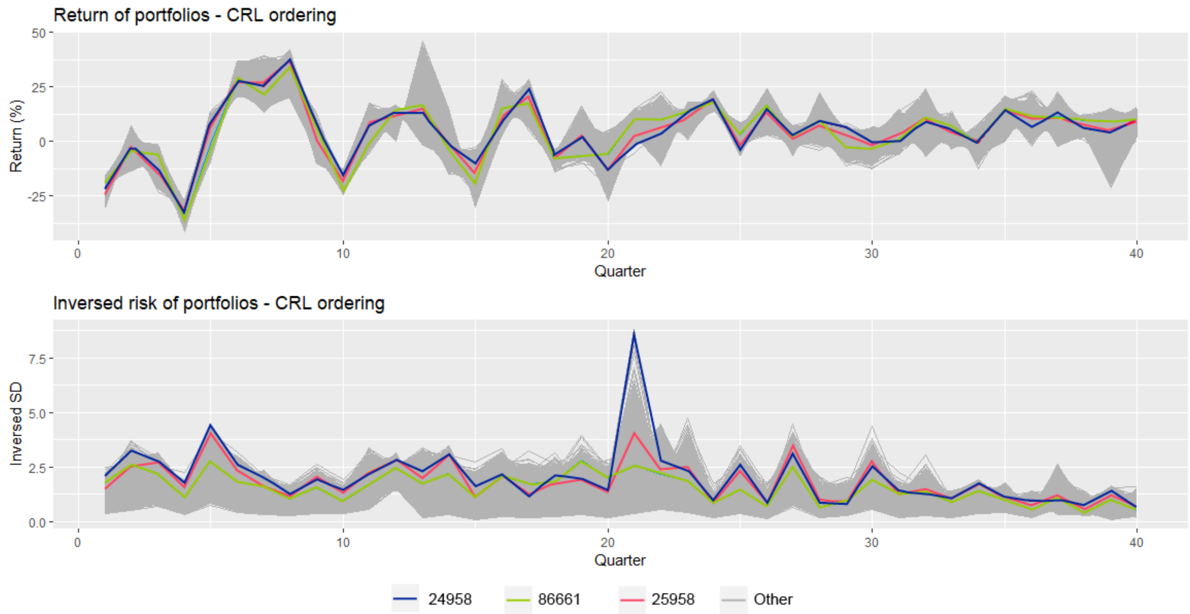


Figure 2 Three portfolios having the most extreme curves (modified CRL metric) 10^5

To examine the robustness of the CRL metric, we simulated more portfolios, specifically 10^6. We aimed to determine whether any selected portfolios would maintain their position in the top 3 or at least the top 10. Table 6 displays the selected portfolio with an expanded simulation. Our previously successful portfolio, 24958, nevertheless secured a position in the top 3. The second number, 86661, is not displayed because it is in the fourth position. This indicates that the robustness of this method is likely satisfactory.

ID	NVDA	AMD	APPL	SMCI	MSFT	MSTR	AMZN	ADBE	GOOGL	GOOG	MU
■ 229042	0,15	0,06	0,06	0,21	0,04	0	0,17	0,07	0,03	0,02	0,19
■ 656080	0,09	0,17	0,15	0,17	0	0,01	0,21	0	0,03	0,13	0,04
■ 24958	0,1	0,15	0,25	0,14	0	0,01	0,24	0,01	0	0,06	0,04

Table 6 Optimal portfolios weights according to our simulation (CRL) 10^6

The simulation produced two new portfolios for us. We have listed the basic characteristics of these portfolios in Table 7. Upon examining Figure 3, it becomes evident that the functions are nearly identical in returns within certain intervals but very different regarding SD. The selected portfolios differ; however, they all share the characteristic of having at least one stock eliminated.

ID	Mean (return)	SD (return)	Min (return)	Max (return)	Mean (risk)
■ 229042	4,65	13,58	-34,95	32,28	0,79
■ 656080	4,3	13,36	-31,36	39,62	0,65
■ 24958	4,71	13,41	-32,26	37,99	0,65

Table 7 Basic characteristics of selected portfolios of our simulation (CRL) 10^6

Many interesting observations can be seen in Figure 3, but we again found quarters 20 to 24 to showcase best how the metric works. The orange portfolio 229042 is the highest of all three but the lowest (worst) for SD. Then blue portfolio 24958 is the best in SD but the worst of the three in return. Remember that the metric works on the whole interval since it is continuous, so these peaks are compensated by the function behavior somewhere else in the interval. This method depends on the selected interval, so the results would be completely different if we cut the first ten quarters. The aim of the following research should be to test this metric on the following data and determine how to stabilize this method.

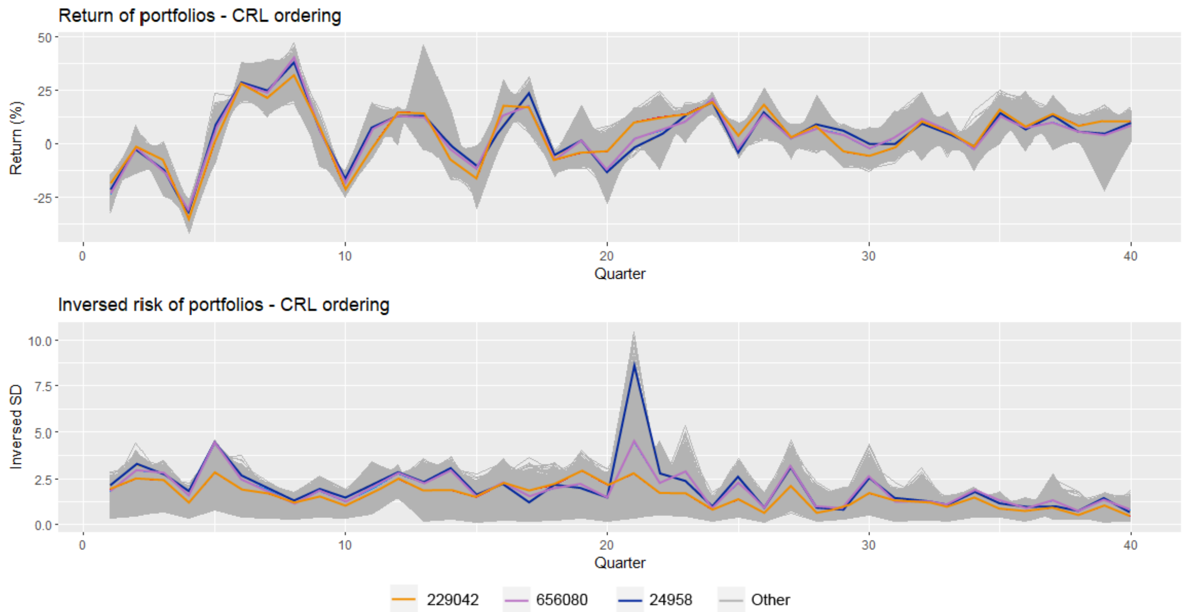


Figure 3 Three portfolios having the most extreme curves (modified CRL metric) 10^6

4 Conclusion and discussion

In this study, we investigated and showed the application of Extreme Rank Length (ERL) and Continuous Rank Length (CRL) metrics for portfolio evaluation. Our findings reveal several key insights and open new avenues for further research.

The results from the ERL metric showed that the portfolios deemed optimal often did not have the best return characteristics nor the lowest mean risk. Relying solely on ERL may not provide a comprehensive assessment of portfolio performance because it computes pointwise ranks and evaluates the duration of the most extreme rank for each function, which does not penalize fluctuation. The significant differences in portfolio weights and the inconsistencies in risk profiles underscore the unfitness of the ERL metric for portfolio analysis. The CRL metric provided more consistent outcomes, as evidenced by the more uniform distributions. The portfolios selected based on CRL exhibited high similarity and consistently maintained rankings across different simulated portfolios. This suggests that the CRL metric may offer a more reliable measure of portfolio performance, particularly in capturing the stability and resilience of portfolios over time.

Despite the promising results, our research is still in its early stages, and several key questions remain unanswered. Comparing it to other portfolio optimization methods should be our next research aim. One of our favorite aspects of this method is that the portfolio's performance is visible on a graph. This allows us some leeway when comparing it to other approaches, as we can always plot the results of the other method. Future research should also explore the performance of these metrics under highly volatile conditions, such as those induced by the COVID-19 pandemic. Additionally, incorporating time-weighted adjustments to the metrics could provide a more accurate reflection of current market conditions. An additional application worthy of investigation is the examination of the preferred stocks within the generated portfolios. We could potentially construct a great portfolio by selecting and combining preferred stocks with others.

In conclusion, while CRL metrics show potential in portfolio evaluation, further investigation and refinement are necessary. Our findings suggest that the CRL metric warrants deeper exploration for its robustness and applicability in diverse market scenarios and should be compared with other methods. By continuing to develop and test these metrics, we aim to contribute to advancing financial analysis and portfolio management.

Acknowledgments

The Grant Agency of South Bohemia University supported this work as an individual grant project under the number 031/2024/S. I want to express my sincere appreciation to my colleague Jan Fesl for providing the essential computational resources.

References

Levy, H. (1990). Stochastic Dominance. In: Eatwell, J., Milgate, M., Newman, P. (eds) Utility and Probability. The New Palgrave. Palgrave Macmillan, London. https://doi.org/10.1007/978-1-349-20568-4_34

Hahn, U. (2015). "A Note on Simultaneous Monte Carlo Tests." Technical report, Centre for Stochastic Geometry and Advanced Bioimaging, Aarhus University.

Markowitz, H. (1952). "Portfolio Selection." The Journal of Finance, 7(1), 77-91.

Mrkvička T., Myllymäki M., Jílek M., Hahn U. (2020). A one-way ANOVA test for functional data with graphical interpretation, Kybernetika, 56/3 432-458. doi: 10.14736/kyb-2020-3-0432

Mrkvička, T., Myllymäki, M., Kuronen, M., & Narisetty, N. N. (2022). "New Methods for Multiple Testing in Permutation Inference for the General Linear Model." Statistics in Medicine, 41(2), 276–297. doi:10.1002/sim.9236.

Myllymäki, M., & Mrkvička, T. (2020). "GET: Global Envelopes in R." arXiv:1911.06583 [stat.ME].

Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H., & Hahn, U. (2017). "Global Envelope Tests for Spatial Processes." Journal of the Royal Statistical Society B, 79, 381–404. doi:10.1111/rssb.12172.

Narisetty, N. N., & Nair, V. J. (2016). "Extremal Depth for Functional Data and Applications." Journal of the American Statistical Association, 111(516), 1705–1714. doi:10.1080/01621459.2015.1110033.

Nasdaq. (2024). "Most Active Stocks Today." Retrieved March 20, 2024, from https://www.nasdaq.com/market-activity/most-active

Ryan JA, Ulrich JM (2024). _quantmod: Quantitative Financial Modelling Framework_. R package version 0.4.26, <https://CRAN.R-project.org/package=quantmod>.

Wavelet Method for Pricing One-Stage Expansion Options under Stochastic Volatility

Dana Černá[1]

Abstract. One-stage expansion options are real options that enable expanding an investment project by a predetermined factor at a certain cost on a specified future date. Assuming that the price of the underlying commodity and its volatility follow a geometric Brownian motion, the valuation model for expansion options is represented by several partial differential equations. This paper aims to introduce a pricing model for one-stage expansion options, propose an efficient wavelet-based method for its numerical solution, and implement the method to address practical problems. The method employs the Crank-Nicolson scheme extended by Richardson extrapolation in conjunction with the wavelet-Galerkin method. Its application to a benchmark problem within the iron-ore mining industry demonstrates the suitability and applicability of the method and highlights its numerical advantages, which are high-order convergence and a small number of iterations necessary to attain the desired accuracy.

Keywords: real option, option to expand, stochastic volatility, cubic spline, wavelet

JEL Classification: C44, G13
AMS Classification: 65M60, 65T60, 35Q91, 91G60

1 Introduction

Real options theory is analogous to financial options theory because real options provide managers with the right, but not the obligation, to make investment decisions based on future market conditions, see (Dixit and Pindyck, 1994; Haque et al, 2014; Myers, 1977). Real options include options to expand, defer, modify, or abandon investments in response to market developments, enhancing the overall value and flexibility of the investment. Expansion options are real options that enable the expansion of an investment project in the future if conditions are favorable. We consider a project tied to the production of a certain commodity whose value depends on the price of this commodity. The aim is to valuate an option to expand the project by a factor κ at time T at the cost of $\mathcal{K}$. Under the assumption that both the commodity price and the variance of the price follow a geometric Brownian motion, the model is represented by partial differential equations (PDEs), which are similar to those representing the Hull-White model for financial options. However, the situation for real options is much more complicated than for financial options because there are several equations to be solved, the payoff function is not given but is determined as the solution of PDEs, and setting appropriate terminal and boundary conditions is also more difficult.

Due to these difficulties, the PDE approach for expansion options has not been studied as widely as for financial options, and it is still an active and challenging area of research. The research on PDE methods for expansion options includes papers (Haque et al, 2014; Li and Wang, 2019; Hozman et al, 2024; Černá, 2023) for the model with constant volatility and (Li et al, 2022) for the model with stochastic volatility.

The objective is to design a model suitable for expansion options. Unlike the stochastic volatility model in (Li et al, 2022), the model in this paper is formulated in logarithmic prices, includes a parameter representing the mean convenience yield on holding one unit of output, equations representing the model are not degenerate, and more precise boundary conditions are established.

The second objective is to propose a numerical method for the model that is efficient in the sense that it is high-order convergent with respect to both time and spatial variables and requires a small number of iterations to solve the resulting system of equations. It has been shown that the wavelet-based methods have these required properties for various option pricing problems, see (Hilber et al, 2013; Černá, 2019; Černá and Fiňková, 2024). Therefore, we use the wavelet-Galerkin method combined with the Crank-Nicolson scheme with Richardson extrapolation. The aim also is to implement the proposed scheme and apply it to a benchmark problem in the iron-ore mining industry.

[1] Technical University of Liberec, Department of Mathematics, Studentská 2, 461 17 Liberec, Czech Republic, dana.cerna@tul.cz

2 Stochastic Volatility Model for Expansion Options

The basic assumption is that both the commodity price P and the instantaneous variance y of P follow a geometric Brownian motion,

$$dP = P\ (r-\delta)\ dt + P\sqrt{y}\, dW, \quad dy = y\,\mu_y\, dt + y\,\sigma_y\, dZ. \tag{1}$$

The variable t represents time, r is the risk-free interest rate, δ is the mean convenience yield on holding one unit of output, μ_y is a constant drift rate of y, σ_y is a constant volatility of y, and W and Z are the standard Wiener processes, which are correlated with the correlation coefficient ρ.

The model contains two projects. The first project P_0 does not include any change in production rate. The second project P_1 involves increasing production by a factor $\kappa > 1$ at time T at the cost of $\mathcal{K}$. These two projects are characterized by the following quantities: the lifetime of the project P_k denoted as T_k^*, the production rate $q_k(t)$ of project P_k, the average cash cost rate of production per unit output $C(t)$, the state royalties rate R, and the company income tax rate denoted by B. The after-tax cash flow rate of project P_k is then given by (Haque et al, 2014; Li and Wang, 2019):

$$D_k(P,t) = q_k(t)\,(P(1-R) - C(t))\,(1-B), \quad k = 0, 1. \tag{2}$$

Based on the theory from (Black and Scholes, 1973; Hull and White, 1987) and employing a similar approach to (Li et al, 2022), we establish that the value $V_k(P, y, t)$ of the project P_k is a solution to the equation

$$\frac{\partial V_k}{\partial t} + \frac{P^2 y}{2}\frac{\partial^2 V_k}{\partial P^2} + \rho P \sigma_y y^{3/2}\frac{\partial^2 V_k}{\partial P \partial y} + \frac{\sigma_y^2 y^2}{2}\frac{\partial^2 V_k}{\partial y^2} + (r-\delta)\,P\frac{\partial V_k}{\partial P} + \left(\mu_y - \lambda\sigma_y\right) y\frac{\partial V_k}{\partial y} - rV_k = -D_k(P,t) \tag{3}$$

where $P \in (0,\infty)$, $y \in (0,\infty)$, $t \in \left(T_k, T_k^*\right)$, $T_0 = 0$, $T_1 = T$, and λ represents the market parameter (Li et al, 2022).

The second-order terms are zero for $P = 0$ and $y = 0$, which means that equation (3) is degenerate. However, the degeneracy can be eliminated using logarithmic transformation and localization. Hence, we define new variables $x = \log P$ and $z = \log y$, choose the minimum price $P_{\min}$ and the maximum price $P_{\max}$ and define $x_1 = \log P_{\min}$ and $x_2 = \log P_{\max}$. Similarly, let $z_1 = \log y_{\min}$ and $z_2 = \log y_{\max}$, where $y_{\min}$ and $y_{\max}$ are chosen minimum and maximum bounds for y, respectively. Then, a new domain is $\Omega = (x_1, x_2) \times (z_1, z_2)$ and the function $U_k(x,z,t) = V_k(e^x, e^z, t)$ satisfies the equation

$$\frac{\partial U_k}{\partial t} + \mathcal{L}_1 U_k = -D_k(e^x, t), \quad (x,z) \in \Omega, \quad t \in \left(T_k, T_k^*\right), \tag{4}$$

where the nondegenerate differential operator $\mathcal{L}_1$ is given by

$$\mathcal{L}_1 U_k = \frac{e^z}{2}\frac{\partial^2 U_k}{\partial x^2} + \rho\sigma_y e^{z/2}\frac{\partial^2 U_k}{\partial x \partial z} + \frac{\sigma_y^2}{2}\frac{\partial^2 U_k}{\partial z^2} + \left(r - \delta - \frac{e^z}{2}\right)\frac{\partial U_k}{\partial P} + \left(\mu_y - \lambda\sigma_y - \frac{\sigma_y^2}{2}\right)\frac{\partial U_k}{\partial z} - rU_k. \tag{5}$$

Boundary conditions for $z = z_1$ and $z = z_2$ are set such that the variance is assumed to be constant z_1 or z_2, and the equation for the model with constant volatility is solved using the method described in detail in (Černá, 2023). Let $g_k^i(x,t)$ be the resulting function for the equation in logarithmic prices for project P_k and variance z_i. Then, the boundary conditions are

$$U_k(x, z_i, t) = g_k^i(x,t), \quad U_k(x_i, z, t) = g_k^1(x_i, t) = g_k^2(x_i, t), \quad x \in (x_1, x_2), \quad z \in (z_1, z_2), \quad i, k = 1, 2. \tag{6}$$

The terminal condition at T_k^* reflects that the project value becomes zero when the end of the lifetime is reached,

$$U_k\left(x, z, T_k^*\right) = 0, \quad (x,z) \in \Omega. \tag{7}$$

Now, we consider an option to increase production by a factor $\kappa > 1$ at the cost $\mathcal{K}$ at time T and denote the value of this option by $H(P, y, t)$. Following the same approach as above, we employ logarithmic transformation and find that the function $F(x,z,t) = H(e^x, e^z, t)$ satisfies

$$\frac{\partial F}{\partial t} + \mathcal{L}_1 F = -D_1(e^x, t) + D_0(e^x, t), \quad (x,z) \in \Omega, \quad t \in (0, T). \tag{8}$$

The terminal condition at time T reflects the investment at price $\mathcal{K}$,

$$F(x,z,T) = \max\left(U_1(x,z,T) - U_0(x,z,T) - \mathcal{K}, 0\right), \quad (x,z) \in \Omega. \tag{9}$$

Boundary conditions have the form

$$F(x, z_i, t) = h_i(x, t), \quad F(x_i, z, t) = h_1(x_i, t) = h_2(x_i, t), \qquad x \in (x_1, x_2), \quad z \in (z_1, z_2), \quad t \in [0, T], \tag{10}$$

where the functions h_1 and h_2 represent the value of the expansion option with constant volatility z_1 and z_2, respectively, which are computed using the method from (Černá, 2023).

Hence, we first compute functions g_k^i and h_i, $i, k = 1, 2$, as solutions of nonstationary equations with one spatial variable using the method from (Černá, 2023). These functions g_k^i and h_i are then used to set boundary conditions for equations representing project values U_k and option values F. Then, we solve equations for project values U_0 and U_1, use functions U_0 and U_1 to set the terminal condition for F, and solve the equation for F. Thus, we have a total of three non-stationary equations with two spatial variables and six non-stationary equations with one spatial variable to solve.

3 Orthogonal Cubic Spline Wavelet Basis

Since the aim is to develop a wavelet-based method, we need an appropriate wavelet basis. Various wavelet bases were already tested for the valuation of expansion options under constant volatility in (Černá, 2023), and it has been shown that an orthogonal cubic spline wavelet basis from (Černá and Fiňková, 2024) is superior to other cubic spline wavelet bases. For this reason, we also use this wavelet basis here. However, the proposed method can also be used with other types of wavelets such as orthogonal wavelets (Černá et al, 2008), biorthogonal spline wavelets (Černá, 2019), and spline multiwavelets (Shumilov, 2013).

Let Ψ^I be an orthogonal cubic spline wavelet basis adapted to homogeneous Dirichlet boundary conditions from (Černá and Fiňková, 2024). This wavelet basis is an orthogonal wavelet basis of the space $L^2(I)$, $I = (0, 1)$, and has a hierarchical structure,

$$\Psi^I = \Phi_{j_0}^I \cup \bigcup_{j=j_0}^{\infty} \Psi_j^I, \quad \Phi_{j_0}^I = \left\{\phi_{j_0,k}, k \in \mathcal{I}_{j_0}\right\}, \quad \Psi_j^I = \left\{\psi_{j,k}, k \in \mathcal{J}_j\right\}. \tag{11}$$

The functions $\phi_{j_0,k}$ are called scaling functions, and the functions $\psi_{j,k}$ are called wavelets. For more details about the concept of a wavelet basis, refer to (Hilber et al, 2013; Černá, 2019).

To solve equations with two spatial variables, we need a two-dimensional wavelet basis defined on $\square = (0, 1)^2$. Such a basis is constructed using the tensor product of functions from Ψ^I. Recall that the tensor product of two functions u and v is defined as $u \otimes v(x, y) = u(x)\, v(y)$ and for two sets Γ_1 and Γ_2 of functions from $L^2(I)$ denote

$$\Gamma_1 \otimes \Gamma_2 = \{\gamma_1 \otimes \gamma_2 : \gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2\}. \tag{12}$$

Using so-called isotropic approach, the multiscale wavelet bases with s levels of wavelets are defined as

$$\Psi^s = \left(\Phi_{j_0}^I \otimes \Phi_{j_0}^I\right) \cup \bigcup_{j=j_0}^{j_0+s-1} \left(\Phi_j^I \otimes \Psi_j^I \cup \Psi_j^I \otimes \Phi_j^I \cup \Psi_j^I \otimes \Psi_j^I\right). \tag{13}$$

4 Wavelet Method

The proposed model is characterized by three PDEs with two spatial variables that all have the same form:

$$\frac{\partial U}{\partial t} + \mathcal{L}_1 U = f(x, t), \quad (x, z) \in \Omega, \quad t \in (\tau_1, \tau_2), \tag{14}$$

and are equipped with boundary conditions

$$\begin{aligned} U(x_1, z, t) &= G_1(t), \; U(x_2, z, t) = G_2(t), \; z \in (z_1, z_2), \; t \in (\tau_1, \tau_2) \\ U(x, z_1, t) &= G_3(x, t), \; U(x, z_2, t) = G_4(x, t), \; x \in (x_1, x_2), \; t \in (\tau_1, \tau_2), \end{aligned}$$

and terminal condition

$$U(x, z, \tau_2) = G_5(x, z), \quad (x, z) \in \Omega. \tag{15}$$

Since the wavelets are defined on the unit square $\square$, we first transform the equation to be also defined on $\square$ using new variables $(x - x_1)/x_d$ and $(z - z_1)/z_d$, where $x_d = x_2 - x_1$ and $z_d = z_2 - z_1$. The next step is the transformation to homogeneous Dirichlet boundary conditions. To this end, we define a function w that satisfies the above boundary conditions as follows:

$$w(x,z,t) = G_3(x,t) + (G_4(x,t) - G_3(x,t))\,z, \quad (x,z) \in \square, \quad t \in [\tau_1, \tau_2]. \tag{16}$$

Then a function $u(x,z,t) = U(x\,x_d + x_1, z\,z_d + z_1, t) - w(x,z,t)$ is the solution of the equation

$$\frac{\partial u}{\partial t} + \mathcal{L}u = p(x,z,t), \quad (x,z) \in \square, \quad t \in (\tau_1, \tau_2), \tag{17}$$

where the differential operator $\mathcal{L}$ is given by

$$\mathcal{L}u = \frac{e^{z\,z_d+z_1}}{2x_d^2}\frac{\partial^2 u}{\partial x^2} + \frac{\rho\sigma_y e^{\frac{z\,z_d+z_1}{2}}}{x_d z_d}\frac{\partial^2 u}{\partial x \partial z} + \frac{\sigma_y^2}{2z_d^2}\frac{\partial^2 u}{\partial z^2} + \left(r - \delta - \frac{e^{z z_d+z_1}}{2}\right)\frac{1}{x_d}\frac{\partial u}{\partial x} + \left(\mu_y - \lambda\sigma_y - \frac{\sigma_y^2}{2}\right)\frac{1}{z_d}\frac{\partial u}{\partial z} - rv, \tag{18}$$

and the right-hand side is

$$p(x,z,t) = f(x\,x_d + x_1, t) - \frac{\partial w}{\partial t}(x,z,t) - \mathcal{L}w(x,z,t). \tag{19}$$

The transformed terminal condition has the form

$$u(x,z,\tau_2) = G_5(x\,x_d + x_1, z\,z_d + z_1) - w(x,z,\tau_2), \ (x,z) \in \square, \tag{20}$$

and the equation is now equipped with homogeneous Dirichlet boundary conditions.

Let $(\cdot,\cdot)$ denote the L^2-inner product and define the bilinear form

$$a(u,z) = (\mathcal{L}(u), z), \quad u, v \in H_0^1(\square). \tag{21}$$

We use the wavelet-Galerkin method for spatial discretization. Let Ψ^s be defined as above, $V^s = \operatorname{span}\Psi^s$ and $\tilde{V}^s$ be the dual of V^s. The objective is to find a function $u \in L^2(\tau_1, \tau_2; V^s) \cap H^1(\tau_1, \tau_2; \tilde{V}^s)$ such that

$$\left(\frac{\partial u}{\partial t}, v\right) + a(u,v) = (p,v) \quad \forall v \in V^s. \tag{22}$$

Furthermore, we utilize the Crank-Nicolson scheme. Let $M \in \mathbb{N}$, and

$$\tau = \frac{\tau_2 - \tau_1}{M}, \ t_m = \tau_1 + m\tau, \ u_m(x,z) = u(x,z,t_m), \ f_m(x,z) = p(x,z,t_m), m = 0, \ldots, M. \tag{23}$$

The scheme is a backward scheme, where at time t_m, $m = M-1, \ldots, 0$, we determine $u_m \in V^s$ using

$$\frac{(u_m, v)}{\tau} + \frac{a(u_m, v)}{2} = p_m(v) \quad \forall v \in V^s, \quad p_m(v) = \frac{(u_{m+1}, v)}{\tau} - \frac{a(u_{m+1}, v)}{2} - \frac{(f_m + f_{m+1}, v)}{2}. \tag{24}$$

Next, we expand the solutions u_m in the basis Ψ^s and denote the vector of coefficients of this expansion as $\mathbf{u}_m^s$,

$$u_m = \sum_{\psi_\lambda \in \Psi^s} (\mathbf{u}_m^s)_\lambda \psi_\lambda. \tag{25}$$

Substituting (25) and $v = \psi_\mu$ into (24), we obtain

$$\mathbf{A}^s \mathbf{u}_m^s = \mathbf{f}_m^s, \quad \mathbf{A}_{\mu,\lambda}^s = \frac{(\psi_\lambda, \psi_\mu)}{\tau} + \frac{a(\psi_\lambda, \psi_\mu)}{2}, \quad (\mathbf{f}_m^s)_\mu = p_m(\psi_\mu), \quad \psi_\lambda, \psi_\mu \in \Psi^s. \tag{26}$$

We choose the generalized minimal residual method (GMRES) with diagonal preconditioning for the solution of the system (26). By employing the wavelet-based method, it can be demonstrated similarly to (Černá and Fiňková, 2024) that the condition numbers of the resulting diagonally preconditioned matrices are uniformly bounded.

Finally, we improve the convergence of the Crank-Nicolson scheme by postprocessing based on Richardson extrapolation. Let $V_{N,M}$ be the computed approximate solution using N basis functions and M time steps. We define a new approximate solution as

$$V_{N,2M}^R = \frac{4V_{N,2M} - V_{N,M}}{3}. \tag{27}$$

Richardson extrapolation has the potential for some types of option pricing problems to improve the convergence rate with respect to τ from $O(\tau^2)$ to $O(\tau^4)$, see (Arciniega and Allen, 2004).

5 Expansion Options for Iron Ore Mining Project

We present numerical results for the benchmark problem regarding an option to expand an investment project in the iron ore mining industry from (Li and Wang, 2019; Li et al, 2022; Černá, 2023). The project data are as follows:

$$B = 0.3, \quad R = 0.05, \quad r = 0.06, \quad \delta = 0.02, \quad \sigma_y = 0.1, \quad \mu_y = 0.1, \quad \rho = 0.1, \quad \lambda = 0.1. \qquad (28)$$

Furthermore, $C(t) = 35e^{0.005t}$ and $q_0(t) = 0.1\, e^{0.007t}$. The production parameters are in billion tons. The cost is $\mathcal{K} = 10$ billion US dollars for the expansion at time $T = 2$ years. In the case of expansion, the production rate is multiplied by $\kappa = 2$ and thus the production rate for the project P_1 is $q_1(t) = q_0(t)$ for $t \in [0, T)$ and $q_1(t) = \kappa\, q_0(t)$ for $t \geq T$. The lifetimes of the projects are $T_0^* = 75.80$ and $T_1^* = 43.62$ years. After the lifetime T_k^*, resources are already exhausted, and thus the production rate $q_k(t) = 0$ for $t > T_k^*$ for $k = 0, 1$. We choose $P_{\min} = 0.01$, $P_{\max} = 800$, $y_{\min} = 0.0001$, $y_{\max} = 1$, and compute the approximate solution using the proposed scheme. The resulting function representing the price of the option at time $t = 0$ is displayed in Figure 1.

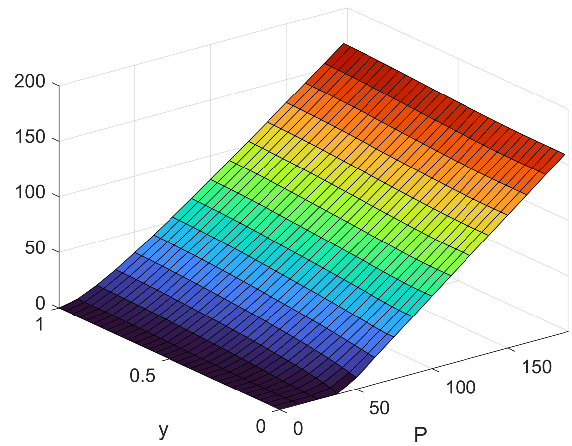


Figure 1 The function $H(P, y, 0)$ representing the value (in billions of US dollars) of the option to expand an iron ore investment project at time $t = 0$ in dependence on iron ore price P and variance y of P.

The quantities characterizing convergence are listed in Table 1. The parameter s denotes the number of wavelet levels, N is the number of basis functions, and $\tilde{M}$ is the number of time steps per unit interval. For the given level s and the corresponding parameters N and $\tilde{M}$, we compute option values $H_s(P) = H(P, 0.25, 0)$ (in billions of US dollars) for $P = 110$ and $P = 130$. To study the convergence behavior, we compute differences $d_s(P) = H_s(P) - H_{s-1}(P)$. These values characterize convergence in the sense that the error of order p implies that differences d_s are of order p. Thus, the high-order convergence of d_s suggests that the method is high-order convergent. The parameter it is the number of GMRES iterations in the last time step if the stopping criterion is that the relative residual is less than 10^{-12} and restart is after every 10 iteration.

s	N	$\tilde{M}$	$H_s(110)$	$d_s(110)$	$H_s(130)$	$d_s(130)$	it
1	144	4	76.6078		97.5039		11
2	576	8	78.0536	1.45e0	99.0660	1.56e0	15
3	2304	16	78.0037	4.99e-2	99.0926	2.67e-2	19
4	9216	32	78.0004	3.33e-3	99.0910	1.66e-3	22
5	36864	64	77.9999	5.23e-4	99.0901	8.70e-4	24

Table 1 Resulting option values $H_s(P)$, their differences $d_s(P)$, and numbers of GMRES iterations it.

6 Conclusion

We proposed the stochastic volatility model for one-stage expansion options. Unlike the stochastic volatility model in (Li et al, 2022), the model in this paper is formulated in logarithmic prices, includes a parameter representing the mean convenience yield on holding one unit of output, equations representing the model are not degenerate, and more precise boundary conditions are established. Furthermore, we proposed a wavelet-based method that uses the Galerkin method with an orthogonal cubic spline wavelet basis from (Černá and Fiňková, 2024) in combination with the Crank-Nicolson scheme with Richardson extrapolation. Numerical experiments for an option to expand an investment project in the iron ore mining industry confirm that the method is relevant and demonstrate the advantages, which are high-order convergence and the fact that the method requires a small number of iterations needed to resolve the problem with the required accuracy. The future aim is to develop an efficient wavelet-based method for more complex models of real options pricing, such as models with stochastic volatility for multi-stage expansion options, models with multiple commodities, and models for other types of real options.

Acknowledgements

This work was supported by grant No. GA22-17028S funded by the Czech Science Foundation.

References

Arciniega A, Allen E (2004) Extrapolation of difference methods in option valuation. Applied Mathematics and Computation 153:637–659. https://doi.org/10.1016/S0096-3003(03)00621-0

Black F, Scholes M (1973) The pricing of options and corporate liabilities. Journal of Political Economy 81:637–659. https://doi.org/10.1086/260062

Černá D (2019) Cubic spline wavelets with four vanishing moments on the interval and their applications to option pricing under Kou model. International Journal of Wavelets, Multiresolution and Information Processing 17:1850061. https://doi.org/10.1142/S0219691318500613

Černá D (2023) Wavelet method for pricing options on two-stage expansion of investment project. In: 2023 International Conference on Applied Mathematics & Computer Science (ICAMCS), pp 28–35, https://doi.org/10.1109/ICAMCS59110.2023.00012

Černá D, Fiňková K (2024) Option pricing under multifactor Black-Scholes model using orthogonal spline wavelets. Mathematics and Computers in Simulation 220:309–340. https://doi.org/10.1016/j.matcom.2024.01.020

Černá D, Finěk V, Najzar K (2008) On the exact values of coefficients of coiflets. Central European Journal of Mathematics 6:159–169. https://doi.org/10.2478/s11533-008-0011-2

Dixit A, Pindyck R (1994) Investment Under Uncertainty. Princeton University Press, Princeton

Haque MA, Topal E, Lilford E (2014) A numerical study for a mining project using real options valuation under commodity price uncertainty. Resources Policy 39:115–123. https://doi.org/10.1016/j.resourpol.2013.12.004

Hilber N, Reichmann O, Schwab C, et al (2013) Computational Methods for Quantitative Finance. Springer, Berlin

Hozman J, Tichý T, Dvořáčková H (2024) Valuation of mining projects under dynamic model framework. Annals of Operations Research 337:1167–1204. https://doi.org/10.1007/s10479-023-05569-y

Hull J, White A (1987) The pricing of options on assets with stochastic volatilities. The Journal of Finance 42:281–300. https://doi.org/10.1111/j.1540-6261.1987.tb02568.x

Li N, Wang S (2019) Pricing options on investment project expansions under commodity price uncertainty. Journal of Industrial and Management Optimization 15:261–273. https://doi.org/10.3934/jimo.2018042

Li N, Wang S, Zhang K (2022) Price options on investment project expansion under commodity price and volatility uncertainties using a novel finite difference method. Applied Mathematics and Computation 421:126937. https://doi.org/10.1016/j.amc.2022.126937

Myers SC (1977) Determinants of corporate borrowing. Journal of Financial Economics 5(2):147–175. https://doi.org/10.1016/0304-405X(77)90015-0

Shumilov B (2013) Multiwavelets of the third-degree Hermitian splines orthogonal to cubic polynomials. Mathematical Models and Computer Simulations 5:511–519. https://doi.org/10.1134/S2070048213060100

Application of the Three-Level Aggregation Model for Evaluating Opinions Under Hesitance for Fuzzy Voting in Spatial Planning Public Decision-Making

Benjamin Emmenegger[1], Miroslav Hudec[2], Frantisek Zapletal[3], Georgiana Bigea[4], Edy Portmann[5]

Abstract. Public planning decisions affect the living conditions of inhabitants and their subgroups differently. Citizens should express their support and/or resistance to each alternative. Since their opinions tend to be subjective, fuzzy voting can be used to express them and reveal to what extent an opinion is in favor or against each alternative. However, inconsistent responses consisting of simultaneous high levels of support and resistance for the same alternative represent a challenge to reach the final decision. To address it, this paper proposes strengthening the consistent answers and weakening the contradictory responses. In the real-life Case Study of Unterdorfstrasse (Switzerland) this consistency is handled and the impact of coalitions among subgroups is explored, as the impact is different if two of the most affected subgroups or two lightly affected subgroups agree on a specific alternative. Using the selected fuzzy measures, we assign weights to subgroups and their coalitions based on geographical features. In addition, to check the robustness of the results, a sensitivity analysis is performed using Monte Carlo simulation. In this way, we emphasize the importance of understanding the dynamics within and between these subgroups to interpret the results. The model accentuates differences in the data and offers a clearer view of the tendencies.

Keywords: fuzzy voting, Choquet integral, spatial planning, decision making, Łukasiewicz t-norm and t-conorm

JEL Classification: C10, C44
AMS Classification: C86

1 Introduction

The use of fuzzy sets and fuzzy logic methods for group decision making has been extensively researched, both theoretically and applied, with the specific characteristics of each application. For example, a fuzzy voting referendum research project was applied in Basel, Switzerland, concluding that it was a welcomed alternative, particularly by undecided voters or who rarely participated in voting [13]. They appreciated the flexibility to express their inclination towards accepting or rejecting. In public matters, important strategic decisions must be made considering diverse stakeholder groups [15]. Especially in local spatial planning settings, subgroups that are affected at varying levels may be considered, depending on their geographic location, among other factors [8]. Traditionally, these issues were addressed through surveys and similar methodologies. However, these approaches frequently encounter limitations as respondents may struggle to provide precise answers [1, 5, 12]. Due to this uncertainty regarding the level of support or resistance in relation to specific issues, the need for more nuanced analytical methods is highlighted and this is where the present study fills a gap in the literature [13].

Since weights are often assigned by expert opinions and/or from data and smaller changes in values might influence solutions (e.g., the order or alternatives), a thorough sensitivity analysis using the Monte Carlo simulation was performed. Monte Carlo replaces the explored parameter(s) with a random variable(s) with assigned probability distribution. The model is then run repeatedly with randomly generated values of random parameters, and the results are analyzed [10].

This paper is structured as follows: the second section contains the Case Study of Unterdorfstrasse, followed by the third section about the methods, detailing the different levels of aggregation, the fourth section describes the results, ending with the analysis and the conclusion.

[1] Human-IST Institute, University of Fribourg, Bd de Pérolles 90, CH-1700 Fribourg, Switzerland, benjamin.emmenegger@unifr.ch
[2] Faculty of Economics, VSB - Technical University of Ostrava, Czech Republic, miroslav.hudec@vsb.cz
[3] Faculty of Economics, VSB - Technical University of Ostrava, Czech Republic, frantisek.zapletal@vsb.cz
[4] Babes Bolyai University, Cluj-Napoca, Romania, georgiana.bigea@ubbcluj.ro
[5] Human-IST Institute, University of Fribourg, Bd de Pérolles 90, CH-1700 Fribourg, Switzerland, edy.portmann@unifr.ch

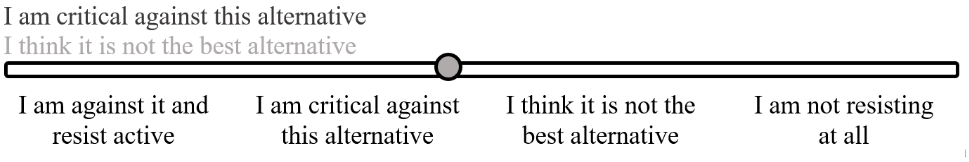


Figure 1: Example of a slider showing CLEs [7]

2 Case Study of Unterdorfstrasse

The Case Study of Unterdorfstrasse was a participatory process in Switzerland, where both traditional and innovative methods were used to engage the community in urban spatial planning. First, the municipality wanted to identify all possible solutions by collecting alternatives from stakeholders. The entire process was preceded by a preliminary decision as to which solution (considered subsequently as alternative 1) should be implemented, which caused great unrest among the population. In total, 13 alternatives were collected from the stakeholders. Eight of them involved only organizational adjustments, while five of them effectively involved new physical routes for motor vehicles, some of which required the construction of new roads. After the collection phase, a fuzzy voting phase followed, where all 13 alternatives were evaluated, together with 540 arguments collected in favor and against them, through online and analogue means. At the voting session, each of the participants expressed their degree of support and/or resistance about each one of the alternatives, using fuzzy sliders. The main reason for assessing these two indicators was to estimate which of the alternative(s) benefited from minimal opposition and maximum support, to prevent future community conflicts and assess the level of potential opposition that would emerge, with the aim of a consensus-based decision-making approach [6].

However, since it is common in public decisions that some decision-makers are more affected by future consequences than others, e.g., only some citizens will be directly affected by the construction of an alternative route on their property and others will not, the question that emerged was whether the agreement of certain groups, who are most affected by the alternative, should be weighed differently than that of others, who are less affected by it, and how would that change the results.

The data was collected by using fuzzy sliders with Comparative Linguistic Expressions (CLEs), used to assess support and resistance [7]. Each slider displayed four statuses, e.g., ranging from 'I do not agree' to 'I agree' as showed in Fig. 1. CLEs were used to allow voters to assess the options more intuitively, in a manner that is more similar to human reasoning.

3 Methods

In this work, we formalized the preference of each respondent (i.e., intensities of support and rejection), followed by the aggregation on a subgroup level and among subgroups. Finally, by using the Monte Carlo simulation we examined the sensitivity of the result.

3.1 Aggregation on the respondent level

When a respondent provides the intensity of support and resistance, his/her reliability should be considered, i.e., the relation between intensities of support and resistance. Let us look at the possible behaviours and their formalization by the logic aggregation perspective:

1. when support (S) and resistance (R) are both high, the consistency of answer is low
2. when support (S) is high and resistance (R) is low, the answer is highly for the alternative
3. when support (S) is low and resistance (R) is high, the answer is highly against the alternative
4. when support (S) is low and resistance (R) is low, the consistency of answer is low

To apply aggregation functions, we transformed resistance into its negation (i.e., $R_n = 1 - R$).

This leads to the adoption of mixed aggregation functions [9]. An option is the convex combination of Łukasiewicz t-norm and its dual t-conorm (for simplicity, we denote support by x and $1 - R$ by y),

$$A_\lambda(x, y) = \lambda \cdot \max(0, x + y - 1) + (1 - \lambda) \cdot \min(1, x + y) \quad (1)$$

where $T_L = \max(0, x + y - 1)$, $S_L = \min(1, x + y)$ and $\lambda \in [0, 1]$. When $\lambda = 1$, we get Łukasiewicz t-norm T_L; when $\lambda = 0$, we get Łukasiewicz t-conorm S_L; while when $\lambda = 0.5$, we get the averaging behavior on the whole domain. Applying one value of λ is not able to handle all the requirements in this work.

The values of (λ) are determined based on the conditions of support (x) and resistance (y) as follows:

- When $(x, y) \in [0, 0.5)^2$, then $\lambda = 0.75$ emphasizing inclination to resistance by conjunctive function.
- When $(x, y) \in (0.5, 1]^2$, then $\lambda = 0.25$ emphasizes the inclination to support by a disjunctive function.
- In the other two cases, $(x, y) \in [0, 0.5] \times [0.5, 1] \cup [0.5, 1] \times [0, 0.5]$, then $\lambda = 0.5$ emphasizes indecisiveness by means of the averaging function.

This methodology suggests that values near 0.5 aggregate very indecisive or conflicting opinions, while values approaching 0 or 1 indicate a high inclination toward rejection or acceptance, respectively.

3.2 Aggregation on the subgroup level

In each subgroup, we have a certain number of respondents, where each answer is in the unit interval. A suitable aggregation for the opinion of a group is by arithmetic mean, due to the neutral logical behavior [4], (i.e., neither inclination to 0 - full resistance, like geometric mean, or inclination to 1 - full support, like quadratic mean):

$$A_z = \frac{1}{m} \sum_{j=1}^{m} A_\lambda(x_j, y_j) \tag{2}$$

where m is the number of respondents in a subgroup and $A_\lambda(x_j, y_j)$ is the aggregated support and rejection of the respondent j by (1).

3.3 Aggregation of subgroups

The importance of coalitions is managed by the discrete Choquet integral [3] and fuzzy measures. Choquet integration is based on not necessarily additive monotone measures $v : 2^{\mathcal{N}} \rightarrow [0, 1]$. A discrete fuzzy measure [11, 14] is a set function on $\mathcal{N} = \{1, 2, \ldots, n\}$ which is monotonic ($v(\mathcal{A}) \leq v(\mathcal{B})$ whenever $\mathcal{A} \subseteq \mathcal{B}$) and satisfies boundary conditions $v(\emptyset) = 0$ and $v(\mathcal{N}) = 1$.

A subset $\mathcal{A} \subseteq \mathcal{N}$ is considered as a coalition, where $v(\mathcal{A})$ explains the importance of coalition. The discrete Choquet integral with respect to a fuzzy measure v is given by

$$C_v(\boldsymbol{x}) = \sum_{i=1}^{n} x_{(i)} \left[v\left(\{j | x_j \geq x_{(i)}\}\right) - v\left(\{j | x_j \geq x_{(i+1)}\}\right) \right] \tag{3}$$

where $\left(x_{(1)}, x_{(2)}, \ldots x_{(n)}\right)$ is a permutation of non-decreasing values, n is the number of subgroups, and $x_{(n+1)} = \infty$. An alternative expression, more suitable for computing is [2]

$$C_v(\boldsymbol{x}) = \sum_{i=1}^{n} \left[x_{(i)} - x_{(i-1)} \right] v(H_i) \tag{4}$$

where $x_{(0)} = 0$ and $H_i = \{(i), \ldots, (n)\}$ is the subset of indices of the $(n - i + 1)$ largest components of vector $\boldsymbol{x}$.

When the collation of two subgroups A and B should be emphasised, we model as $v(A + B) > v(A) + v(B)$ (super additivity). When it should be attenuated, we model as $v(A + B) < v(A) + v(B)$ (sub additivity) and finally when coalition does not bring nothing new, we model it as $v(A + B) = v(A) + v(B)$ (additivity). In all three cases, we should keep the monotonic property.

4 Results

The data collected from the Case Study of Unterdorfstrasse was processed using the methods, and the following results have been obtained.

4.1 Collected Data

All data were gathered utilizing sliders. In this approach, support and resistance levels for various alternatives were captured, and the feedback was analyzed. This method ensured a nuanced understanding of stakeholder positions. The support and resistance were assessed using CLEs. Two slider dimensions were offered for each alternative, each with four CLEs and direct visual feedback, as shown in Fig. 1 [7].

The data collected in the Case Study of Unterdorfstrasse consisted of three relevant tables that outline a structure designed to capture the voting per alternative. In Fig. 2 three tables can be seen. The table `singleVote` holds

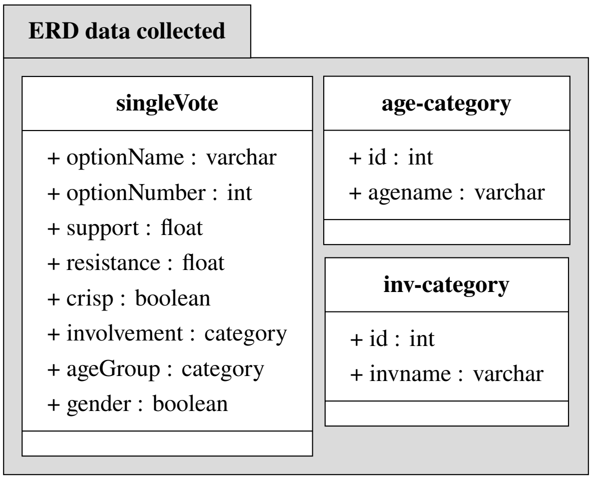


Figure 2: Captured entities

a variety of attributes for each vote, including unique identifiers and the fuzzy-voting measures of `support` (S) or `resistance` (1-R). It also links to demographic data such as the voter's involvement level and age group, which are categorized in the tables `inv-category` and `age-category` tables respectively. These auxiliary tables are referenced to analyze voting behavior more granularly. They provide predefined categories for the voters' involvement, distinguishing between non-affected and affected, depending on the degree of affectedness per each alternative.

4.2 Weights of Coalitions

Based on the involvement of people, four groups were ultimately identified which, for this case, show a reasonable size and differentiation from each other. Purely geographically, seven groups of participants were identified together with the local planner. Since the alternatives are local solutions with physical changes in the streets and property, the groups were assigned to neighborhoods. The results originating from three of the neighborhoods were not taken into account, because fewer than four people participated from them. From the other four neighborhoods, 14 to 27 [1] people took part. The four groups and all combinations thereof ultimately result in 16 different possible sets of groups. The initial weights were qualitatively determined for an initial starting point. The weights were created with respect to the impact of the combinations and individual groups on the implementation of the road infrastructure (alternatives). The initial weights are shown and described below:

- Grp1 (0.179): Somewhat important; influence depends on which solution is chosen and whether they are affected or not.
- Grp2 (0.071): Not important; just want a solution to the problem.
- Grp3 (0.286): Important to the decision in any case.
- Grp4 (0.464): Most impact on the decision since they are landlords.
- $v(\text{Grp1}, \text{Grp2})$ (0.320): Not good since not much affected.
- $v(\text{Grp1}, \text{Grp3})$ (0.970): Only neighbors of some cases but not directly affected by the problem today.
- $v(\text{Grp1}, \text{Grp4})$ (0.226): More impact since it includes more relevant and affected people.
- $v(\text{Grp2}, \text{Grp3})$ (0.161): Some impact and affected in some cases.
- $v(\text{Grp2}, \text{Grp4})$ (0.226): Impact since it includes many people who are affected.
- $v(\text{Grp3}, \text{Grp4})$ (0.258): Most impact on the decision since affected in all solutions, positively or negatively.
- $v(\text{Grp1}, \text{Grp2}, \text{Grp3})$ (0.217): Impact since it includes many people.
- $v(\text{Grp1}, \text{Grp2}, \text{Grp4})$ (0.246): Some impact but not higher than the groups with 3 and 4.
- $v(\text{Grp1}, \text{Grp3}, \text{Grp4})$ (0.275): All affected groups are included.
- $v(\text{Grp2}, \text{Grp3}, \text{Grp4})$ (0.261): Includes very many people.
- $v(\text{Grp1}, \text{Grp2}, \text{Grp3}, \text{Grp4})$ (1): All groups are affected - value required by theory.

Based on these weights per coalition, the Choquet integration (4) was calculated at the third level.

[1] N = 16, 27, 14 and 25

4.3 Sensitivity analysis

The weights of coalitions are expected to be the driving factor of the results. At the same time, these weights can hardly be precisely set based on some measuring. Therefore, uncertainty can play a potentially important role for this factor. This uncertainty is explored using Monte Carlo simulation [10]. The weights of single groups are considered random and uniformly distributed. In total, 2,000 runs were used.

The setting for the analysis is given as follows. We change the weight of the single groups (v_1 to v_4), but the synergies of the coalitions are preserved (the ratios between the coalitions and the single groups are kept; for example, if the original weights are $v_1 = 0.5$, $v_2 = 0.8$ and $v_{12} = 0.65$, then under new weights $v_1 = 0.4$, $v_2 = 0.6$ we get $v_{12} = 0.5$). In this setting, we assume the change of the single weights by not more than ±10 percentage points (±0.1) while keeping the range $[0, 1]$. Thus, 0.2 can be changed to any value between 0.1 and 0.3; 0.07 can be changed to any value between 0 and 0.17.

The rankings, resulting from Monte Carlo simulation, can be found in Tab. 1. As expected, the rankings of some alternatives are absolutely stable and reached the same position for all scenarios (alternatives 3, 6, 8, and 10, denoted in bold). However, the rest are potentially affected by the changes in weight considered. The greatest uncertainty was found for alternative 9, which can be ranked at five different positions (from 7th to 11th). Alternative 9 is the 0 option to do nothing and leave the problem as it is today. From Tab. 1 you can also interpret what the distribution of alternatives looks like considering the ranking position. The most uncertain ranking is expected at the 8th to 11th positions where from 3 to 4 different alternatives can be found (e.g. ranking 8th contains Alt2, Alt5, Alt9 and Alt11). Since these are rankings, the lower ranks are most likely to be less stable than the top ones. Nevertheless, it is interesting that the ranks 12 and 13 are very clear.

When comparing the obtained results with those calculated using the initial weights, they are consistent. The mean ranking coming from Monte Carlo simulation (see Tab. 1) is identical with the original ranking. Thus, the results are considered stable enough. On the other hand, it was shown that the uncertainty in weights can possibly impact some positions in the ranking, which should be taken into consideration when using the results for future decisions.

Table 1: Resulting ranking from Monte Carlo simulation

Ranking	Alt1	Alt2	Alt3	Alt4	Alt5	Alt6	Alt7	Alt8	Alt9	Alt10	Alt11	Alt12	Alt13
Min	**10**	**8**	1	**11**	**8**	13	**3**	2	**7**	6	**7**	**4**	**3**
Max	**12**	**10**	1	**12**	**10**	13	**4**	2	**11**	6	**8**	**5**	**5**
Mean	**11.04**	**9.61**	1.00	**11.95**	**9.06**	13.00	**3.04**	2.00	**8.24**	6.00	**7.11**	**4.23**	**4.73**

5 Discussion

The top ranks are very clear regardless of the weights of the groups. For the third rank, there is a slight deviation, and the fourth and fifth ranks deviate somewhat up or down. Rank 6 is again very clear, followed by very different evaluations. The last rank (13th) in the table is also clear.

The clearly favored alternative is the one that directs traffic control away from the affected perimeter (those involved). The second ranked solution is the one that attempts to solve the problem (heavy traffic) organizationally at the level of those responsible. Both are understandably given high priority. The third, most controversial, rank (alternative 7) is, interestingly, an organizational solution that does not require any major structural interventions. Also noteworthy is the fact that the solution originally proposed by the Commission (alternative 1) ranks between 10th and 12th.

It must be said that this transdisciplinary approach is more population-oriented than technical. Therefore, it may happen that the first ranking alternatives have a lower probability of implementation in further engineering-based elaborations or regional departments. Then, the subsequent ranks become relevant for the municipality. This will continue until the 0 option of doing nothing is reached.

In regards to decision making and stability despite groups and weights, it is shown that even with a deviation of 10 percentage points, there are very strong preferences regardless of the initial situation. This is of course provided that 1st and 2nd level aggregations have already taken place, which also reinforce the result. With no deviation across 2,000 data sets, this represents a clear weighting of the decision-makers' will. These alternatives must therefore definitely be pursued by the municipality and taken towards the next step.

Furthermore, various weighting scenarios should now be considered, and the previous phase (1st and 2nd levels) should be subjected to a Monte Carlo simulation again to look for differences and patterns. In this way, the method can be further assessed, and the data set can be evaluated. A larger overall data set with a similar method would then provide a significant result.

6 Conclusion

The evaluation of voting alternatives is a demanding task, especially when citizens express their support and resistance to developments that influence their quality of life. A usual voting might be considered as oversimplified. Citizens in a municipality are not equally influenced by any construction plan. Voting supported by weights does not solve the problem. It is a hard task to assign weights, since there might be a different number of citizens in differently influenced areas. In addition, citizens are hesitant in supporting or resisting proposed construction alternatives.

In this work, we proposed an evaluation that considered the individual uncertainty of answers and relevance of subgroup coalitions, assuming that it is not the same impact when two the most affected subgroups or two lightly affected subgroups agree. Although weights are assigned according to the geographical distance and influence, the robustness of the results is realized by simulation. The analysis has shown that the solution is stable, and differences in order of alternatives are minor.

Acknowledgements

This article has been produced with the financial support of the European Union under the REFRESH – Research Excellence For Region Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition, the SGS project no. SP2024/083 of the Ministry of Education, Youth and Sports of the Czech Republic.

Further thanks go to the sponsors of the Case Study of Unterdorfstrasse in particular to the municipality of Geuensee, Kost+Partner A.R. and various contributors from the Lucerne University of Applied Sciences and Arts.

References

[1] Altig, D., Barrero, J. M., Bloom, N., Davis, S. J., Meyer, B., and Parker, N. (2022). Surveying business uncertainty. *Journal of Econometrics*, 231(1):282–303.

[2] Beliakov, G., James, S., and Wu, J.-Z. (2020). *Discrete Fuzzy Measures*. Springer, Cham.

[3] Choquet, G. (1954). *Theory of capacities*, volume 5.

[4] Dujmović, J. (2018). *Soft Computing Evaluation Logic: the LSP Decision Method and Its Applications*. Wiley–IEEE Computer Society.

[5] Ellerby, Z., Wagner, C., and Broomell, S. B. (2022). Capturing richer information: On establishing the validity of an interval-valued survey response mode. *Behavior Research Methods*, 54(3):1240–1262.

[6] Emmenegger, B. (2022). Final report about unterdorfstrasse cas at municipality of geuensee. Technical report, ajato GmbH and Municipality of Geuensee.

[7] Emmenegger, B., Bigea, G., and Portmann, E. (2023). Applied large-scale group decision making using systemic consensus and fuzzy method of comparative linguistic expressions. In *13th Conference of the European Society for Fuzzy Logic and Technology, EUSFLAT 2023*, volume 14069 of *Lecture Notes in Computer Science*, pages 124–136. Springer.

[8] Emmenegger, B., Bigea, G., and Portmann, E. (2024). Opportunities of fuzzy voting in decision-making and negotiation techniques in municipal and regional contexts. *Institute of Electrical and Electronics Engineers Inc.*

[9] Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2009). *Aggregation Functions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press.

[10] Harrison, R. L. (2010). Introduction to monte carlo simulation. In *AIP conference proceedings*, volume 1204, page 17. NIH Public Access.

[11] Horanská, u. and Šipošová, A. (2018). A generalization of the discrete choquet and sugeno integrals based on a fusion function. *Information Sciences*, 451-452:83–99.

[12] Kouatli, I. (2022). The use of fuzzy logic as augmentation to quantitative analysis to unleash knowledge of participants' uncertainty when filling a survey: Case of cloud computing. *IEEE Transactions on Knowledge and Data Engineering*, 34(3):1489–1500.

[13] Schwarz, D., Emmenegger, B., Fivaz, J., Hudec, M., and Portmann, E. (2024). Effects of fuzzy voting on citizen participation: Empirical evidence from switzerland. *Institute of Electrical and Electronics Engineers Inc.*

[14] Wang, Z. and Klir, G. J. (1992). *Fuzzy Measure Theory*. Springer Science + Business Media, New York.

[15] Zapletal, F., Hudec, M., Švaňa, M., and Němec, R. (2023). Three-level model for opinion aggregation under hesitance. *Soft Comput.*, 27(10):6653–6669.

Assessing tourism efficiency in traditional beach touristic centers in Mexico: Application of dynamic two-stage DEA model and fuzzy time series forecasting

Martin Flegl[1], Carmen Lozano[2], Patrick Joaquín Cruz[3], Marketa Matulova[4]

Abstract. Tourism plays an important role in the Mexican economy contributing around 7.6% to the Gross Domestic Product. That is why it is important to evaluate performance of the tourism sector for detecting possible areas of improvement. In this article, a combination of dynamic two-stage DEA model and fuzzy time series forecasting is used to investigate hospitality and site attraction efficiency in seven traditional beach touristic centers between January 2015 and December 2021. The results indicate significant differences between Stage 1 and Stage 2 efficiencies. Low site attraction efficiencies indicate opportunities for new policy decision-making strategies to strengthen tourism position both nationally and internationally.

Keywords: Data Envelopment Analysis, Fuzzy logic, Mexico, Tourism, Window Analysis

JEL Classification: C44, C61, L83
AMS Classification: 62M20, 91G70, 03C98.

1 Introduction

Tourism is one of the most important economic sectors in Mexico, because it has positioned itself as one of the main tourist destinations internationally and has promoted national, regional, and local development. In other words, the participation of the tourism sector went from contributing 6.9% in 2020 to 7.6% in 2021 of the Gross Domestic Product (GDP) at current prices. The arrival of national and international tourists to Mexico has been constantly growing during the last two decades. In 2022, Mexico registered arrival of 38.326 million of international tourists, ranking the country as the 6th most visited in the world. Constant growth of tourists arrivals resulted in a direct growth of a hospitality capacity across the country. At the end of 2022, Mexico offered 881,022 hotel rooms, representing 14.55% bigger capacity compared to 2016. Such a growth creates imminent pressure to guarantee efficiency in the whole tourism sector (SECTUR, 2023a).

To assess the tourism efficiency in Mexico, we used the Data Envelopment Analysis (DEA), one of the most used methodologies for measuring efficiency and performance (Assaf and Tsionas, 2019; Emrouznejad and Yang, 2018). DEA has also many applications in tourism with a focus on either hotel industry level efficiency or regional/state level efficiency. For example, in the hotel industry level Hathroubi et al. (2014) evaluated technical efficiency of 42 hotels in Tunisia considering environmentally responsible attributes. Higuerey et al. (2020) investigated the efficiency of 147 hotels in Ecuador during a period 2013-2017. Oukil et al. (2016) provided an analysis of the hospitality industry in Oman with a focus on hotel characteristics. In the case of regional/state level, Flegl et al. (2023) evaluated a hospitality efficiency in 32 Mexican states for a period from 1998 to 2018. Kido-Cruz et al. (2021) evaluated 59 touristic municipalities in Mexico to observe a link between tourism efficiency and poverty. Furthermore, several authors applied the DEA methodology to evaluate museums efficiency or nature/culture attractions. For example, Huang et al. (2022) evaluated 31 forest park provinces in China for a period of 2009 and 2018 regarding production and service efficiency. Wu and Lin (2022) assessed the performance of cultural tourism in 14 tourist destinations in Asia from 2015 to 2019.

Little attention has been placed on the combination of hospitality/tourism efficiency and local museum visits. Several studies evaluated the impact of culture, nature and local activities on hospitality (Oukil et al., 2016).

[1] Tecnologico de Monterrey, School of Engineering and Sciences, Calle Puente 222, Coapa, Arboledas del Sur, Tlalpan, 14380, Mexico City, Mexico, martin.flegl@tec.mx. ORCID: 0000-0002-9944-8475.
[2] Facultad de Gastronomía y Turismo, Universidad Anáhuac México, mcarmen.lozano@anahuac.mx. ORCID: 0000-0002-8579-4027.
[3] Tecnologico de Monterrey, School of Engineering and Sciences, Calle Puente 222, Coapa, Arboledas del Sur, Tlalpan, 14380, Mexico City, Mexico, A01027636@tec.mx.
[4] Mendel University, Faculty of Business and Economics, Zemědělská 1, 613 00 Brno, Czech Republic, marketa.matulova@mendelu.cz

However, the evaluation does not consider a network DEA model structure to evaluate the direct impact of the hospitality on sites visits. That is why, we constructed a two-stage dynamic DEA model using monthly data for seven traditional beach touristic centers in Mexico for a period 2015-2021. Stage 1 evaluates the hospitality efficiency considering hotels capacity and tourists arrival, whereas Stage 2 focuses on museums and archeological zones visits efficiency taking into account locations' attractivity. Further, we combined the obtained efficiency results with fuzzy time series forecasting to expose a future efficiency trend in each traditional beach touristic center.

2 Materials and methods

2.1 Data Envelopment Analysis

Data Envelopment Analysis is a non-parametric data-oriented approach for evaluating efficiency and/or performance of a set of homogeneous decision-making units (DMUs) according to their capability to transform m different inputs to s different outputs (Cooper et al., 2011). If a more complex production process is required, multistage models are used. In the presented analysis, a two-stage DEA production process is considered (Figure 1). In this case, we assumed that each DMU_j $(j = 1,2,\dots,n)$ has m inputs x_{ij}^A $(i = 1,2,\dots,m)$ used in the hospitality operations stage, which generates D intermediate outputs z_{dj} $(d = 1,2,..,D)$ and s outputs y_{rj}^A $(r = 1,2,\dots,s)$. Then the intermediates outputs become the inputs to the site attraction process stage with m additional inputs x_{ij}^B $(i = 1,2,\dots,m)$. Finally, Stage 2 generates s outputs y_{rj}^B $(r = 1,2,\dots,s)$ (Tisová and Flegl, 2023).

To observe changes in DMUs' efficiency over multiple time periods, the Window Analysis (WA) approach based on moving averages can be used. The WA approach is valuable in situations with insufficient number of DMUs in comparison to the number of available inputs and outputs. Using the Window Analysis, an efficiency of a DMU in a particular period is compared with its efficiency in other periods along with the efficiency of the other DMUs (Cooper et al., 2011). Therefore, there is $n.k$ DMU in each window, where n is the number of DMUs in each period and k is the width of each window. This feature increases the discriminatory ability (degree of freedom) of a DEA model (Dyson et al., 2001), as the total number of T periods is divided into series of overlapped periods (windows), each with a width $k(k < T)$ leading to $n.k$ DMUs. The first window has $n.k$ DMUs for periods $\{1,\dots,k\}$, the second period has $n.k$ DMUs and periods $\{2,\dots,k+1\}$, and so on, until the last window has $n.k$ DMUs and periods $\{T-k+1,\dots,T\}$. In total, there are $T-k+1$ separate analyses where each analysis examines $n.k$ DMUs (Cooper et al., 2011).

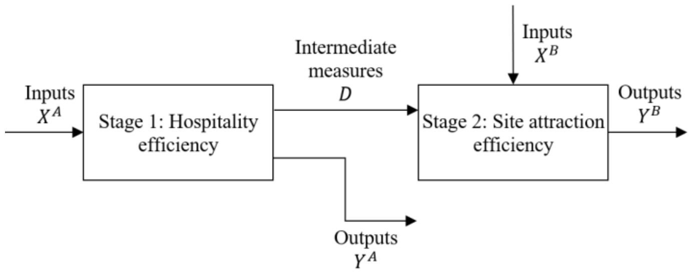


Figure 1 Two-stage production process for tourism efficiency analysis

2.2 Data and model structure

The analysis uses monthly data related to hospitality activities in Mexico (SECTUR, 2023b) in seven traditional beach touristic centers since January 2015 to December 2021: Acapulco, Cozumel, La Paz, Manzanillo, Mazatlán, Puerto Vallarta and Veracruz – Boca del Río. The hospitality activities are divided into a two-stage production process consisting of hospitality and site attraction stages. As Lee et al. (2010) and Oliani et al. (2011) pointed out, hotels' quality and capacity are important factors in tourism evaluation. In this case, star rating is commonly used to express hospitality quality. Therefore, rooms availability in each touristic center divided by star rating is utilized for the inputs in Stage 1. The efficiency of the hospitality process is evaluated regarding the rooms occupation rate (ROR) and total number of tourist nights (TN). The ROR is taken as the only output of Stage 1 (y_{rj}^A), whereas the TN is treated as an intermediate output (z_{dj}). Similar model structures were used, for example, by Hathroubi et al. (2014), Oukil et al. (2016), or Flegl et al. (2023).

The objective of Stage 2 is to evaluate each beach touristic center's efficiency in site attraction process, i.e., how well they promote museums and archeological zones. The intermediate input of tourist nights represents a set of potential visitors. The more tourists visit each touristic center and the more nights these tourists stay, the bigger the chance of site attraction visits. This chance of visits is stimulated by the availability of site attractions, which is not only limited to sites located within each touristic center but should also include attraction sites around each touristic center. So, first, we calculated each site quality as a normalized value between 0 and 100 taking into account the total number of visits during the last 22 years (separately for 104 museums and 182 archeological zones operated by Instituto Nacional de Antropología e Historia). Second, the impact of the site attraction process is limited to a certain distance. Therefore, we decided to limit the impact by 2-hour radius (considering an average speed of 80 km/h), where the impact within 30 minutes is 100%, and then decreases by 25% every 30 minutes. Museums or archeological zones outside the radius of 2 hours have zero impact on the site attraction efficiency. As a result, Stage 2 includes two inputs x_{ij}^{B} reflecting a weighted sum of museums (MUS) and a weighted sum of archeological zones (AZ) attraction within the 2-hour radius in each touristic center. Finally, the outputs of Stage 2 include the total number of museums visits (MUSVI) and total number of archeological zones visits (AZVIS), using the same weighted sum procedure as in the case of the inputs.

Table 1 summarizes the descriptive statistics of the used variables. The analysis covers 84 months ($T = 84$) considering national and international tourists together. The variable returns to scale DEA model was used, and the WA considers 12-month window ($k = 12$) to eliminate the impact of seasonality. The MaxDEA Ultra 7 software was used for all calculations.

Variable	1-star rooms	2-star rooms	3-star rooms	4-star rooms	5-star rooms	Tourists night	Rooms occupation rate	Museums	Archeological zones	Museum visits	Archeological zones visits
Minimum	0.00	4,152.00	5,562.00	16,138.00	7,452.00	0.00	0.00	0.50	0.00	0.00	0.00
Maximum	49,276.00	59,187.00	117,371.00	221,134.00	218,220.00	1,272,233.00	0.90	10.50	43.50	265,558.00	10,365,507.00
Mean	12,927.00	23,270.00	48,311.00	85,472.00	98,248.00	334,390.00	0.48	3.04	7.04	20,155.00	753,909.00
StDev	14,228.00	19,239.00	31,766.00	57,776.00	73,265.00	257,239.00	0.17	3.63	14.93	47,670.00	2,068,636.00

Table 1 Descriptive statistics of the used inputs and outputs

2.3 Fuzzy time series

Consider the universe of discourse $U \subset R$ in which the fuzzy sets $A_1, A_2, \dots, A_n$ with membership functions μ_{A_i} are defined. Let $\{y(t), t = 0,1, \dots, N\}$ be a time series taking values in U. A fuzzy time series (FTS) $\{F(t), t = 0,1, \dots, N\}$ is a fuzzification of $\{y(t)\}$. In this work, the fuzzification of the time series $\{y(t)\}$ will be the FTS given by $F(t) = \{\mu_1(y(t)), \cdots, \mu_n(y(t))\}$. An FTS model is a collection of rules in the form $precedent \rightarrow consequent$ that shows how the fuzzy sets A_i relates to each other over time. The process of analyzing a time series using fuzzy logic can be summarized in a train phase and a forecast phase. The train phase consists in the following steps (Lucas et al., 2021): Step 1- The partitioning of the universe of discourse. Step 2- The fuzzification of the time series to turn it into an FTS. Step 3- Temporal pattern extraction. Extract information of the FTS F due to previous observations using the rules assigned in the model of the FTS. The forecast phase can also be divided into three steps: Step 1- Fuzzification of the sample that is going to be used in the forecasting. Step 2- For each element $f(t) \in F$, find the rules that contains $f(t)$ as precedent. Then apply these rules. Step 3- Deffuzification, that is, transforming the value $\hat{F}(t+1)$ into a crisp value $\hat{y}(t+1)$. In this work, we use the Weighted Multivariate Fuzzy Logical Rule Group (WMFLRG) provided in the package pyFTS to perform the extraction of the temporal pattern and to match the rules in the Step 2 of the forecast phase. In a similar manner, we also used the package pyFTS to perform the other Steps both in the train and forecast phases.

3 Results

The results are divided into two main parts: first, we describe the obtained results regarding the hospitality process efficiency and, second, we describe the results in the site attraction efficiency. In both parts, we also analyze the results of the forecasting part of the analysis.

Table 2 summarizes the yearly average efficiencies of the seven traditional beach touristic centers. The average efficiency of all centers during the analyzed period was 0.758. This result can be considered as a high efficiency level. The highest efficiency can be observed in the case of La Paz (0.839) and Manzanillo (0.801), whereas the lowest efficiency was reported in Veracruz - Boca del Río (0.669, -0.089 below the average) and Acapulco (0.677, -0.081). Considering the yearly averages, we can see a stable slight growth from 2015 (0.781) to 2019 (0.831), which was interrupted in 2020 due the global pandemic crisis of Covid-19 and average efficiency of 0.519. The hospitality efficiency was then partially recovered in 2021 reaching an average efficiency of 0.755.

Touristic center	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**Average**
Acapulco	0.598	0.583	0.756	0.769	0.826	0.462	0.746	**0.677**
Cozumel	0.815	0.889	0.815	0.829	0.848	0.441	0.838	**0.782**

La Paz	0.910	0.887	0.836	0.863	0.885	0.678	0.800	**0.839**
Manzanillo	0.802	0.834	0.841	0.854	0.849	0.583	0.849	**0.801**
Mazatlán	0.746	0.758	0.765	0.772	0.806	0.670	0.819	**0.763**
Puerto Vallarta	0.818	0.843	0.857	0.837	0.863	0.496	0.717	**0.776**
Veracruz – Boca del Río	0.782	0.741	0.706	0.768	0.744	0.410	0.522	**0.669**
Average	**0.781**	**0.791**	**0.797**	**0.813**	**0.831**	**0.519**	**0.756**	**0.758**

Table 2 Yearly average efficiency scores in the hospitality efficiency

To analyze the results more in detail, Figure 2 displays the evolution of the monthly hospitality efficiencies. First, we can observe a similar evolution of the efficiencies in all seven analyzed beach touristic centers. We can notice highest monthly efficiencies in July of each year, followed by the decline in September. This is due to the summer vacations. Further, across the whole evaluated period, we can notice low deviations around the average. Only Veracruz - Boca del Río and Acapulco show higher deviations from the average. Second, we can observe the impact of the Covid-19, as all touristic centers were literarily shut down due to the government restrictions. In the second half of 2020 and 2021, the hospitality sector slowly recovered. The figure indicated that the recovery pace was different, as for example, Veracruz - Boca del Río recorded an average efficiency of only 0.522 in 2021, compared to Manzanillo 0.849, Cozumel 0.838, and Mazatlán 0.819.

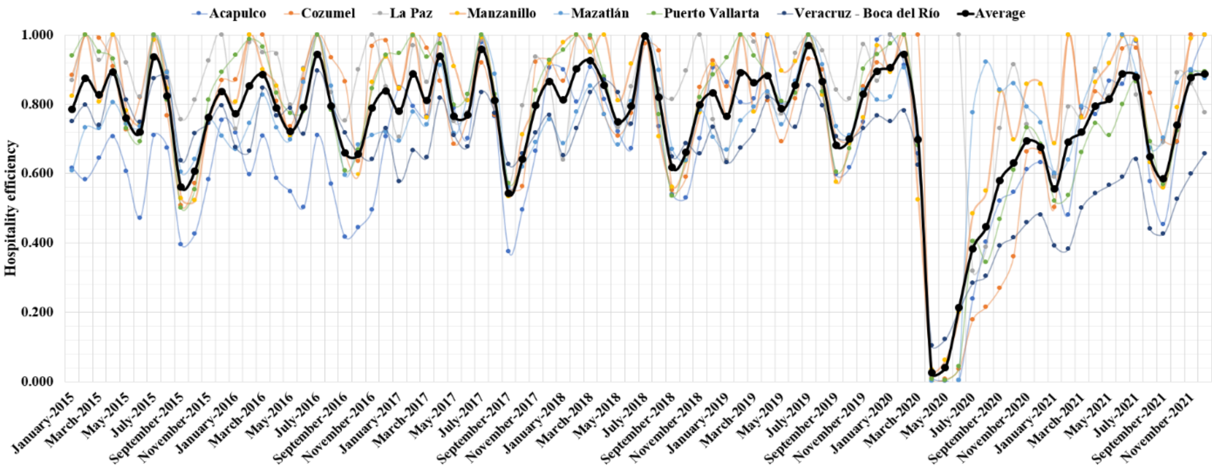


Figure 2 Efficiency evolution in the hospitality efficiency from 01/2015 to 12/2021

In Stage 2, the average efficiency of the site attraction process was 0.545 (Table 3). This is a significant difference compared to Stage 1 (-0.213). In this case, the results revealed much bigger differences between the touristic centers and much higher deviations from the average. Cozumel was evaluated as the best with an efficiency of 0.893, followed by Manzanillo (0.831). On the other side, the worst evaluated touristic center was Puerto Vallarta with very low efficiency of 0.165 (-0.728 compared to Cozumel), Acapulco (0.278) and La Paz (0.399). Surprisingly, the Covid-19 did not cause a significant drop in the efficiency in 2020 and 2021 as in both years the average efficiency was higher than the average (0.698 and 0.662 respectively).

Touristic center	2015	2016	2017	2018	2019	2020	2021	Average
Acapulco	0.255	0.249	0.272	0.241	0.234	0.240	0.462	**0.278**
Cozumel	0.879	0.895	0.899	0.923	0.962	0.839	0.849	**0.893**
La Paz	0.224	0.231	0.397	0.290	0.332	0.710	0.856	**0.399**
Manzanillo	0.755	0.835	0.779	0.761	0.922	0.972	0.928	**0.831**
Mazatlán	0.528	0.444	0.489	0.506	0.697	0.629	0.798	**0.559**
Puerto Vallarta	0.149	0.031	0.013	0.176	0.319	0.028	0.436	**0.165**
Veracruz – Boca del Río	0.702	0.737	0.643	0.681	0.701	0.609	0.741	**0.693**
Average	**0.499**	**0.489**	**0.499**	**0.511**	**0.595**	**0.698**	**0.662**	**0.545**

Table 3 Yearly average efficiency scores in the site attraction efficiency

Figure 3 indicates that the distinct impact of the Covid-19 on the site attraction efficiency was caused by different approaches established in each touristic center. For example, sites around Cozumel were not completely shut down, whereas the sites around La Paz were completely shut down only in April, August and September in 2020. Considering a combination of the Window Analysis approach and the selected moving averages of 12 months, together with lower numbers of site visit compared to the hospitality, Covid-19 had distinct impact on the efficiencies. Further, the effect of summer vacations in July and September is not that visible in this stage. Similarly, the impact of Covid-19 shows significant disruption in the post-Covid-19 period. Although the efficiency indicates a recovery and growth, the pre-pandemic pattern was not established. If we analyze the hospitality efficiency in Figure 2, we can see reestablished pattern in 2021 with a drop in September following the summer vacations.

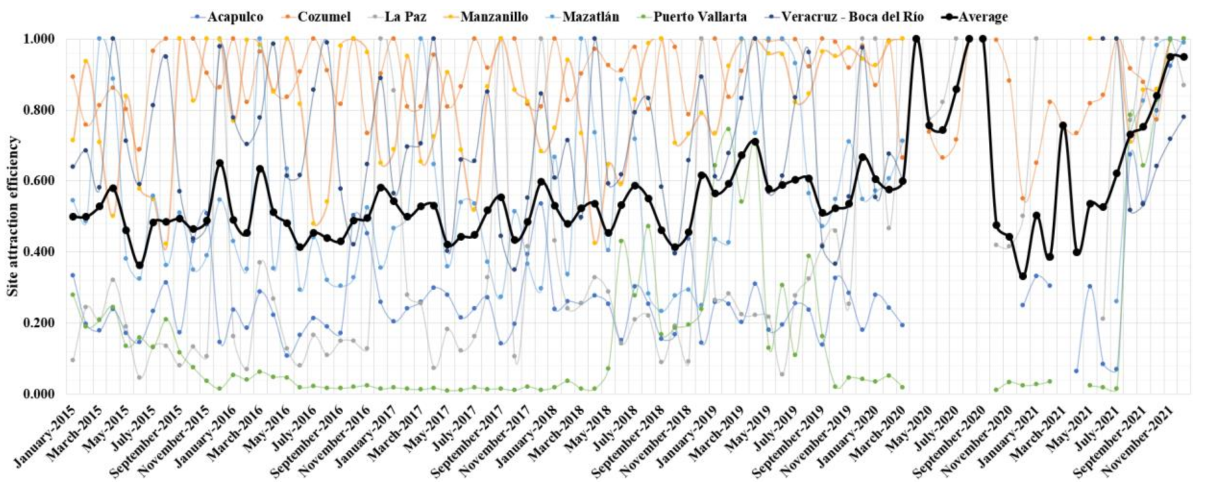


Figure 3 Efficiency evolution in the site attraction efficiency from 01/2015 to 12/2021

Finally, to see the evolution of the efficiencies in both stages behind the evaluated period, Figure 4 and Figure 5 illustrates the forecasted values until January 2025 (i.e., 37 months). It is clear that the average hospitality efficiency remains high 0.882 (with a lower bound 0.827 and an upper bound 0.882) and all beach touristic centers retook their pre-covid pattern. Similarly, the site attraction efficiency remains low 0.535 (lower bound 0.499 and upper bound 0.570). Further, the results indicate significant differences between the beach touristic centers, but with lower deviations magnitudes.

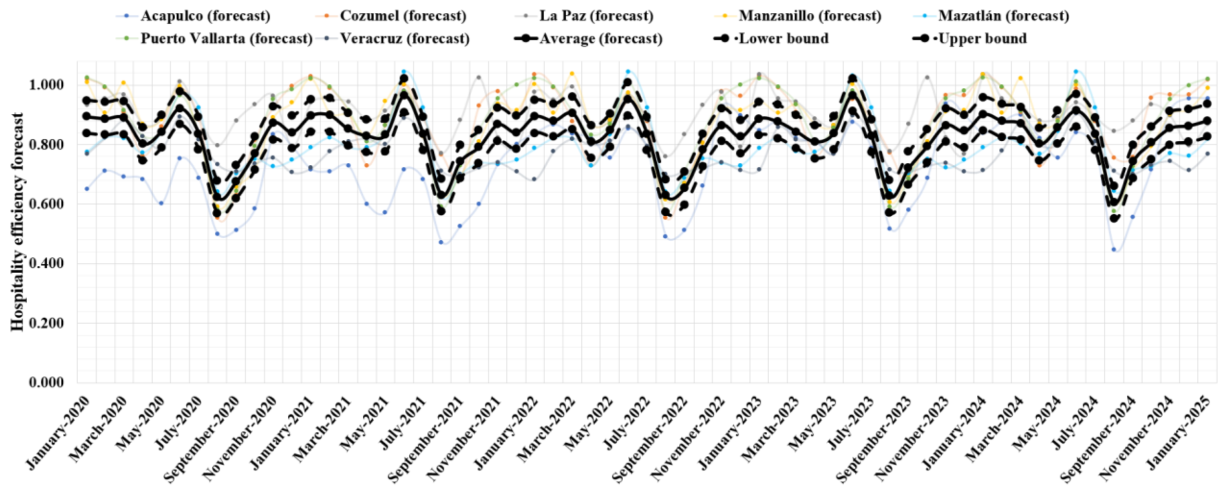


Figure 4 Forecast of the hospitality efficiency towards January 2025

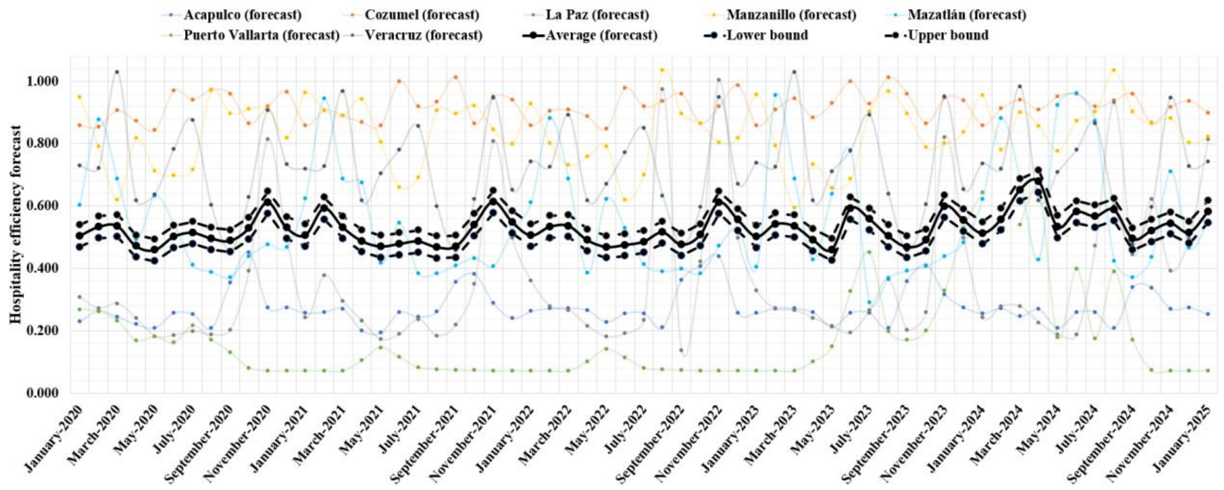


Figure 5 Forecast of the site attraction efficiency towards January 2025

4 Conclusion

The presented analysis presented a unique two-stage DEA process connecting hospitality and site attraction processes for seven traditional beach touristic centers in Mexico. For this purpose, the Window Analysis approach was used on a period since January 2015 until December 2021. The results revealed a high average hospitality efficiency with low level of deviations and, on the other hand, low site attraction efficiency with high level of deviations. The further steps in the research can go in several directions. First, the analysis would include more years (months) as data for 2022 and 2023 were recently published. Second, the analysis can include more beach touristic centers to observe differences between the traditional centers and 19 more beach touristic centers registered by the Secretariat of Tourism. Third, the analysis would also use the Malmquist Index approach to investigate the effect of a technological change on the efficiency.

References

Assaf AG, Tsionas MG (2019) A review of research into performance modeling in tourism research - Launching the Annals of Tourism Research curated collection on performance modeling in tourism research. *Annals of Tourism Research* 76: 266-277. https://doi.org/10.1016/j.annals.2019.04.010

Cooper W, Seiford LM, Zhu J (2011) *Handbook on Data Envelopment Analysis*. New York: Springer.

Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA (2001) Pitfalls and protocols in DEA. *European Journal of Operational Research* 132: 245-259. http://dx.doi.org/10.1016/S0377-2217(00)00149-1

Emrouznejad A, Yang G-L (2018) A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences* 61: 4-8. https://doi.org/10.1016/j.seps.2017.01.008

Flegl M, Cerón-Monroy H, Krejčí I, Jablonsky J (2023) Estimating the hospitality efficiency in Mexico using Data Envelopment Analysis. *OPSEARCH* 60(1): 188-216. https://doi.org/10.1007/s12597-022-00619-8

Hathroubi S, Peypoch N, Robinot E (2014) Technical efficiency and environmental management: The Tunisian case. *Journal of Hospitality and Tourism Management* 21: 27-33. https://dx.doi.org/10.1016/j.jhtm.2014.03.002

Higuerey A, Viñan-Merecí C, Malo-Montoya Z, Martínez-Fernández V-A (2020) Data Envelopment Analysis (DEA) for Measuring the Efficiency of the Hotel Industry in Ecuador. *Sustainability* 12(4): 1590. https://doi.org/10.3390/su12041590

Huang X-J, An R, Yu M-M, He F-F (2022) Tourism efficiency decomposition and assessment of forest parks in China using dynamic network data envelopment analysis. *Journal of Cleaner Production* 363: 132405. https://doi.org/10.1016/j.jclepro.2022.132405

Kido-Cruz A, Cruz-Pérez JP, Ortiz-Zavala A (2021) Turismo y pobreza: Un análisis de la envolvente de datos para municipios en México. *Inquietud empresarial* 21(2): 85-95. https://doi.org/10.19053/01211048.12334

Lee C, Huang H-I, Yeh H-R (2010) Developing an evaluation model for destination attractiveness: sustainable forest recreation tourism in Taiwan. *Journal of Sustainable Tourism* 18(6): 811-828. https://dx.doi.org/10.1080/09669581003690478

Lucas PO, Orang O, Silva PC, Mendes EMAM, Guimaraes FG (2021) A Tutorial on Fuzzy Time Series Forecasting Models: Recent Advances and Challenges. *Learning and Nonlinear Models* 19(2): 29-50. https://dx.doi.org/10.21528/lnlm-vol19-no2-art3

Oliani LGN, Rossi GB, Gervasoni VC (2011) What are the attractiveness factors that influence the choice of a tourist destination—A study of Brazilian tourist Consumer. *Chinese Business Review* 10(4): 286-293. https://dx.doi.org/10.17265/1537-1506/2011.04.005

Oukil A, Channouf N, Al-Zaidi A (2016) Performance evaluation of the hotel industry in an emerging tourism destination: The case of Oman. Journal of Hospitality and Tourism Management, 29: 60-68. https://dx.doi.org/10.1016/j.jhtm.2016.05.003

Tisová P, Flegl M (2023) Underwriting and Investment Efficiency in the Czech Life Insurance Sector: A Two-stage DEA Window Analysis Approach. *In 41st International Conference on Mathematical Methods in Economics 2023*. Conference proceeding. Prague University of Economics and Business, Prague, pp 407-413.

SECTUR (2023a) *Ranking Mundial del Turismo Internacional*. Secretaría de Turismo de México [Mexican Secretary of Tourism]. https://www.datatur.sectur.gob.mx/SitePages/RankingOMT.aspx (accessed 15 January 2024).

SECTUR (2023b) *Compendio Estadístico del Turismo en México 2022.* DATATUR. https://datatur.sectur.gob.mx/SitePages/CompendioEstadistico.aspx (accessed 10 January 2024).

Wu Y-C, Lin S-W (2022) Efficiency evaluation of Asia's cultural tourism using a dynamic DEA approach. *Socio-Economic Planning Sciences* 84: 101426. https://doi.org/10.1016/j.seps.2022.101426

Model of network interconnections

Petr Fiala[1], Renata Majovská[2] Lorena Caridad López del Río[3]

Abstract. In today's economy, many activities are networked. The importance of network industries that deliver products and services is growing. Traditionally, only the effects of competition have been analysed within networks. Increasingly, companies are discovering that even cooperation between competitors can bring benefits to all involved. Network interconnections are analysed where networks not only provide services within their own network but also provide access to others' networks. This paper models and analyses co-opetition in network industries using biform games as a combination of non-cooperative and cooperative game theory. The authors propose a division of biform games into sequential and simultaneous games. A proposed basic model of a network industry with two networks is defined to analyse the interconnections. The model can be solved as a sequential biform game and as a two-stage non-cooperative game. The results are analysed and an illustrative example is solved.

Keywords: network industry, competition, cooperation, co-opetition, games

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

The network economy is the name given to today's global relationships between economic actors, which are characterized by strong interconnections (Bobzin, 2006, Shy, 2001). The main role in the new era is played by the interconnection of anything with anything in a vast network of relationships at different levels, where resources and activities are shared, markets are expanded and risk costs are reduced. Today's network industries provide the infrastructure and foundation for the functioning of the economy and society. Specific network effects can be observed both in physical networks and in virtual networks of users of the same products. Network industries take many forms and include material networks such as transport, logistics, communications, energy, water and wastewater, but also abstract economic, financial, social and knowledge networks. Significant technological changes, in particular digitalisation, have further transformed traditional public services and given rise to network industries, in particular digital platforms.

The traditional economic approach is devoted to the analysis of competition in network industries. However, it has also been found that cooperation can bring benefits to all involved. The concept of co-opetition (Brandenburger and Nabuleff, 2011) is a business strategy that enriches the rules of competition and cooperation by combining the benefits of both. Biform game models (Brandenburger and Stuart, 2007) offer themselves as a suitable modelling tool for the concept of co-opetition because they combine cooperative and non-cooperative game theory approaches (Mayerson, 1997).

Using this modelling framework based on the concept of co-opetition and biform games, specific models for network industries can be designed and analysed. The models analyse network industries in terms of total number of members and number of cooperating members in terms of service prices and network access fees and are analysed using simultaneous biform games. Network interconnections are also analysed, where networks not only provide services on their own network but also provide access to other networks. The model can be analysed as a sequential biform game.

The rest of the article is organized as follows. Section 2 introduces the elements of the co-opetition concept. Basic biform games are used to model co-opetition problems. Section 3 introduces the basic model of network industries. A sequential biform game for analysing the basic model is presented in Section 4. Section 5 is devoted to the

[1] Prague University of Economics and Business, Department of Econometrics, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic, pfiala@vse.cz

[2] University of Finance and Administration, Prague, Department of Computer Science and Mathematics, Estonská 500, 101 00 Praha 10, Czech Republic, renata.majovska@mail.vsfs.cz

[3] University of Seville, Department of Financial Economics and Operations Management, Avenida Ramón y Cajal 1, 41018 Seville, Spain, lcaridad@us.es

analysis of the basic model using a two-stage noncooperative game. An illustrative example is solved in Section 6. Section 7 presents the conclusions.

2 Co-opetititon concept

The traditional literature on network analysis focuses on the study of competitive relationships (Armstrong 1998, Laffont-Rey-Tirole 1998a, 1998b). However, it has also been found that cooperation can bring benefits to all involved. Considering competitive and cooperative pressures simultaneously in the analysis of network industries is a challenge for further research. The aim of this paper was to propose a modelling framework for the analysis of network industries and to explore the simultaneous observation of the effects of competition and cooperation in special models.

The concept of co-opetition (Brandenburger and Nabuleff, 2011) is a business strategy that enriches the rules of competition and cooperation by combining the advantages of both. This concept delineates the different elements of the system under study and focuses on the search for so-called complementors, which are competitors that bring added value as a result of the extension of the capabilities of the analysed entity, thus extending competition by cooperation. The corresponding PARTS model contains five items that form the basis of the functioning of the concept of co-opetition in the search for a solution to the situation under analysis: Players, Added Value, Rules, Tactics, Scope.

Players are divided by type into producers, customers, suppliers, competitors and complementors. Complementors provide added value if their products provide an extension of the possibilities for the producers' own products. The rules structure the negotiation between players. Tactics are sequences of activities that constitute the monitoring of the negotiation process by other players. The scope is determined by the interrelationships between the PART elements of the game model and the other possible games in which the players of that model participate.

The concept of co-opetition can be analysed using game theory models (Okura and Carfi, 2014). Biform game models (Brandenburger and Stuart 2007) are offered as a suitable modelling tool for the concept of co-opetition as they combine cooperative and non-cooperative game theory approaches. We propose a division of biform games into sequential and simultaneous games (Fiala and Majovská 2020).

The sequential biform game can be solved consequently by application of non-cooperative and cooperative approaches. The game consists of two phases. In one phase, one type of game models (cooperative or non-cooperative) is used, and in the other phase, the additional type following the previous part is used in the sequence. In this article, the first phase will be cooperative and the second will be non-cooperative. In the first phase, the maximum output for the coalition of all players is calculated and certain parameters are determined cooperatively with regard to the interests of the participants. In the second phase, players compete on other parameters of the model and the solution of the situation is determined non-cooperatively, for example using the concept of Nash equilibrium.

The sequential biform game can then be solved using non-cooperative and cooperative approaches consequently. The game consists of two phases. In one phase, one type of game models (cooperative or non-cooperative) is used and in the second phase another type is used following the previous part of the sequence. In this paper, the model where the first phase is cooperative and the second is non-cooperative will be specifically analysed. In the first phase, the maximum output for a coalition of all players is calculated and certain parameters are determined cooperatively with respect to the interests of the participants. In the second phase, the players compete for other parameters of the model and the solution to the situation is determined non-cooperatively, for example using the concept of Nash equilibrium. When analysing network industries, there are also models where the order of phases will be reversed.

The simultaneous biform game has one phase in which cooperative and non-cooperative approaches are used simultaneously to find a consensual solution to the situation. Finding this consensus solution, however, can be done through multi-round negotiations. The situation is influenced by the composition of the coalition of cooperating players and their level of cooperation. Relationships between players can be cooperative and non-cooperative at the same time. For example, the relationships between producers and complementors are competitive because complementors provide competitive substitute products and at the same time cooperative because they add value by extending the use of producers' products. It is possible to formulate a number of network models in which a simultaneous biform game will be used for the analysis (e.g. Fiala, 2023). Different constraints affect the players who are under pressure and thus determine the level of cooperation. The level of cooperation can change over time and can be measured by several criteria. This offers an extension of the models to dynamic and multi-criteria.

3 Basic model

A network industry with two networks is defined for the interconnection analysis (Kim 2004). This simplified situation allows for more clarity in the results. The interconnection model considers the following simplification of the network industry situation. Bertrand's duopoly model is used as a basis, which is analysed as a game in which strategic decisions are based on prices. We consider the interaction of two networks providing substitutable network services. This is a model based on the concept of co-opetition, where the players are the networks and the users of the networks. The networks compete with each other in the network services they provide, but at the same time they provide added value in access to the interconnection of the two networks. The networks thus become complementors.

We will use symmetric demand functions in the standard linear form for substitution services

$$q_i(p_i, p_j) = b_0 - b_1 p_i + b_2 p_j,\ i, j = 1, 2, \tag{1}$$

where q_1, q_2 are demands for the network services, p_1, p_2 are network competitive prices and b_0, b_1, b_2 are parameters of the demand functions.

Networks have the same composition of cost, consisting of a fixed component f and a variable component c. Variable costs consist of the costs of entering and leaving the network, each of size c_0, and the costs in between c_1. Total variable cost is therefore equal to $c = 2c_0 + c_1$.

Networks have the same composition of costs, consisting of a fixed component f and a variable component c. Variable costs consist of the costs of entering and leaving the network, each of size c0, and the costs in between, c1. Thus,

Networks provide interconnection options between networks and this requires actions to enter a competing network, which are charged using so-called access fees (a_1, a_2). However, these access fees may be set arbitrarily by the networks or by a third party.

The profit of each network is given by the following formula

$$\pi_i(p_i, p_j, a_i, a_j) = [(p_i - c)q_i - f] + \left[(a_i - c_0)q_j - \left(a_j - c_0\right)q_i\right],\ i, j = 1, 2. \tag{2}$$

The networks try to maximize their profits.

Two different game schemas are used for analysis:

- The sequential biform game: in the first phase of the game, networks choose access fees in a cooperative way; in the second phase, networks compete non-cooperatively in prices.
- Two-stage non-cooperative game: in the first phase, networks compete on access fees; in the second phase, networks compete on prices.

In the following sections in this paper, these approaches will be analysed, compared and illustrated with an example.

4 Sequential biform game

The basic model can be analysed as a sequential biform game. In the first phase, the goal of cooperation is to maximize the total profit for the network industry.

The total profit for the network industry

$$\Pi(p_1, p_2) = [(p_1 - c)q_1 - f] + [(p_2 - c)q_2 - f] \tag{3}$$

does not include access payments because total access payments between networks are offset. However, the access payment affects the price and hence the profit.

Assume that the networks themselves or with the help of a third party, cooperatively agree on a reasonable level of the reciprocal access fee a_r for both networks. We will return to the discussion on the level of the access fee a_r after the non-cooperative pricing phase.

In the second phase, the networks compete for prices (p_1, p_2). We use the concept of Nash equilibrium to deal with this non-cooperative situation. The profit of each network at a given access fee a_r is given by the following formula

$$\pi_i(p_i, p_j) = [(p_i - c)q_i - f] + \left[(a_r - c_0)q_j - (a_r - c_0)q_i\right],\ i, j = 1, 2. \tag{4}$$

The Nash equilibrium is determined by solving the following system of equations. The first order conditions for profit maximization are

$$\frac{\partial \pi_i(p_i,p_j)}{\partial p_i} = 0,\ i = 1, 2. \tag{5}$$

Given the symmetry between networks, there is a symmetric Nash equilibrium

$$p_r = p_1 = p_2 = \frac{b_0+b_1c}{2b_1-b_2} + \frac{b_1+b_2}{2b_1-b_2}(a_r - c_0). \tag{6}$$

The total price can be divided into two components. The first component is determined by competition on prices. The second component is driven by the cost of access to the network.

Discussion of the level of the access fee:

The final price is the same as the price in the typical Bertrand competition, if the access fee is set to $a_r = c_0$,

$$p^* = \frac{b_0+b_1c}{2b_1-b_2}. \tag{7}$$

If network services are to function as a public service and maximize user welfare, then network profit is zero. In such a situation, the access charge a_0 and the price p_0 are denoted.

The access fee a_0 may be reciprocally calculated from (5)

$$a_0 = c_0 + \frac{2b_1-b_2}{b_1+b_2}\left(p_0 - \frac{b_0+b_1c}{2b_1-b_2}\right). \tag{8}$$

From the relation (8) it follows

$$\text{if } a_0 < c_0 \text{ then } p_0 < p^*. \tag{9}$$

Cooperation through setting the access fee can effectively control the price. The relation $a_0 < c_0$ does not imply that the network gains a loss from providing access to the network. The requirements for access to competing networks are balanced at equilibrium.

5 Two-stage non-cooperative game

To address this situation, a two-stage non-cooperative game model is used to determine access fees and prices for network services. In the first stage, the networks calculate the access fees by solving the non-cooperative game. The second stage is devoted to calculating prices also using the non-cooperative game. We will use backward induction. The profit of each network is given by formula (2). The calculation starts by determining the prices, taking the non-reciprocal access fees as given.

The first order conditions for profit maximization are

$$\frac{\partial \pi_i(p_1,p_2)}{\partial p_i} = 0,\ i = 1, 2. \tag{10}$$

The solution of the system provides an expression of the network profits in terms of access fees. We then proceed with the first stage of calculating the access fees backwards.

The symmetric Nash equilibrium can be calculated from the following conditions

$$\frac{\partial \pi_i(a_1,a_2)}{\partial a_i} = 0,\ i = 1, 2. \tag{11}$$

The solution of the system (10) gives a symmetric access fee $a_c = a_1 = a_2$ and a symmetric final price

$$p_c = p_1 = p_2 = \frac{b_0+b_1c}{2b_1-b_2} + \frac{(b_1+b_2)}{2b_1-b_2}(a_c - c_0). \tag{12}$$

The access fees a_r and a_c differ depending on how they were set, whether cooperative or non-cooperative. The prices p_r and p_c are then determined depending on the access fees a_r and a_c. There is a relationship between the equilibrium price and the equilibrium access fee such that the price increases if the access fee also increases.

The price is higher when networks compete on prices and also on access charges than when access charges are set cooperatively, because access charges can be agreed below marginal costs c_0, which is offset by the benefit to users and accompanied by an increase in the number of users. The access fees set non-cooperatively will be higher than marginal costs.

For the situations analysed, the following relationship between prices and access charges applies

$$p_0 < p^* < p_c \text{ and } a_0 < c_0 < a_c. \tag{13}$$

Determining the correct level of access fees is influenced by a number of implications that need to be considered. The concept of pressure can be used to model and analyse this situation, where the negotiating parties change the aspiration levels of the criteria in an attempt to reach consensus.

Networks have to respond to pressures that, on the one hand, push prices down to attract more network users and, on the other hand, raise access charges to raise competitors' prices. Raising access charges raises prices and thus puts negative pressure on the welfare of network users. On the other hand, lowering access charges will lead to increased cooperation between networks and possible tacit agreement.

6 Illustrative example

We illustrate the approaches of the sequential biform game and the two-stage non-cooperative game on a numerical example and compare the results. We set parameters of the demand functions $b_0 = 400$, $b_1 = 20$, $b_2 = 10$ and cost parameters $f = 0$, $c_0 = 3$, $c_1 = 4$, $c = 2c_0 + c_1 = 10$.

We solve the problem as the sequential biform game. In the second phase, the networks compete on prices (p_1, p_2). The symmetric Nash equilibrium is given using the formula (6)

$$p_r = p_1 = p_2 = 20 + (a_r - 3).$$

In the first phase, negotiate networks on the access fee a_r (see discussion of the level of the access fee).

If $a_r = c_0 = 3$, the price is given using the formula (7)

$$p^* = 20,$$

total profit for the network industry is given using formula (3)

$$\Pi(p^*, p^*) = [(20 - 10)200] + [(20 - 10)200] = 4000,$$

and profits for networks are given using the formula (4)

$$\pi_i(p^*, p^*) = [(20 - 10)200] = 2000, i = 1, 2.$$

If $a_r < c_0$ then $p_r < p^*$ and profits will be lower. For e.g. $a_r = 1$, we get

$$p_r = 18, \pi_i(p_r, p_r) = [(18 - 10)220] = 1760, i = 1, 2,$$

$$\Pi(p_r, p_r) = 3520.$$

For a two-stage non-cooperative game, we get from (12)

$$p_c = p_1 = p_2 = 20 + (a_c - 3).$$

If $a_c > c_0$ then $p_c > p^*$ and profits will be higher. For e.g. $a_c = 4$, we get

$$p_r = 21, \pi_i(p_c, p_c) = [(21 - 10)190] = 2090\ i, j = 1, 2, \Pi(p_c, p_c) = 4180.$$

The difference is in the process of setting access fees, whether it is cooperative or non-cooperative. Setting access fees controls prices and profits.

7 Conclusion

Behaviour in network industries is an important subject of intensive economic research. There has been a shift from the initial investigation of only network competition to the investigation of richer relationships in networks that include cooperation between competitors. This paper presents a basic modelling framework for analysing network co-opetition. The concept of co-opetition captures the joint operation of competitive and cooperative relationships within network industries. The modelling framework is based on the concept of co-opetition, which is modelled using biform games as a combination of non-cooperative and cooperative games. The biform games are divided into sequential and simultaneous games according to the way non-cooperative and cooperative game theory approaches are combined.

The modelling framework allows to create specific models to analyse certain relationships in networks. By connecting and disconnecting these specific models, it is possible to analyse a targeted part of interest within the network industry. This paper analyses a specific model that captures the number of cooperating units in a network and the interconnectedness of networks. The model is simple, yet allows important managerial implications to be drawn.

The proposed model framework and the specific problem are very flexible and can be enriched and extended with additional features. Possible modifications to the model framework include the use of additional general tools for modelling co-opetition on networks and procedures for their analysis, the analysis of different options for linking cooperating and non-cooperating parts, the use of multi-criteria evaluation in general, additional negotiation procedures, etc. Another important part is to create other specific models and to consider ways of connecting and disconnecting specific models. Experimentation with models can have important managerial implications that can be translated into improvements in real policies affecting network industries.

Acknowledgements

This work was supported by the grant No. IGA F4/18/2024, Faculty of Informatics and Statistics, Prague University of Economics and Business, Prague.

References

Armstrong M (1998) Network Interconnection in Telecommunications. The Economic Journal 108(448): 545–564

Bobzin H (2006) Principles of Network Economics. Springer, Berlin

Brandenburger AM, Nalebuff BJ (2011) Co-opetition. Crown Business, New York

Brandenburger A, Stuart H (2007) Biform games. Management science 53 (4): 537-549

Fiala P (2023) Model of a network industry. In: Proceedings of the 41st International Conference on Mathematical Methods in Economics, Praha: 74–79

Fiala P, Majovská R (2020). Modeling the Design Phase of Sustainable Supply Chains. In: Khan SAR (ed) Global Perspectives on Green Business Administration and Sustainable Supply Chain Management. IGI Global, Hershey PA

Kim C (2004) A Simple Approach to Network Competition. Journal of Economic Development, 29(2): 163-176

Laffont JJ, Rey P, Tirole J (1998a) Network Competition: I. Overview and Nondiscriminatory Pricing. RAND Journal of Economics 29(1): 1-37

Laffont JJ, Rey P, Tirole J (1998b) Network Competition: II. Price Discrimination. RAND Journal of Economics 29(1): 38-56

Myerson RB (1997) Game Theory: Analysis of Conflict. Harvard University Press, Cambridge

Okura M, Carfì D (2014) Coopetition and Game Theory. Journal of Applied Economic Sciences 9(3): 1-29.

Shapley LS (1953) A value for n-person games. In: Tucker AW, Luce RD (eds) Contributions to the Theory of Games II. Princeton University Press, Princeton

Shy O (2001) The Economics of Network Industries. Cambridge University Press, Cambridge

Accessibility of public charging infrastructure for electric vehicles across Central European countries: a geospatial analysis.

Jindra Lacko[1], Tomáš Formánek[2]

Abstract. This study models spatial distribution of electric vehicle (EV) charging stations across eight EU nations, a pivotal aspect in the decarbonization strategy of the transportation domain. Employing geolocated open source dataset, augmented with official statistical data, we conduct a quantitative analysis on a finely granulated hexagonal grid. Our findings demonstrate that zero-inflated two stage models exhibit superior performance compared to conventional Poisson count model in this specific context. Moreover, our modeling endeavors reveal that a substantial portion of the variability in the geographic dispersion of EV chargers can be explained by country-specific fixed effects and the population count of a given grid cell.

Keywords: electric vehicles, charging infrastructure, spatial analysis, spatial modelling

JEL Classification: O33, O38, Q48
AMS Classification: 62M30

1 Introduction

Decarbonization of the transport segment is a long term goal of a number of regional and global endeavours European Commission, 2020 in order to the limit impact of the climate change. Replacing internal combustion engine (ICE) based mobility by electric vehicle (EV) based one is a complex task, one which can be by necessity fully achieved only on a generational scale. It can be viewed as a set of several sub-problems, such as decarbonizing long distance cargo transport, personal mobility and replacing current ICE infrastructure by EV based one.

A network of EV charging stations is on one hand a key enabling factor of EV adoption, as EV users need to charge their vehicles regularly, and on the other hand cannot be understood merely as a drop in replacement of the current network of gas stations.

Key differences between ICE and EV infrastructure include the need of EV charging stations to be connected to the main grid, which has capacity constraints, the possibility of home charging of EV vehicles and the relative slow speed of charging compared to filling a tank with gasoline. Important consideration should be also given to the fact that EV adoption is currently seen as desirable, and as such receives varying levels of state support - this includes direct and indirect support for EV infrastructure building, including public and private charging facilities.

The role of public charging network in EV adoption is subject to scholarly debate He et al, 2022, Metais et al, 2022. Questions include the leading or lagging role of charging infrastructure with regards to overall EV adoption, and role of public versus private chargers in residential areas.

The usage of private home charging stations raises an important question with regards to accessibility of EV vehicle ownership for people living in apartment buildings, who generally do not have access to private parking spaces. This is a particularly important issue in urban areas, where the majority of the population lives in apartment buildings. The issue is further complicated by the fact that the majority of apartment buildings in urban areas are not equipped with charging infrastructure, and retrofitting them with charging infrastructure is a complex and costly task.

This raises the question of how to ensure that people living in apartment buildings have access to charging infrastructure, and how to ensure that the expected transition to EVs is equitable and inclusive.

[1] Department of Econometrics, Faculty of Informatics and Statistics, Prague University of Economics and Business, Prague, Czechia, jindra.lacko@vse.cz
[2] Department of Econometrics, Faculty of Informatics and Statistics, Prague University of Economics and Business, Prague, Czechia, formanek@vse.cz

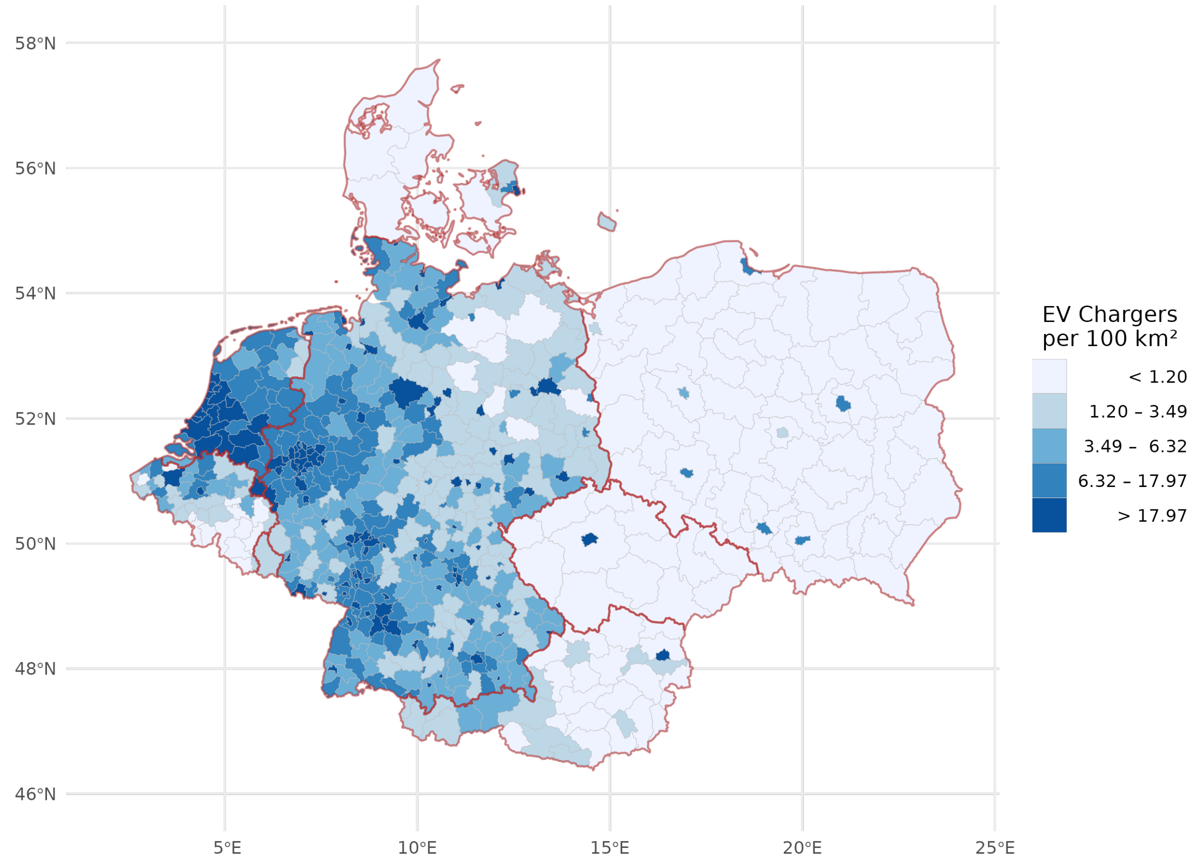


Figure 1 Public EV chargers per 100 km² at NUTS3-level

2 Data

Our analysis is based on three spatial datasets: we use a highly granular raster of level-8 hexagons based on the H3 architecture Sahr et al, 2003 with edge sizes of approximately 440 meters and we combine them with Open charge map (OCM) OpenChargeMap, 2023 information containing locations and additional information about public EV charging stations. The dataset is further enriched with distances to nearest shopping mall, obtained via OpenStreetMap OpenStreetMap contributors, 2023.

The eight countries used in our analysis are covered by 1,003,060 hexagons (level-8 H3 cells, continental Europe), and there are 35,353 public charging stations.

The EV charging stations are distributed over H3 cells highly unevenly. This is clear both from NUTS3 overview map 1 and from histogram of chargers per cell 2.

More formally we test the randomness of spatial distribution of EV chargers per H3 cell using in Moran's I test. The value of Moran's I statistic when using spatial contiguity neighborhood definition is 3.222751e-01, translating to z-score of 479.61, and 1.672074e-01, translating to z-score of 1072.3, when using distance based neighborhood definition with distance threshold 5 kilometres between H3 cell centroids. The hypothesis of spatial randomness of EV charger count per H3 cell can be therefore rejected, with demonstrated robustness with regards to choice of neighborhood matrix.

Data used for spatial analysis and model estimation are based on multiple open access sources. The Central European (CE) area covered in this analysis comprises of eight countries (NUTS0 units): Austria, Belgium, Czechia, Denmark, Germany, Luxembourg, the Netherlands, and Poland. Overseas dependencies (e.g. Antilles, Greenland, etc.) are excluded form the analysis. The choice of level-8 hexagons is based on the highest granularity of publicly available population density data, provided by the Humanitarian Data Exchange (HDX) Dicko, 2023. Observed EV charging stations, their exact locations, capacity / speed and additional data are retrieved from the Open charge map (OCM) OpenChargeMap, 2023, which is an open source and open access resource focused on EV infrastructure.

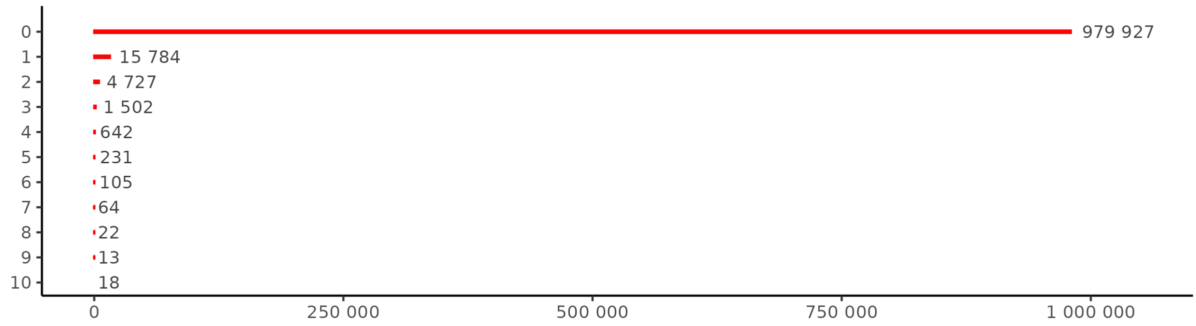


Figure 2 Histogram of EV chargers per H3 cell

The variables considered in our analysis are as follows:

- *evChargeCount*$_i$ is the main (target) variable of our analysis. It is an OCM-based OpenChargeMap, 2023 number of public EV chargers located within a given ith 440m hexagon as of October 2023.
- $\log(Population)_i$ represent the log-transformed number of inhabitants for a given 440m hexagon. Data come from the Kontur Population Dataset as of October 2023. Log transformation is used for econometric modelling. For the sake of log-transformation, zero values are transformed to 0.1 (10% of the lowest observed count).
- *Unempl_m*$_i$ depicts the unemployment level. Data are available at the NUTS2-level. We use Eurostat-based Eurostat, 2024 unemployment level for the age group 15-74 (code of the data set used is: lfst_r_lfu3rt).
- *ACPr_m*$_i$ is the ratio of economically active to total population, calculated at the NUTS2 level (Eurostat data sets lfst_r_lfp2act and demo_r_pjangroup were used).
- *ACPrYoung_m*$_i$ is the relative proportion of economically active young population (age groups 15-24 and 25-34 combined) to the total economically active population (15 years and older). Mean values for the period 2017 to 2022 are calculated.
- $\log(WAGr\_m)_i$ describes relative NUTS2-level wages based on Eurostat's dataset nama_10r_2lp10. For the calculation, compensations of employees per hour worked are given as percentages of EU27 average (i.e. 100 = average EU hourly compensation for a given year). Mean value is calculated across 2017 to 2022 observations.
- $\log(dToMall)_i$ is the distance from hexagon's center-point to the nearest shopping mall (as registered in the OSM database). Log-transformation is used for easier interpretation.
- *NUTS0*$_i$ is the root level of *Nomenclature des Unites Territoriales Statistiques*, i.e. the EU member state. This variable is a proxy for country level differences in social and economical customs, including effects of applicable EV regulatory regime. We are modelling the *NUTS0* variable as a factor, setting the reference level to the Czech Republic.

3 Methods

Basic intuition tells us that a good starting point for count data is the Poisson model, with equation

$$P(X = k) = \frac{e^{-\lambda} \times \lambda^k}{k!} \tag{1}$$

Where X is a discrete stochastic variable, λ is the single parameter of Poisson distribution (mean and variance) and k is a non-negative integer.

Such a model can be fitted with a relative ease. It however encounters serious issues in model validation phase, namely evaluating for expected and observed zero counts.

Instead we consider a two regime model, following Greene, 2012. In one regime, the outcome is always zero (sometimes described as a structural zero to differentiate from zeroes from the second regime). In the other, the usual Poisson process is at work, which can also produce a zero outcome - or some other.

Let z denote a binary indicator of regime 1 ($z = 0$) or regime 2 ($z = 1$) and let y denote the outcome of the Poisson process in regime 2. The vectors $\boldsymbol{x}$ and $\boldsymbol{w}$ stand for model covariates, which need not be the same for both regimes. Following Greene, 2012 we use $\boldsymbol{\gamma}$ for parameters of the logistic function $F()$.

$$g(y) = \begin{cases} Prob(z_i = 0|\boldsymbol{w}_i) = F(\boldsymbol{w}_i, \boldsymbol{\gamma}) & \text{Regime 1: } y \text{ will equal (structural) zero} \\ Prob(y_i = k|\boldsymbol{x}_i, z_i = 1) = \frac{e^{-\lambda_i} \times \lambda_i^k}{k!} & \text{Regime 2: } y \text{ will be a count outcome} \end{cases} \quad (2)$$

For our study we consider three models:

- [P] baseline model with Poisson component only and no zero inflation component
- [ZIP-1 Poisson model with zero inflation component determined by country and (log of) population variables
- [ZIP-2] Poisson model with zero inflation component considering all available variables

The aim of the more constrained [ZIP-1] model is to provide a parsimonious alternative to both baseline [P] model and the more comprehensive [ZIP-2] model, which uses the full set of 14 variables in both zero inflation and count components.

While the number of dependent variables is not of major concern in relation to the number of observations - the total number of H3 cells exceeds 1 million, and even the smallest country (Luxembourg) contains 3916 cells - we consider a smaller model beneficial with regards to interpretation.

The aim of the two variables (log of population and member state as a factor) in [ZIP-1] is to separate differences attributable to individual member states while respecting the specifics of truly uninhabited areas - such as the IJsselmeer inland sea in the Netherlands or the alpine regions of Austria - where no EV chargers are to be reasonably expected.

The models were fitted in R statistical programming language using modelling platform glmmTMB Brooks et al, 2017.

Coefficient estimates and key metrics of the three models are presented in table 1. The estimation follows the two components structure: firstly the zero inflation component, where the probability of structural zero is considered, and secondly the Poisson count component (which can also lead to zero observation).

In interpreting the coefficients in the two component model, care must be taken to interpret the zero inflation components in light of the first regime of equation 2, i.e. a positive coefficient means a variable is associated with a higher count of structural zeroes. Whereas coefficients from the count model need to be interpreted in second regime of equation 2, i.e. a positive coefficient means a variable is associated with a higher count of EV chargers.

4 Empirical Analysis

Based on our modelling we observe that both [ZIP-1] and [ZIP-2] zero inflated models significantly outperform the baseline [P] model. This conclusion holds consistently for AIC, BIC and expected number of zeroes criteria.

The difference between the two zero inflated models is not as conclusive, as the [ZIP-2] model performs slightly better on basis of information theory based criteria (AIC and BIC), while the [ZIP-1] model shows closer fit in criterion of expected count of zero values.

Given the relatively small difference between the [ZIP-1] and [ZIP-2] model results we will give our preference to the [ZIP-1] model, since it achieves a broadly comparable result with a smaller set of parameters.

While the actual values of model coefficients do not lend themselves to interpretation as easily as would be the case with a linear model, due to transformations inherent to GLM modelling technique, we can consider their signs and relative importance.

Given the prevalence of zeroes in our dataset the chief driver seems to be the zero inflation component, with comparatively limited contribution of the count (Poisson) process.

In the zero inflation component the most important variables are population (with a strong negative association with presence of a EV charger) and country specific fixed effects. It seems noteworthy that Denmark, a Scandinavian country and thus belonging to a region often associated with positive attitude to EVs, does not show statistically significant difference from the reference level (the Czech Republic).

For the other countries we can interpret the country specific coefficients as log of odds relative to the reference level (which is the Czech Republic). Thus a given H3 cell in Poland is $\exp(1.09665) = 2.994119$ times more likely to contain no EV charger than a comparable hexagon in the Czech Republic. On the other hand a German cell is $\exp(-1.44344)^{-1} = 4.23524$ times less likely to contain no EV charger compared to a Czech one with the same population.

	[P]			[ZIP-1]			[ZIP-2]		
Count component	Estimate	p-value[1]		Estimate	p-value[1]		Estimate	p-value[1]	
Intercept	-6.563672	< 2e-16	***	-3.331999	< 2e-16	***	-3.922101	< 2e-16	***
log(*population*)	0.849847	< 2e-16	***	0.264717	< 2e-16	***	0.308643	< 2e-16	***
Unempl_m	-0.087469	< 2e-16	***	-0.046453	8.62e-11	***	-0.001984	0.843283	
ACPr_m	3.159733	1.37e-15	***	3.957138	< 2e-16	***	2.349086	0.000240	***
ACPryoung_m	0.688849	0.046952	*	-0.847857	0.00136	**	4.400792	3.07e-14	***
log(*WAGr_m*)	-0.213216	0.001324	**	0.432679	< 2e-16	***	-0.044458	0.677168	
log(*distToMall*)	-0.438865	< 2e-16	***	-0.378921	< 2e-16	***	-0.235436	< 2e-16	***
NUTS0AT	1.207566	< 2e-16	***				0.134183	0.413218	
NUTS0BE	1.330639	< 2e-16	***				0.431791	0.020226	*
NUTS0DE	1.878724	< 2e-16	***				0.921093	5.62e-10	***
NUTS0DK	0.360063	0.000437	***				-1.238121	2.59e-08	***
NUTS0LU	1.516667	< 2e-16	***				-0.003560	0.991093	
NUTS0NL	2.145562	< 2e-16	***				0.553506	0.000228	***
NUTS0PL	-1.282094	< 2e-16	***				-0.262853	0.160572	
ZI component	Estimate	p-value		Estimate	p-value		Estimate	p-value	
Intercept				7.84152	< 2e-16	***	2.98254	1.17e-08	***
log(*population*)				-0.86523	< 2e-16	***	-0.73426	< 2e-16	***
Unempl_m							0.09915	1.37e-10	***
ACPr_m							-2.26041	0.014	*
ACPryoung_m							5.99652	1.22e-13	***
log(*WAGr_m*)							0.07698	0.607	
log(*distToMall*)							0.36586	< 2e-16	***
NUTS0AT				-0.89810	< 2e-16	***	-1.41011	1.21e-10	***
NUTS0BE				-0.57723	2.29e-11	***	-1.01558	4.85e-05	***
NUTS0DE				-1.44344	< 2e-16	***	-1.15040	8.23e-09	***
NUTS0DK				0.14474	0.131		-2.25843	4.37e-11	***
NUTS0LU				-0.81394	1.44e-06	***	-1.93891	2.14e-05	***
NUTS0NL				-2.05374	< 2e-16	***	-2.17279	< 2e-16	***
NUTS0PL				1.09665	< 2e-16	***	0.97060	8.07e-06	***
AIC	186 384			177 446			176 171		
BIC	186 550			177 635			176 502		
Log Likelihood	−93 178			−88 707			−88 057		
Num. observations	1 003 060			1 003 060			1 003 060		
Num. parameters	14			16			28		
$\sum_i \hat{f}_i(0)$[2]	975 319			979 912			980 000		

[1] With stars for significance according to convention: 0 = ***, 0.001 = **, 0.01= *, 0.05 = .

[2] The actual number of hexagons with zero EV chargers in dataset is 979 927

Table 1 Estimation of the models

We find somewhat surprising the comparatively low impact of the *WAGr_m* variable, representing regional level of wages / personal income, in the full [ZIP2] model. Considering the cost aspect of EV mobility this seems both somewhat counter-intuitive and against the trend of published research, e.g. Ščasný et al, 2018.

This conundrum could be possibly explained by:

- income differences at country level are already reflected in the *NUTS0* variable, with income differences within country insignificant in comparison
- income effect apparent in the decision to buy an EV vehicle itself, but not in the structure of charging network due to majority of EV owners utilizing private (and thus unreported in our dataset) chargers at home – and the charging network meeting rather the needs for fast charging resulting from transit traffic, which are less sensitive to local income levels
- income effect masked by other factors, such as state subsidies for EV infrastructure and capacity constraints of local power network

5 Conclusions

Our research demonstrates that the distribution of electric vehicle (EV) charging stations in Central European countries can be effectively modeled using a blend of physical and socioeconomic factors. Considering the critical role of EV adoption in the green transition of the transportation sector and the potential limitations of the charging network capacity impeding widespread EV usage, further investigation into this area is essential.

6 Acknowledgement

This work was supported by the grant No. IGA F4/24/2023, Faculty of Informatics and Statistics, Prague University of Economics and Business.

References

Brooks ME, Kristensen K, van Benthem KJ, et al (2017) glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R Journal 9(2):378–400. https://doi.org/10.32614/RJ-2017-066

Dicko A (2023) rhdx: A Client for the Humanitarian Data Exchange platform API

European Commission (2020) Sustainable and Smart Mobility Strategy – putting European transport on track for the future

Eurostat (2024) Eurostat database. online, https://ec.europa.eu/eurostat/data/database

Greene W (2012) Econometric analysis, international edition edn. Pearson series in economics, Pearson, Boston, Mass. Munich

He SY, Kuo YH, Sun KK (2022) The spatial planning of public electric vehicle charging infrastructure in a high-density city using a contextualised location-allocation model. Transportation Research Part A: Policy and Practice 160:21–44. https://doi.org/https://doi.org/10.1016/j.tra.2022.02.012

Metais MO, Jouini O, Perez Y, et al (2022) Too much or not enough? Planning electric vehicle charging infrastructure: A review of modeling options. Renewable and Sustainable Energy Reviews 153:111719

OpenChargeMap (2023) OpenChargeMap. https://openchargemap.org

OpenStreetMap contributors (2023) OpenStreetMap. https://www.openstreetmap.org

Sahr K, White D, Kimerling AJ (2003) Geodesic discrete global grid systems. Cartography and Geographic Information Science 30(2):121–134. https://doi.org/10.1559/152304003100011090

Ščasný M, Zvěřinová I, Czajkowski M (2018) Electric, plug-in hybrid, hybrid, or conventional? Polish consumers' preferences for electric vehicles. Energy Efficiency 11(8):2181–2201. https://doi.org/10.1007/s12053-018-9754-1, https://doi.org/10.1007/s12053-018-9754-1

Asymmetries in Savings-Investment Nexus: Global Perspectives

Lukáš Frýd[1]

Abstract. This study investigates the intricate dynamics between savings and investment, focusing on potential asymmetries and heterogeneous effects across different investment quantiles in both Large and Non-Large economies. Building upon empirical evidence, we explore whether the relationship between savings and investment varies according to economic conditions, particularly during the growth or decline phases. In contrast to the symmetric effects commonly assumed, our analysis unveils significant asymmetries in the impact of savings shocks on investment changes, notably observing stronger connections for adverse shocks compared to positive ones. Moreover, we identify variations in this relationship across different investment quantiles and distinguish between the responses of Large and Non-Large economies. Specifically, adverse savings shocks demonstrate stronger associations with investment changes during periods of economic downturn or decapitalization, while positive shocks exhibit heightened effects during growth phases. These findings underscore the importance of recognizing asymmetries and quantile-specific effects, providing a nuanced understanding of the interplay between savings and investment dynamics.

Keywords: quantile regression, common latent factors, saving-investment nexus, asymmetries

JEL Classification: C22, E22, F21
AMS Classification: 91B84

1 Introduction

The positive correlation between the ratio of investments (I) to GDP and savings (S) to GDP at the national level is a well-documented stylized fact. This relationship has been empirically validated numerous times through both time series and cross-sectional analyses. While empirical findings consistently reveal a statistically significant positive linear relationship between I/GDP and S/GDP, the consensus on the strength of this relationship has diminished. Contemporary economic theory tries to reconcile empirical realities with theoretical models, often employing structural models to explain this puzzle. In this paper, rather than adding to the extensive literature attempting to explain this puzzle, we focus on analyzing the heterogeneity in the investment-savings linkage. Our analysis is motivated by the observation that periods of economic growth (measured by GDP) are more frequent than periods of economic downturns. However, investment slumps tend to be more severe than investment increases. This suggests that the relationship between savings and investment shocks is contingent on the specific quantile of investment, highlighting a heterogeneous distributional effect. This leads us to question whether agents, on average, respond symmetrically and/or whether the response varies depending on unobserved conditions and react asymmetrically. Empirical analyses based on conditional mean estimation fail to capture these phenomena at the tails of the dependent variable distribution, which arise due to unexpected events, business cycles, and diverse structural and institutional factors. Additionally, we explore whether the perceived asymmetric investment response to positive or negative savings shocks in conditional mean analyses is misleading due to the asymmetric distribution of investment.

Our paper investigates the presence of a heterogeneous distributional effect and examines whether the strength of the investment-savings link differs in response to positive or negative savings shocks across various quantiles. Specifically, we analyze whether the decline or increase in investment within an economy correlates asymmetrically with positive or negative savings shocks and whether these shocks manifest differently in economies experiencing investment declines versus increases. We employ quantile regression to capture, analyze, and test for heterogeneity in the strength of the investment-savings association across different quantiles of the investment distribution. Furthermore, we define a model that relaxes the strong (and often unrealistic) assumption of cross-sectional independence among economics. Our model allows for both investment and savings in individual economies to be influenced by common latent factors.

[1] University of Economics in Prague, Department of Econometrics, Winston Churchill Square 4, 13067 Prague, Czech Republic, lukas.fryd@vse.cz

We purposely discuss the link between investments and savings and do not attempt a causal interpretation. We merely focus on the study of the short-run heterogeneous distributional impact of positive/negative savings shocks on the respective quantiles of log changes in investments, controlling for common global factors and heterogeneity of slope across economics. We run the analysis for different country clusters, specifically the impact of Large economies (the ten largest economies in the world) versus Non-Large economies.

2 Methodology

2.1 Model

The analysis of the relationship between investment and savings is based on a now well-known paper by Feldstein and Horioka 1980 (FH). In this paper we assume the following model:

$$\begin{aligned} \Delta inv_{it} &= \alpha_i + \beta_{1i}\Delta s_{it}^{+} + \beta_{2i}\Delta s_{it}^{-} + \boldsymbol{\omega}_i^T \boldsymbol{f}_t^l + \epsilon_{it} \\ \Delta s_{it}^{+} &= \alpha_i^{+} + \boldsymbol{\psi}_{1i}^T \boldsymbol{f}_{it}^o + \boldsymbol{\gamma}_{1i}^T \boldsymbol{f}_t^l + \epsilon_{1it} \\ \Delta s_{it}^{-} &= \alpha_i^{-} + \boldsymbol{\psi}_{2i}^T \boldsymbol{f}_{it}^o + \boldsymbol{\gamma}_{2i}^T \boldsymbol{f}_t^l + \epsilon_{2it} \end{aligned} \tag{1}$$

where subscripts $i = 1, ..., N$ and $t = 1, ..., T$ denote the country and the year, respectively, and

- $inv = \log I$, $s = \log S$,
- $\Delta s_{it}^{+} = \max(0, \Delta s_{it})$ a $\Delta s_{it}^{-} = \min(\Delta s_{it}, 0)$, see Shin et al 2014.
- $\boldsymbol{f}_t^l$ is an $r\times 1$ vector of latent common factors follow covariance stationary process and distributed independently of ϵ_{it}, ϵ_{1it},ϵ_{2it},
- $\boldsymbol{f}_{it}^o$ is an $r \times 1$ vector of observed country specific factors follow covariance stationary process and distributed independently of ϵ_{it}, ϵ_{1it},ϵ_{2it},
- $\boldsymbol{\gamma}_i$, $\boldsymbol{\omega}$, $\boldsymbol{\psi}$ is a vector of factor loadings,
- α_i, α_i^{+}, α_i^{-} are national fixed factors possibly correlated with regressors,
- α_i^{+}, α_i^{-} are national fixed factors,
- ϵ_{it}, ϵ_{1it},ϵ_{2it}, are exogenous random component that we assume to have zero mean and serially and cross-sectionally independent.

Kahneman and Tversky 2013, Shiller 2015 show that asymmetry is a fundamental aspect of the human condition. The asymmetric relationship between investment and savings is discussed in Raza et al 2018, where the results suggest that negative changes in saving have a stronger effect on investment than positive change. The possible transmission mechanism of Raza et al 2018 is explained by the close link between savings and investments with GDP, where GDP responds asymmetrically to monetary policy. Next, the relationship between I and S (to GDP) is strongly influenced by frictions and market failures Coeurdacier et al, 2020, 2015. However, the impact of frictions and market failures is different in the case of a growing economy or a shrinking economy; for example, Coeurdacier et al 2015 find that the strength of the correlation increases in turbulent periods and decreases in good times. Hence, the asymmetry in the tightness of the constraint leads to different responses of investments.

Moreover, we consider that both the change in investment and the change in savings are affected by the same global factors $\boldsymbol{f_t}$. As specific factors affecting the correlation between I and S, we consider the world interest rate, global productivity shock, technology shocks Baxter and Crucini, 1993; Glick and Rogoff, 1995, global GDP development, financial indicators and spatial dependence. It is clear from the equation 1 that the effect of unobserved factors would result in omitted variable bias. However, the equation 1 cannot be estimated due to unobserved factors $\boldsymbol{f}_t$. To include additional latent common, global as well as country-specific factors, we follow the approach by Harding and Lamarche, 2014; Harding et al, 2020 based on Pesaran 2006 and approximate these common factors $\boldsymbol{f}_t$ by cross-sectional averages of dependent and independent variables. The model is built in the following way:

$$\Delta inv_{it} = \alpha_i + \beta_{1i}\Delta s_{it}^{+} + \beta_{2i}\Delta s_{it}^{-} + \delta_i^T \bar{z}_t + \epsilon_{it} + O_p(N^{-1/2}), \tag{2}$$

where $\bar{z}_t$ is vector of cross-sectional averages of dependent (Δinv_t) and independent variables ($\Delta s_{it}^{+}, \Delta s_{it}^{-}$), $u_{it} = \epsilon_{it} + O_p(N^{-1/2})$ where $O_p(N^{-1/2})$ is associated with approximation of latent factors $\boldsymbol{f}_t$ with cross-section averages. The heterogeneous distributional effects of the shocks are estimated using conditional quantile regression:

$$Q_{\Delta inv_{it}}(\tau|\Delta s_{it}^{+}, \Delta s_{it}^{-}, \bar{z}_t) = \alpha_i(\tau) + \beta_{1i}(\tau)\Delta s_{it}^{+} + \beta_{2i}(\tau)\Delta s_{it}^{-} + \delta_i^T(\tau)\bar{z}_t. \tag{3}$$

where $Q(.|.)$ represents the function for conditional quantile τ and regressors, see Koenker and Bassett Jr 1978.

The resulting parameter estimates are obtained according to Harding and Lamarche 2014, Harding et al 2020 as

$$\hat{\boldsymbol{\pi}}(\tau) = \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\pi}}_i(\tau) \tag{4}$$

where $\boldsymbol{\pi}_i(\tau) = (\alpha_i(\tau), \beta_{1i}(\tau), \beta_{2i}(\tau))$. This approach is called quantile common correlated effect mean group estimator (QMG) Harding et al 2020. Variance covariance matrix for $\hat{\boldsymbol{\pi}}(\tau)$ allow for possible unit roots in factors and spatial forms of weak cross-sectional dependence (Pesaran 2006,Harding and Lamarche 2014). The estimation of the variance-covariance matrix $Var(\hat{\boldsymbol{\pi}}(\tau))$ is conducted using the bootstrap method according to Hagemann 2017.

3 Empirical part

3.1 Data

We employ unbalanced panel data from the World Bank database. The data was obtained for 217 economics from 1994 to 2019. The analyzed data does not include the pandemic Covid. We removed observations from the dataset for periods in which there was a war in the country. We further excluded economics with more than ten missing observations for the dependent variable to at least partially balance the panel. After these adjustments, we obtained a panel with a cross-sectional dimension of 132 economics. The investments are defined as Gross fixed capital formation (current US), and we define savings as Gross domestic savings (current US). Economics are further categorized into 'Large' and 'Non-Large' economics, with 'Large' referring to the top 10 economies worldwide based on their GDP.

3.2 Preliminary analysis

We tested for cross-sectional dependence using the CD test for weak cross-sectional dependence by Pesaran 2021. In the case of Δinv, the test statistic is equal to 110.44, and the p-value was 0.00. We reject H_0 of no cross-sectional dependence. In the case of Δs, the value of the test statistic is 84.55, and the p-value is equal to 0.00. Again, we reject H_0 of no cross-sectional dependence. The unit root test is performed for the variable s and inv for lags 0 to 3. We choose the number of lags according to the rule of thumbs Pesaran 2007. We perform the second generation unit root test CIPS by Pesaran 2007. We test unit root for augmented Dickey-Fuller regression with drift and further with deterministic trend. For both s and inv we can not reject the null hypothesis of a unit root in all cases. After the subsequent first difference, we reject H_0 for $\alpha = 0.01$.

3.3 Main analysis

In the empirical part we analyse the extended equation 1 into the following form:

$$Q_{\Delta inv_{it}}(\tau|\Delta s_{it}^{+}, \Delta s_{it}^{-}, \bar{z}_t) = \alpha_i(\tau) + \beta_1(\tau)\Delta s_{it}^{+} + \beta_2(\tau)\Delta s_{it}^{-} + \gamma_{i0}(\tau)D + \gamma_{i1}(\tau)D\Delta s_{it}^{+} + \gamma_{i2}(\tau)\Delta s_{it}^{-} + \boldsymbol{\delta}_i^T(\tau)\bar{z}_t. \tag{5}$$

where D represents a dummy variable for economics in the Large category. The estimation results of the equation 5 for Large vs Non-Large are in Table 1, where the columns labelled q represent the estimation for the respective quantile. Further, we test whether there is an asymmetric response of investment to a positive/negative shock in savings. The Wald-type test statistics by Koenker and Bassett 1982 is used for this purpose. The null hypothesis for Non-Large takes the form

$$H_0 : \beta_1(\tau_i) = \beta_2(\tau_i) \ vs.\ H_1 : \beta_1(\tau_i) \neq \beta_2(\tau_i).$$

For the group of Large economics, the null hypothesis is

$$H_0 : \beta_1(\tau_i) + \gamma_1(\tau_i) = \beta_2(\tau_i) + \gamma_2(\tau_i) \ vs.\ H_1 : \beta_1(\tau_i) + \gamma_1(\tau_i) \neq \beta_2(\tau_i) + \gamma_2(\tau_i).$$

In the case of Non-Large economics, the findings are summarized in Table 1. A noteworthy observation is that the positive shock in savings (Δs^{+}) lacks statistically significant impact only at the q_{10} quantile, while for the other quantiles, its effect holds statistical significance. In the context of the q_{25} quantile, the influence on the change in investment is rather modest, as indicated by the parameter estimate of 0.05. Therefore, in situations where the economy's capital is diminishing, a link between positive savings shock and investments appears to be negligible. As for Δs^{+}, a quantile-specific pattern emerges. The impact of Δs^{+} becomes more pronounced

as we move from quantile q_{25} to quantile q_{90}. The sole exception is the comparison between quantiles q_{10} and q_{25}, where the effect of Δs^{+} on investment change remains indistinguishable. For the other quantiles, the data support a heterogeneous distributional effect of positive saving shock. Significant parameters govern the influence of Δs^{-} at quantiles q_{10}, q_{25}, q_{50}, and q_{75}. For economics experiencing a decline in capital stock (q_{10} and q_{25}), the relationship between investment and a negative savings shock is particularly pronounced. In the case of a quantile q_{10}, the parameter estimate β even reaches 0.836. Consequently, a 1% negative savings shock is associated with an investment reduction of 0.8%. For quantile q_{25}, this reduction is 0.4%. In the context of a negative savings shock, the influence of Δs^{-} diminishes as quantiles increase. However, the effect remains marginal for quantile q_{75}, and at quantile q_{90}, it is no longer statistically significant. In Non-Large economics, investments respond asymmetrically to positive and negative shocks, as illustrated in Table 2. Only at the q_{50} quantile is it impossible to distinguish between the impact of positive and negative shocks.
However, when we account for the influence of a Large economy, we can observe that, unlike Non-Large economics, both positive and negative shocks in savings have statistically significant effects across all quantiles. See Table 1, specifically rows $\Delta s^{+}Large$ and $\Delta s^{-}Large$ for details. In contrast to Non-Large economics, a positive savings shock has a statistically significant impact even in the case of a capital decline within the economy. This is evident in Table 1, where you can find the relevant data in rows $\Delta s^{+} \times Large$ and columns q_{10}, q_{25}, and q_{50}. Therefore, even as capital decreases, a portion of domestic savings is utilized to finance a portion of the depreciated capital. Comparing Non-Large and Large economics, we observe that Large economics exhibit a stronger link between a positive shock in savings and investment for quantiles q_{10}, q_{25}, and q_{50}. However, when an economy is expanding its capital stock (q_{75} and q_{90}), the data does not support a statistically different response in investments to positive savings shocks between the Large and Non-Large groups.

Moving on to a negative savings shock in Large economics, we can observe a statistically significant difference in the connection compared to the rest of the world, as indicated in Table 1, excluding q_{25}, where the variance is higher due to a smaller number of observations, and thus the parameter is not statistically significant. However, we suspect this to be a Type II error. The fixed effect of a major economy estimated through parameter γ_0 in Equation 3 is marginal. At the same time, data for Large economics support the presence of an asymmetric investment response to positive/negative shocks only for quantiles q_{10} and q_{25}, as seen in Table 2 in column Large.

Table 1: Estimation results from quantile regression

Variable	$q_{0.1}$	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	$q_{0.9}$
Intercept	−0.0906***	−0.0404***	0.0085	0.0603***	0.126
	(0.0102)	(0.0085)	(0.0052)	(0.0053)	(0.0071)
Δs^{+}	0.0338	0.0545***	0.178***	0.353***	0.456***
	(0.0214)	(0.021)	(0.0377)	(0.0418)	(0.0424)
Δs^{-}	0.836***	0.424***	0.142***	0.037***	−0.0167
	(0.0975)	(0.1147)	(0.0422)	(0.0143)	(0.0322)
$Large$	0.0426***	−0.0027	−0.0274***	−0.04***	−0.0713***
	(0.0145)	(0.0128)	(0.0095)	(0.0097)	(0.0099)
$\Delta s^{+} \times Large$	0.258*	0.217*	0.178*	0.0109	0.0069
	(0.1372)	(0.1312)	(0.0995)	(0.1101)	(0.0967)
$\Delta s^{-} \times Large$	0.676**	0.465	0.488***	0.558***	0.568***
	(0.3169)	(0.3266)	(0.1862)	(0.1547)	(0.1061)
$\Delta s^{+}Large$	0.2917**	0.2717***	0.3555***	0.3635***	0.4633***
	(0.1335)	(0.1271)	(0.0953)	(0.1037)	(0.1021)
$\Delta s^{-}Large$	1.512***	0.8894***	0.6297***	0.5945***	0.5514***
	(0.3002)	(0.3038)	(0.1952)	(0.1407)	(0.0968)

Note: Asterisks denote significance at the 10% (*), 5% (**), and 1% (***) levels. Data in brackets represent standard errors.

3.4 Discussion

For Non-Large economies, various factors cause a positive savings shock to have either no or negligible connection to investment during periods of capital decline (quantiles q_{10} and q_{25}). Interpreting this causally, a positive shock to

Table 2: Hypothesis of the asymmetric effect

Quantile	Non-Large	Large
$q_{0.1}$	−7.51 ***	−3.13 ***
$q_{0.25}$	−3.29 ***	−1.62 *
$q_{0.5}$	0.69	−1.24
$q_{0.75}$	6.47 ***	−1.10
$q_{0.9}$	6.96 ***	−0.59

Note: The numerical values represent the values of the test statistic of the test. Asterisks denote significance at the 10% (*), 5% (**), and 1% (***) levels.

savings does not mitigate the fall in domestic investment; instead, the additional savings are likely invested outside the domestic economy. Conversely, a negative savings shock is strongly linked to investment during periods of capital stock reduction. This suggests that these economies heavily rely on domestic savings to maintain their capital levels. In economies that are expanding their capital stock (quantiles q_{75} and q_{90}), there is a statistically significant association between changes in investment (Δinv) and positive changes in savings (Δs^{+}), while the association with negative changes in savings (Δs^{-}) is statistically and economically insignificant. Causally, this implies that positive changes in savings are partially used for domestic investment, whereas declines in domestic savings are offset by foreign financing.

For large economies, the relationship between investment and both positive and negative savings shocks is stronger than that of non-large economies. During periods of investment slump (quantiles q_{10} and q_{25}), there is a statistically and economically significant association with positive savings shocks. Causally, this indicates that positive savings shocks can partially mitigate capital decline in Large economies, a pattern not observed in Non-Large economies. Moreover, negative savings shocks are also more strongly associated with investment declines in these quantiles, implying that negative shocks cause significant drops in investment. In Large economies, the investment response is generally homogeneous, except for a notable quantile effect at q_{10} and q_{25}, where data supports an asymmetric effect. This asymmetry is not observed in the remaining quantiles.

Another interesting aspect is the statistically insignificant difference in the linkage between positive savings shocks and investments for quantiles q_{75} and q_{90} in both Large and Non-Large economies. This suggests that both types of economies react similarly to positive savings shocks during periods of capital stock expansion. This similarity can be attributed to the significant impact of Large economies on the global system; when Large economies thrive, the global economy benefits. However, the response to negative savings shocks differs between Large and Non-Large economies. For Large economies, negative savings shocks significantly impact investment even at quantiles q_{75} and q_{90}. A negative savings shock in Large economies leads to a higher decline in investments during periods of capital stock growth. This can be explained by the inability of Large economies to attract sufficient foreign capital to offset domestic savings deficits. Conversely, Non-Large economies can replace deficits in domestic savings with foreign capital, allowing capital stock expansion to continue despite negative shocks (quantiles q_{75} and q_{90}).

Of course, all of this holds under ceteris paribus conditions. In a connected world, the relevance of the assumption that Large economies experience negative savings shocks while Non-Large economies remain unaffected is open to question. Further research could explore the spillover effects of savings.

Conclusion

This study highlights the nuanced and asymmetric dynamics between savings and investment across different quantiles, distinguishing between Large and Non-Large economies. Our findings reveal that adverse savings shocks substantially impact investment changes more than positive shocks, particularly during economic downturns. For Non-Large economies, positive savings shocks show negligible effects during capital declines, whereas adverse shocks significantly influence investment, underscoring these economies' reliance on domestic savings for maintaining capital levels. In Large economies, positive and negative savings shocks exhibit strong correlations with investment changes. Positive savings shocks can partially offset capital declines, a relationship not observed in Non-Large economics. Moreover, the effect of adverse savings shocks in Large economies is pronounced even during periods of capital growth.

Acknowledgements

The work was supported by the Internal Grant Agency of Prague University of Economics and Business under Grant F4/24/2023.

References

Baxter M, Crucini MJ (1993) Explaining saving–investment correlations. The American Economic Review pp 416–436

Coeurdacier N, Guibaud S, Jin K (2015) Credit constraints and growth in a global economy. American Economic Review 105(9):2838–2881

Coeurdacier N, Rey H, Winant P (2020) Financial integration and growth in a risky world. Journal of Monetary Economics 112:1–21

Feldstein M, Horioka C (1980) Domestic saving and international capital flows. The economic journal 90(358):314–329

Glick R, Rogoff K (1995) Global versus country-specific productivity shocks and the current account. Journal of Monetary economics 35(1):159–192

Hagemann A (2017) Cluster-robust bootstrap inference in quantile regression models. Journal of the American Statistical Association 112(517):446–456

Harding M, Lamarche C (2014) Estimating and testing a quantile regression model with interactive effects. Journal of Econometrics 178:101–113

Harding M, Lamarche C, Pesaran MH (2020) Common correlated effects estimation of heterogeneous dynamic panel quantile regression models. Journal of Applied Econometrics 35(3):294–314

Kahneman D, Tversky A (2013) Prospect theory: An analysis of decision under risk. In: Handbook of the fundamentals of financial decision making: Part I. World Scientific, p 99–127

Koenker R, Bassett G (1982) Tests of linear hypotheses and l" 1 estimation. Econometrica: Journal of the Econometric Society pp 1577–1583

Koenker R, Bassett Jr G (1978) Regression quantiles. Econometrica: journal of the Econometric Society pp 33–50

Pesaran MH (2006) Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica 74(4):967–1012

Pesaran MH (2007) A simple panel unit root test in the presence of cross-section dependence. Journal of applied econometrics 22(2):265–312

Pesaran MH (2021) General diagnostic tests for cross-sectional dependence in panels. Empirical economics 60(1):13–50

Raza H, Zoega G, Kinsella S (2018) Asymmetries exist in the feldstein–horioka relationship. The Journal of International Trade & Economic Development 27(6):667–684

Shiller RJ (2015) Irrational exuberance: Revised and expanded third edition. In: Irrational exuberance. Princeton university press

Shin Y, Yu B, Greenwood-Nimmo M (2014) Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ardl framework. Festschrift in honor of Peter Schmidt: Econometric methods and applications pp 281–314

An Empirical Efficiency Comparison of Downside Risk and Drawdown Risk in Dynamic Portfolio Optimization

Qian Gao[1]

Abstract. Through an in-depth exploration of risk, we have come to recognize the crucial role of downside risk in identifying potential extreme risks, and drawdown risk focuses on losses under adverse conditions. These two types of risks contribute to decision optimization, but their actual benefits within portfolios remain unclear. Consequently, this paper aims to utilize empirical data from diverse market environments, employing these two risk types to construct dynamic portfolios and conduct multi-dimensional comparisons. The chosen risk measures comprise maximum drawdown, conditional value at risk, entropic value at risk, conditional drawdown at risk, and entropic drawdown at risk. In our empirical study, we demonstrate the performance of downside risk and drawdown risk under different market conditions, comparing indicators of optimal holding periods for maximum returns across various time periods using Z-scores. The results indicate the effectiveness of risk-constructed portfolios, with drawdown risk exhibiting notable advantages. Additionally, the efficiency varies across different time periods and market conditions. This further elucidates that various markets may possess unique risk measurements tailored to their respective characteristics.

Keywords: portfolio optimization, downside risk, drawdown risk, time periods

JEL Classification: C58
AMS Classification: 91G10

1 Introduction

In finance, efficient risk management is paramount to secure optimal long-term performance for investment portfolios. Central to this endeavor are sophisticated metrics such as downside risk and drawdown risk measures. In the pursuit of dynamic portfolio optimization, there arises a compelling need to harmonize diverse risk metrics and the intricate dynamics of the market. This integration is pivotal for attaining a state of effectiveness and efficiency in portfolio management and optimization endeavors.

In order to expand risk measurement, various risk methods have been proposed. The Entropic Value-at-Risk (EVaR) concept provides an approximation to the Conditional Value-at-Risk (CVaR) method. It corresponds to the strict upper bound obtained from the Chernoff inequality of VaR and CVaR. (Ahmadi-Javid, 2012). Compared to the CVaR approach, portfolios using the EVaR method exhibit superior optimal, average, and worst-case return rates and Sharpe ratios. Portfolios based on EVaR may hold a competitive advantage (Ahmadi-Javid, 2019). Furthermore, an extended portfolio optimization framework called Entropic Drawdown at Risk (EDaR) has been proposed, which is based on the application of EVaR to drawdown distributions.

The conventional Maximum Drawdown (MDD) method relies on a single observation of maximum loss, potentially encompassing significant statistical errors. As an enhancement to portfolio return sample paths, a new single-parameter risk function called Conditional Drawdown-at-Risk (CDaR) has been proposed, expected to predict future risks better and provide a more stable and obtainable portfolio (Chekhlov, 2004). It has been shown that the portfolio's dynamic performance in out-of-sample testing improves significantly when risk constraints are tightened, leading to higher portfolio returns and lower losses and drawdowns when compared to CVaR (Krokhmal, 2005).

Nevertheless, an increasing number of MDD methods have also been applied to portfolios, with noticeable drawdowns observed since the market turbulence post-COVID-19. All models tend to exhibit fatter tails than assumed by index distributions (Geboers et al., 2023). Another method, Monte Carlo simulations for estimating MDD (MDaR), has also proved helpful in investment strategies and can rapidly respond to volatility-level changes. (Mendes and Lavrado, 2017)

[1] VŠB – Technical University of Ostrava, Faculty of Economics, Department of Finance, 17. listopadu 2172/15, 708 00 Ostrava-Poruba, Czechia, qian.gao@vsb.cz.

Developing more advanced algorithms based on these methods further validates their effectiveness in market applications. A dynamic portfolio trading system is divided into two phases: stock trend prediction and portfolio optimization. Pixel Graph Networks (PGN) have been particularly successful in predicting buying trade signals in the following days, input into the Mean-CDaR optimization model, notably (Alamdari et al., 2024). The new hybrid possibilistic and flexible robust model CDaR-HPFRM, considering CDaR metrics, provides a novel method for constructing effective portfolios in uncertain environments (Savaei et al., 2024). Portfolio optimization methods based on multi-objective planning algorithms have also demonstrated superior performance under different market conditions (Drenovak et al., 2022).

Based on this, the paper aims to evaluate downside risk (CVaR, EVaR) and drawdown risk (MDD, CDaR, EDaR) in dynamic portfolio optimization, comparing their performance (Sharpe and Calmar ratio) and advantages. The structure of the paper is as follows: Section 2 will introduce the relevant framework and computational methods, Section 3 will present the data and empirical results, and Section 4 will provide a summary and conclusion.

2 Methodology

This section will introduce a risk portfolio optimization framework based on classical portfolio optimization. We will explain the design process and how we empirically validated this framework. Then, we display the two types of risks we use: downside and drawdown risks, totaling five risks. Finally, we present the performance evaluation indicators after implementing portfolio optimization.

In our framework, we have included factors related to transaction costs and holding periods. Initially, the total weights of each portfolio may not add up to 1, as we adjust for incurred costs by deducting them from the weights. We calculate $w_{i,t}$ mainly based on the difference from the previous trading instance, understanding that each adjustment may not require a complete transition from 0 to 1, which may differ from the initial trading weight. For each $t \in \{1,2,3, \dots\}$ we solve the optimization problem as Equation (1),

$$\begin{aligned} min \quad & \rho_{(w_{i,t})}, \\ s.t. \quad & \sum_i w_{i,t} = 1 - \theta, \\ & w_{i,t} \geq 0, i = 1, \dots, N, \\ & \theta = \sum_i \left| w_{i,t} - w_{i,t-1} \right| \cdot \upsilon, \\ & w_{i,0} = 0, \end{aligned} \tag{1}$$

where $\rho_{(w_{i,t})}$ is the various types of risks, θ is the ratio of transaction costs. $w_{i,t}$ is the portfolio stock weight at the end of period t, i is the stock in the N asset pool (index), where $\sum \left| w_{i,t} - w_{i,t-1} \right|$ for the costs associated with weight changes, υ is the set transaction fee rate based on the average level of market traders. The specific construction pattern can be observed in Figure 1.

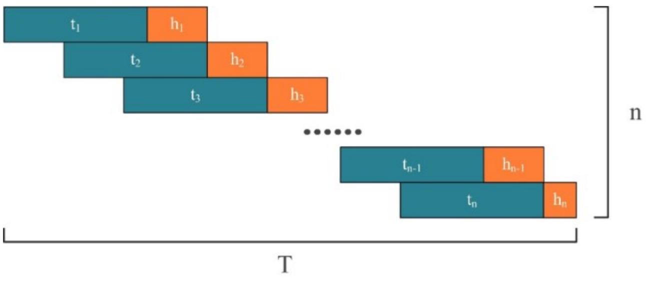


Figure 1 Designed portfolio optimization framework time diagram

Our framework depicted in Figure 1, t_n is the training period (in-sample), and h_n is the holding period (out-of-sample). It begins with training on data within t_n to determine optimal weights for that period. These weights are then held for h_n to evaluate the model's effectiveness. The training and holding periods are fixed, while variations in the holding period help assess risk and return levels. The holding period doesn't exceed the training period to avoid increased noise and instability.

Next, it is imperative to introduce the five risks. CVaR and EVaR belong to the downside risk type, while MDD, CDaR, and EDaR pertain to the drawdown risk type. All specified significance levels in our study are set to 5%.

CVaR extends this concept by incorporating the average value of losses that exceed the VaR threshold. The formulas are outlined as follows,

$$VaR_\alpha(X) = -inf_{u\in(0,t)}\{X_u \in R: F_X(X_u) > \alpha\},$$
$$CVaR_\alpha(X) = VaR_\alpha(X) + \frac{1}{\alpha t}\sum_{u=1}^{t} max(-X_u - VaR_\alpha(X), 0), \tag{2}$$

where $F_X(X_u)$ is the distribution function of the stock return X at u, u means a trading day in t, inf means the lower bound, $\frac{1}{\alpha t}$ is an adjustment factor, $max(-X_u - VaR_\alpha(X), 0)$ is truncated and positive for losses above VaR.

The EVaR corresponds to the tightest possible upper bound obtained by Chernoff's inequality for VaR and $CVaR$. Its formula is

$$EVaR_\alpha(X) = inf_{z>0}\left\{z \cdot ln\left(\frac{M_{X_u}(z^{-1})}{\alpha}\right)\right\}, \tag{3}$$

where inf still represents the search for the lower bound among all positive real values of z. $M_{X_u}(z^{-1})$ is the moment-generating function of the stock return X at u, and α is the significance level.

Next, we need to introduce drawdown risk. Firstly, we need to understand MDD, which measures the maximum loss of a portfolio or asset during a specific period. The formula is as follows,

$$DD(X_u) = max_{\tau\in(0,u)}(X_\tau - X_u)$$
$$MDD(X) = max_{u\in(0,t)}[DD(X_u)] \tag{4}$$

where $max_{\tau\in(0,u)}$ is the search for the maximum X_τ, $max_{u\in(0,t)}$ is the search for the maximum $DD(X_u)$.

CDaR represents the expected value of the worst-case drawdowns exceeding a specified threshold. The formula is as follows,

$$DaR_\alpha(X) = max_{u\in(0,t)}\{DD(X_u) \in R: F_X\big(DD(X_u)\big) < 1-\alpha\},$$
$$CDaR_\alpha(X) = DaR_\alpha(X) + \frac{1}{\alpha t}\sum_{u=0}^{t} max[DD(X_u) - DaR_\alpha(X), 0], \tag{5}$$

where α is the confidence level. X_u represents the value of the stock returns. $DD(X_u)$ denotes the drawdown of X at time u.

Similar to EVaR, the formula for EDaR is

$$EDaR_\alpha(X) = inf_{z>0}\left\{z \cdot ln\left(\frac{M_{DD(X_u)}(z^{-1})}{\alpha}\right)\right\}. \tag{6}$$

Next, we need to evaluate the portfolio, first is the standard deviation (Std), which is used to measure the dispersion of data points relative to the mean. The formula is as follows,

$$Std = \sqrt{\frac{1}{t}\sum_{u=1}^{t}(X_u - \bar{X}_t)^2}, \tag{7}$$

where Std means standard deviation, t is the number of days/observations, X_u is the stock return at a trading day u, and $\bar{X}_t$ means the average value of asset returns over the training period t. Sed is a statistical measure of the negative deviation below the mean in a data set. Its formula is as follows,

$$Sed = \sqrt{\frac{1}{t}\sum_{u=1}^{t} min\,(X_u - \bar{X}_t, 0)^2}, \tag{8}$$

where Sed is the semi-deviation, $min\,(X_u - \bar{X}_t, 0)^2$ means the difference between data point X_u and its average value $\bar{X}_t$, select the smallest value compared to 0. The Sharpe ratio adjusts returns for risk by quantifying the excess return of risk relative to the risk-free rate. The formulas are as follows,

$$Sharpe = \frac{R_p - R_f}{\sigma_p}, \tag{9}$$

where Sharpe is the Sharpe ratio, R_p is the portfolio return, R_f is the risk-free rate, σ_p is the portfolio return standard deviation, see Equation (7). The Calmar ratio compares the average annual return to the maximum drawdown. The formulas are as follows,

$$Calmar = \frac{R_p - R_f}{MDD}. \tag{10}$$

After calculating all evaluation indicators, our target is to facilitate their comparison on a standardized scale. To achieve this aim, we compute Z-scores. It serves as a standardized measure, quantifying the deviation of data points from the mean of the given dataset. This quantification is expressed in terms of standard deviations. The formula is expressed as

$$Z = \frac{\mu - \bar{\mu}}{\sigma_\mu}, \tag{11}$$

where Z means the Z-score, μ represents the computed indicators, $\bar{\mu}$ denotes the mean within the same indicator. Std and Sed use negative Z value.

3 Empirical Results

In our empirical comparison, we aimed to ensure the completeness, availability, and timeliness of the data while selecting representative and relevant (industry) market indexes. Therefore, we chose blue-chip stocks from two markets, the DAX30 and SSE50 index components, as last adjusted before 01/01/2018. The performance of these indexes serves as our benchmark, spanning from 01/01/2018-31/12/2023. Using the framework discussed above, we set the training period to 252 trading days, with a fixed holding period of 126. Considering the rise of online brokers, the trend toward zero transaction fees, and the one-year treasury yield at the beginning of 2018. Thus, we set the transaction fee rate at 0.5% and the risk-free rate at 2%. We have obtained attractive results by evaluating the portfolio performance in two different market environments. Figure 2 illustrates the holding cumulative returns of various risk portfolios. The dashed line indicates portfolio rebalancing at this time.

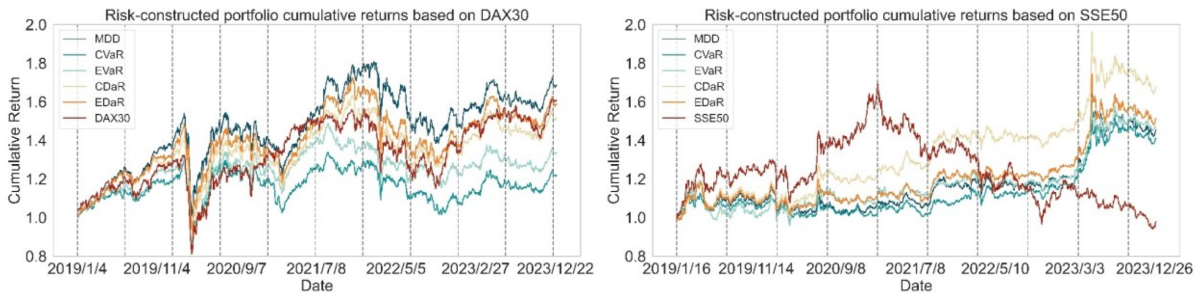


Figure 2 Risk-constructed portfolio cumulative returns

Firstly, the cumulative returns of DAX30-based portfolios show stable growth, mirroring the index. However, there was a significant decline and rebound during the three months of 2020. Initially, all five risk portfolios outperformed the benchmark index. From 2021 onwards, the index and drawdown risk returns surpassed those of the downside risk portfolios. Notably, drawdown risk portfolios yielded relatively higher returns, indicating their strong performance within the DAX market environment.

In contrast, the cumulative returns of SSE50-based portfolios depict a different trajectory. After an initial stable period, the portfolios faced a moderate decline in 2020, followed by a swift recovery. From early 2021 to 2023, the portfolios experienced a prolonged decline. Although the five risk portfolios' returns remained stable, they consistently lagged the benchmark index. From 2023, the risk portfolios' returns increased rapidly, deviating 0.4-0.7 from the benchmark index: drawdown risk portfolios, particularly CDaR, showed notably higher returns.

These findings highlight the pronounced differences in the performance of portfolios based on the DAX30 and SSE50 in varying market environments. We present the cumulative return distributions in Figure 3 to provide a more accurate assessment of their returns.

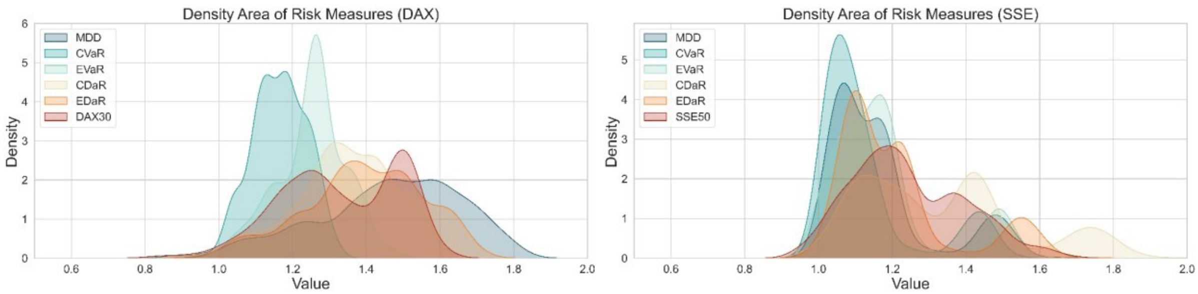


Figure 3 Cumulative return distribution of risk portfolio

The distribution of DAX and SSE in Figure 3 illustrates the risk characteristics of portfolios in different market environments. The DAX index and drawdown risks are more like the normal distribution, suggesting relatively uniform fluctuations. However, the distribution of downside risks is more concentrated, indicating that risk events in the DAX market are more likely to be confined within a specific range, potentially leading to higher volatility due to extreme events.

In contrast, the SSE market shows a distinct right-skewed distribution, indicating positive skewness. This implies more frequent low-value occurrences across all indicators than a normal distribution. Unlike the DAX, the SSE market has a small distribution extending to the right, suggesting the presence of high values or outliers under specific conditions. Similar to the DAX, downside risks in the SSE market are more concentrated, implying potential significant fluctuations.

We speculate that the relatively normal distribution in the DAX market indicates a more balanced risk profile, likely influenced by typical market fluctuations. In contrast, SSE may be more affected by event-driven factors leading to deviations from normalcy. Therefore, additional indicators for evaluating risk portfolios are needed. Figure 4 presents a comparison of Z-scores across multiple indicators.

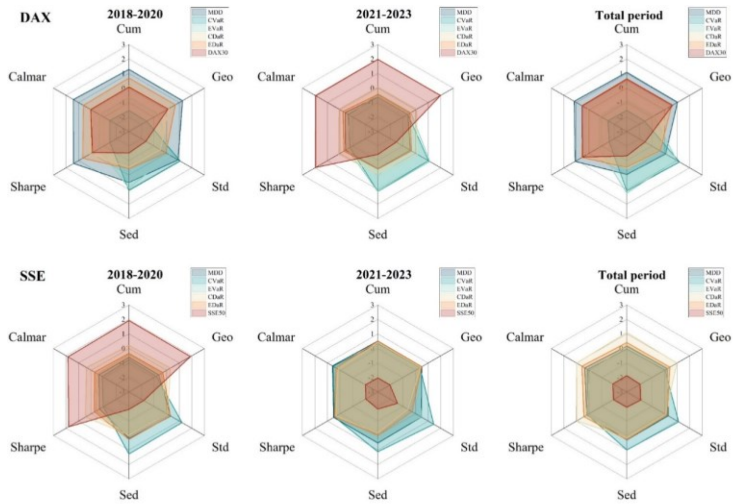


Figure 4 Radar charts for Z-score evaluation of risk portfolios across different time periods

Note: Cum is the cumulative return, Geo is the geometric mean cumulative return, Std is the standard deviation, Sed is the semi-deviation, Sharpe is the Sharpe ratio, Calmar is the Calmar ratio.

To demonstrate the strengths of various portfolios under different indicators, we utilized Z-scores to represent evaluation metrics in radar charts, as shown in Figure 4. A relative comparison within fixed intervals reveals that CVaR and EVaR perform exceptionally well in standard deviation and semi-deviation, showing significant advantages and notable stability across all periods in both market environments. In contrast, drawdown risk indicators form shapes closer to hexagons, suggesting a more balanced distribution but lacking stability.

Furthermore, there are similarities between specific periods in the two market environments. The risk portfolio Z-score indicators distribution from 2021 to 2023 for DAX and from 2018 to 2020 for SSE are highly similar. Referring to Figure 2, cumulative return variations during these periods are relatively stable, particularly with SSE failing to surpass the index value. This means low-risk preferences yield stable returns but may limit potential gains.

4 Conclusion

Comparison between downside and drawdown risks in dynamic portfolio optimization has led to several noteworthy conclusions. Firstly, both methods exhibit distinct characteristics and performance attributes. Downside risk serves as a comprehensive measure of overall risk exposure, proving to be an effective tool for swiftly assessing portfolio risk under diverse market conditions. Conversely, drawdown risk is particularly crucial during market turbulence or irrational fluctuations as it focuses on capturing adverse scenarios of portfolio risk.

Secondly, empirical findings suggest that the performance of downside and drawdown risks may vary across different market environments. Downside risk may tend to overly emphasize risk control in relatively stable market conditions, potentially resulting in loss of returns. On the other hand, drawdown risk demonstrates resilience and accuracy in reflecting portfolio risk exposure, especially during periods of market pressure.

Additionally, downside risk appears to exhibit a more concentrated distribution in portfolio optimization performance and consistently performs better in terms of standard deviation and semi-deviation. In comparison, drawdown risk demonstrates a broader distribution, with instances of completely nonlinear distribution observed in the SSE index environment. Its performance in evaluation metrics is relatively balanced but unstable.

Based on these results, it is evident that the choice between downside risk and drawdown risk methods depends on various factors, including investor preferences, risk tolerance, and market conditions. Therefore, a comprehensive approach should be adopted, considering the unique characteristics and performance attributes of both methods, to make informed decisions in dynamic portfolio optimization scenarios.

Acknowledgements

This work has been supported by the SP2024/047, an SGS research project of VSB-TUO.

References

Ahmadi-Javid A (2012). Entropic Value-at-Risk: A New Coherent Risk Measure. Journal of Optimization Theory and Applications 155(3): 1105–1123. https://doi.org/10.1007/s10957-011-9968-2

Ahmadi-Javid A, Fallah-Tafti M (2019). Portfolio optimization with entropic value-at-risk. European Journal of Operational Research 279(1): 22–241. https://doi.org/10.1016/j.ejor.2019.02.007

Alamdari M. K, Esfahanipour A, Dastkhan H (2024). A portfolio trading system using a novel pixel graph network for stock selection and a mean-CDaR optimization for portfolio rebalancing. Applied Soft Computing 152: 111213. https://doi.org/10.1016/j.asoc.2023.111213

Cajas D (2021). Entropic Portfolio Optimization: A Disciplined Convex Programming Framework (SSRN Scholarly Paper 3792520). https://doi.org/10.2139/ssrn.3792520

Chekhlov A, Uryasev S, Zabarankin M (2004). Portfolio Optimization With Drawdown Constraints. World Scientific Book Chapters: 209–228. https://doi.org/10.1142/9789812562586_0013

Drenovak M, Ranković V, Urošević B, Jelic R (2022). Mean-Maximum Drawdown Optimization of Buy-and-Hold Portfolios Using a Multi-objective Evolutionary Algorithm. Finance Research Letters 46: 102328. https://doi.org/10.1016/j.frl.2021.102328

Geboers H, Depaire B, Annaert J (2023). A review on drawdown risk measures and their implications for risk management. Journal of Economic Surveys 37(3): 865–889. https://doi.org/10.1111/joes.12520

Krokhmal P, Uryasev S, Zrazhevsky G (2005). 29. Numerical Comparison of Conditional Value-at-Risk and Conditional Drawdown-at-Risk Approaches: Application to Hedge Funds. In Applications of Stochastic Programming (pp. 609–631). Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898718799.ch29

Mendes B. V. de M, Lavrado R. C (2017). Implementing and testing the Maximum Drawdown at Risk. Finance Research Letters 22: 95–100. https://doi.org/10.1016/j.frl.2017.06.001

Savaei E. S, Alinezhad E, Eghtesadifard M (2024). Stock portfolio optimization for risk-averse investors: A novel hybrid possibilistic and flexible robust approach. Expert Systems with Applications 250: 123754. https://doi.org/10.1016/j.eswa.2024.123754

Evaluating the Energy and Emission Efficiency of the 27 Countries of the European Union during 2011-2022 Period Using DEA

Jana Hanclova[1], Lucie Chytilova[2], and Dominika Bordacsova[3]

Abstract. The European energy market is characterized by a diverse mix of energy sources, including fossil fuels, nuclear power, and an increasing share of renewable energy sources such as wind, solar, and biomass. The market operates within the framework of the European Union (EU), with regulations aimed at promoting competition, ensuring security of supply, and reducing greenhouse gas emissions.

The main goal of this paper is to evaluate the efficiency of energy and emissions in 27 EU countries using data envelopment analysis (DEA) during the 2011-2022 period. Attention is paid to inputs of energy consumption in the industry or transport sector, the share of renewable sources, and undesirable output of air pollutants and greenhouse gases. By incorporating various input and output variables, the study assesses how effectively these countries utilize their energy resources while minimizing environmental impacts. We use a direct distance function, non-proportional changes, and a non-orientated DEA model with variable returns to scale. The results document that the critical output is precisely the share of renewable resources and that there is no significant improvement in the examined period. The second critical undesirable output is the production of emissions. It was also found that the problem for most inefficient countries was the high energy consumption in the transport sector compared to the industry sector. The results provide insight into best practices and highlight areas for improvement, contributing to policy recommendations to improve energy sustainability and reduce emissions across the EU.

Keywords: data envelopment analysis, energy efficiency, emission efficiency, European Union, renewable energy

JEL Classification: C61, D24, M11
AMS Classification: 90C08, 90C90

1 Introduction

Global climate change represents one of the most significant challenges facing humanity today. It is primarily driven by the accumulation of greenhouse gases (GHGs) in the atmosphere, largely resulting from human activities, such as the burning of fossil fuels for energy. Addressing this challenge requires urgent action on multiple fronts, with **energy efficiency and emission reduction** playing crucial roles. *Energy efficiency* refers to using less energy to perform the same task or produce the same result. It is a fundamental strategy for reducing energy consumption and mitigating climate change. Key benefits of energy efficiency include reduction in greenhouse gas emissions, use of energy-efficient technologies, and practices that can lead to significant cost savings. Emission reduction involves decreasing the amount of pollutants, particularly greenhouse gases, released into the atmosphere. Strategies for emission reduction include **transitioning to renewable energy sources**, implementing carbon capture and storage technologies, and promoting circular economy principles to reduce waste and enhance resource efficiency. Energy efficiency and emission reduction are interlinked and mutually strengthening strategies. Efficient energy use directly contributes to lower emissions by reducing the demand for energy from fossil fuels. Furthermore, investments in renewable energy and low-carbon technologies improve energy

[1] VSB – Technical University of Ostrava, Faculty of Economics, Department of Systems Engineering and Informatics, 17. listopadu 2172/15, 708 00 Ostrava-Poruba, Czech Republic, jana.hanclova@vsb.cz
[2] VSB – Technical University of Ostrava, Faculty of Economics, Department of Systems Engineering and Informatics, 17. listopadu 2172/15, 708 00 Ostrava-Poruba, Czech Republic, lucie.chytilova@vsb.cz
[3] VSB – Technical University of Ostrava, Faculty of Economics, 17. listopadu 2172/15, 708 00 Ostrava-Poruba, Czech Republic, dominika.bordacsova.st@vsb.cz .

efficiency while simultaneously cutting emissions. Figure 1 presents the development of the percentage of total energy consumption from renewable sources for 27 countries in the European Union (Eurostat, 2024b). It has been chosen for the assessment of the progress towards the objectives and targets of the EU Sustainable Development Strategy. This data set covers the indicator for monitoring progress towards renewable energy targets of the Europe 2020 strategy implemented by Directive 2009/28/EC on the promotion of the use of energy from renewable sources (RED I) and the Fit for 55 strategies under the Green Deal implemented by Directive (EU) 2018/2001 on the promotion of the use of energy from renewable sources (RED II). The overall *share of renewable energy (RES)* in the EU's total energy consumption has been *steadily increasing* since 2004 from almost 10 % to the level of 23 % in 2022. The Renewable Energy Directive (RED I), adopted in 2009, set a binding target for the EU to achieve a 20% share of energy from renewable sources by 2020, which was successfully achieved (22%). The EU aims to increase this to 32% by 2030, under the revised Renewable Energy Directive (RED II). The share of RES in transport is lower compared to other sectors but has grown due to the increased use of biofuels and electric vehicles (1,6 % in 2004 to 9,6 % in 2022). Renewable energy sources, such as wind, solar, hydro, and biomass, contribute significantly to the electricity sector (from 16 % in 2011 to 41,2 % in 2022). In some EU countries, renewables account for a large proportion of electricity generation. For example, Denmark and Germany have a large share of wind and solar energy, respectively.

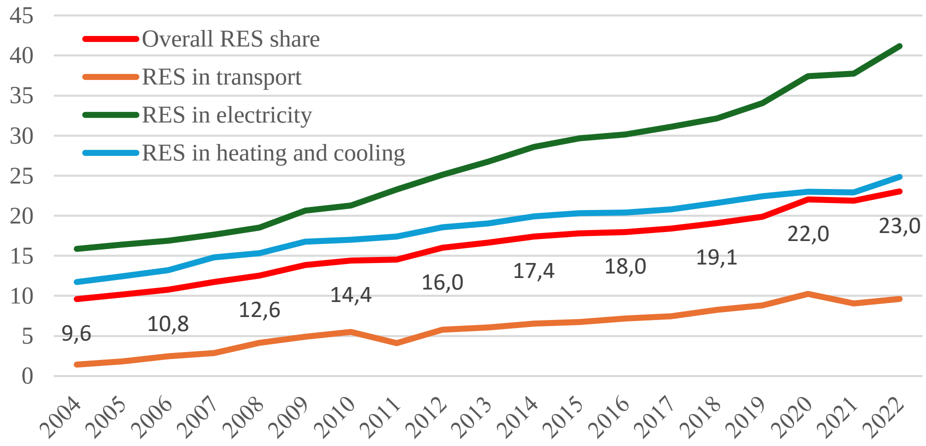


Figure 1 Share of energy from renewable sources (RES, percentage)

Source: Data extracted from the Eurostat database [nrg_ind_ren] (Eurostat, 2024b)

Regarding trends in the development of *emission reduction* in the 27 EU countries, there is a noticeable positive reduction of greenhouse gases from 3 478 936 thousand tonnes in 2011 to 2 897 292 in 2022, i.e., by about 17%. The GHGs in mining and quarrying, which fell by almost 33%, are the biggest contributors to this.

In the context of global climate change, energy efficiency and emission reduction are vital strategies that offer substantial environmental, economic, and social benefits. Policymakers, businesses, and individuals must work collaboratively to implement effective measures that improve energy efficiency and reduce emissions, contributing to global climate goals and the overall well-being of society.

The main goal of this paper is to evaluate the efficiency of energy and emissions in 27 EU countries using data envelopment analysis (DEA) during the 2011-2022 period.

The article consists of four sections. The following Section 2 is devoted to the literature review. Section 3 presents the main features of the methodology, and describes data and variables. Section 4 briefly discusses the results obtained, while the main conclusion and directions for further research are presented in Section 5.

2 Review of the literature

Although significant strides have been made in *assessing energy and emission efficiency across various regions and sectors, there is* a notable gap in the comprehensive analysis that specifically targets **EU countries using recent data**. Many researchers in energy and environmental economics use data envelopment analysis (DEA) to evaluate the efficiency of a production system. Data Envelopment Analysis is a powerful non-parametric method used to evaluate the efficiency of decision-making units (DMUs) such as countries, companies, or other entities. Compare the inputs and outputs of each DMU to identify those operating on the efficiency frontier, thus determining which units are performing best and which need improvement.

Three basic review publications systematically summarise previous research on the use of DEA in the energy and environment. Zhou et al. (2008) classified 100 such publications for the period 1983 – 2006 and found that benchmarking of electricity utility accounts is the basis for the largest number of studies, followed by the areas of modeling environmental performance and energy efficiency study. Meng et al. conducted a comprehensive survey of 46 empirical studies published in 2006-2015 in China's regional energy and carbon emission efficiency. The authors paid attention to the selection of inputs and outputs for the production system, various types of DEA models, and the incorporation of dynamic changes. The third paper by Sueyoshi et al. (2017) summarised 693 research works on DEA applied to energy and environmental efficiency from the 1980s to the 2010s. The results confirmed a dramatic increase in the number of such professional articles, particularly after the 2000s. The results of these studies document the benefits of using DEA in guiding large policy and business issues such as global warming and climate change in the world.

Energy efficiency, or energy saving and conservation, is a measure of managing and restraining a growth in energy consumption (Meng, et al., 2016). One of the most popular research areas for measuring energy efficiency is comparison among industries/sectors, regions, and countries. Another prominent application area is agriculture production. Zhou et al. (2008) assesses energy efficiency in 21 OECD countries. When considering energy consumption as input and GDP as output, the study identified the most efficient countries and provided information on best practices. There are not many research articles devoted to energy and emission efficiency in the European Union. Gökgöz et al. (2023) deal with the analysis of environmental, energy, and economic efficiency for the energy market in the European Union. According to environmental efficiency, the authors found that the Nordic countries are primarily more efficient, while, in terms of economic innovation efficiency, Germany is ranked as the top-performing country. Cámara-Aceituno et al. (2024) examined only **emission efficiency** (CO2, CH4) for 27 countries in the EU during the period 2012-2020. The results showed that only 12 countries had emission efficiency and that relatively high and rich countries are closer to achieving high environmental efficiency than less developed countries. Borozan and Starcevic (2019) explained the technical efficiency of the DEA and its determinants in the European energy industry for the period 2005-2016. They noted that the differences in efficiency may be explained by company-specific characteristics and environmental factors. In particular, energy efficiency may be improved by ensuring employee education with a focus on the internalization of business and open to competition.

Past research using DEA has contributed significantly to our understanding of energy efficiency and emission efficiency. These studies have provided valuable benchmarks, identified key areas for improvement, and informed policy and strategic decisions. As policymakers continue to seek sustainable development, the insights from DEA studies remain crucial in guiding effective and impactful strategies. Although significant strides have been made in assessing energy and emission efficiency through DEA in various regions and sectors, there is a notable gap in the comprehensive analysis that specifically targets EU countries using recent data.

3 Materials and methods

Data Envelopment Analysis (DEA) is a powerful nonparametric method used to evaluate the efficiency of decision-making units (DMUs) such as countries, regions, companies, or other entities. For this study, we focus on a specific DEA model suitable for evaluating energy and emission efficiency across EU countries, which is the **non-radial directional distance function (DDF) model**. This model is particularly relevant, as it allows the simultaneous consideration of multiple inputs and outputs, including *undesirable outputs* like CO_2 emissions. Unlike radial models, which assume a proportional reduction of all inputs or expansion of outputs, the nonradial DDF model allows for different rates of adjustment for each input and output. This is crucial for accurately assessing efficiency in contexts where certain inputs or outputs cannot be easily scaled. The DDF extends traditional DEA by incorporating a direction vector that guides *the reduction of inputs* and *undesirable outputs* or *the expansion of desirable outputs*. This flexibility makes it particularly suitable for environmental efficiency analysis, where reducing emissions (undesirable outputs) and maintaining or increasing economic output (desirable outputs) are objectives.

Zhou et al. (2012) presented **a nonradial DDF approach** to modeling energy and emission efficiency in electricity generation for 129 countries in the year 2005. It was found that the OECD countries had better carbon emission efficiency and integrated energy emissions efficiency than the non-OECD countries, while the difference in energy efficiency was not significant. Take Wang et al. (2016) applied a non-radial DEA model with a focus on both energy consumption and emissions for 30 provinces in China from 1996 to 2012. The results showed that carbon emission efficiency among the 30 areas varied significantly closely related to regional economic development. Furthermore, the radial DDF model might *overestimate the efficiency of the emission.* Wang and Wei (2014) confirmed that economically well-developed cities demonstrated higher efficiency.

To study energy and emission efficiency, we will formulate **a nonradial DEA model with DDF** and the possibility that inputs, desirable outputs, and undesirable outputs can be adjusted nonproportionally. For each decision-making unit (DMU_j) $j = 1, 2, \dots, N$, we will consider M inputs ($i = 1, 2, \dots, M$), R desirable outputs R (r = 1, 2,..., R) and undesirable outputs ($s = 1, 2, \dots, S$). Now we can formulate an extended and nonradial DEA model that is nonoriented with DDF according to (Zhou et al., 2012 and Wang et al., 2014):

$$z = \boldsymbol{D_T} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{b}, \boldsymbol{g}^x, \boldsymbol{g}^y, \boldsymbol{g}^b) = max\left\{\boldsymbol{w}^{x'} \cdot \boldsymbol{\beta}^x + \boldsymbol{w}^{y'} \cdot \boldsymbol{\beta}^y + \boldsymbol{w}^{b'} \cdot \boldsymbol{\beta}^b\right\} = \beta$$

$$\begin{aligned} &\sum_{j=1}^{N} \lambda_j \cdot x_{ij} \le (1 - \beta_i^x) \cdot x_{io} && i = 1, 2, \dots, M \\ &\sum_{j=1}^{N} \lambda_j \cdot y_{rj} \ge (1 + \beta_r^y) \cdot y_{ro} && r = 1, 2, \dots, R \\ &\sum_{j=1}^{N} \lambda_j \cdot b_{sj} = (1 - \beta_s^b) \cdot b_{so} && s = 1, 2, \dots, S \\ &VRS : \sum_{j=1}^{N} \lambda_j = 1, \quad \lambda_j \ge 0 && j = 1, 2, \dots, N \\ &\beta_i^x, \beta_r^y, \beta_s^b \ge 0 \quad \forall i, r, s. \end{aligned} \tag{1}$$

where the inputs, desirable outputs, and undesirable outputs for DMU_j are respectively denoted by $\boldsymbol{x}' = (x_{1j}, x_{2j}, \dots, x_{Mj}), \boldsymbol{y}' = (y_{1j}, y_{2j}, \dots, y_{Rj}), \boldsymbol{b}' = (b_{1j}, b_{2j}, \dots, b_{Sj})$. λ_j intensity variables for connecting inputs and outputs by a convex combination. The nonzero vector $\boldsymbol{g}' = (\boldsymbol{g}^x, \boldsymbol{g}^y, \boldsymbol{g}^b) = (-\boldsymbol{x}_o, \boldsymbol{y}_o, -\boldsymbol{b}_o)$ is the *directional vector for inputs, desirable outputs, and undesirable outputs.* The normalized weight vector $\boldsymbol{w} = (\boldsymbol{w}^x, \boldsymbol{w}^y, \boldsymbol{w}^b)'$ is determined by the number of inputs and desirable and undesirable outputs. It is assumed that the weight of all inputs, desirable and undesirable outputs is successively 1/3. The vector $\boldsymbol{\beta}' = (\boldsymbol{\beta}^x, \boldsymbol{\beta}^y, \boldsymbol{\beta}^b)$ denotes the scaling factors vector for inputs, desired outputs, and undesirable outputs, i.e. β_i^x expresses the intensity of reduction of *i-th* input, β_r^y increase of *r-th* desirable output, β_s^b and reduction of *s-th* undesirable output for each DMU_j. The zero $\beta_i^x, \beta_r^y, \beta_s^b$ expresses **the effective DMU_j** for input x_i, desirable output yr, and undesirable output b_s in turn. The mentioned DEA model (1) is established under the condition of *variable return to scale (VRS).*

In this paper, the research sample includes 27 DMUs, which are countries of the European Union (EU) presented and is based on data drawn from 2011 to 2022.

indicators	**explanatory variables**	**meaning/ data code in the Eurostat database**	**unit**
inputs	empl	employment. [nama_10_a10_e]	thousand hours worked
	fast	nonfinancial assets – stocks [nama_10_nfa_st]	million euros
	fcind	final energy consumption in an industry sector [ten00124]	thousand tonnes of oil equivalent
	fctra	final energy consumption in the transport sector [ten00124]	thousand tonnes of oil equivalent
desirable outputs.	gdp	gross domestic product [nama_10_gdp]	million euros
	apgg	air pollutants and greenhouse gases; [env_ac_ainah_r2]	thousand tonnes;
undesirable outputs	srene	share of renewable energy in gross final energy consumption [sdg_07-40]	percentage.

Table 1 The indicators descriptions (Eurostat (2024a, 2024b))

variable	empl	fast	fcind	fctra	gdp	apgg	srene
Minimum	348 792	23 516	40	181	6 848	1 764	1,9
maximum	62 168 126	10 263 318	57 410	52 247	3 017 499	797 609	66,0
means	12 188 234	1 493 743	8 779	10 171	439 279	116 928	21,2
std. dev.	12 361 829	2 392 920	11 924	13 957	685 939	158 205	11,8

Table 2 Descriptive statistics of indicators for 27 countries in the EU from 2011 to 2022

4 Results

By optimizing the non-radial and nonorientated DEA model (1) with DDF as well as including the undesirable output, the nonproportional measure, the total beta coefficient β , and other partial beta coefficients were obtained for each of the 27 EU countries in the individual years from 2011 to 2022. It used GAMS software 43.3.1.

The number of effective countries is shown in Table 3. The results document that the number of effective units was 11 to 13, that is, 40,7 to 44,4 %. This result corresponds to the research study by Cámara-Aceituno et al. (2024), who identified 12 countries from the EU during the period 2012-2020. Among the effective EU countries in the entire monitored period were Germany (DE), Denmark (DK), Estonia (EE), France (FR), Ireland (IE), Luxembourg (LU), Latvia (LV), Malta (MT), Netherlands (NL), Poland (PL) and Sweden (SE).

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
number	12	12	13	11	12	11	11	12	13	12	12	12
%	44,4	44,4	48,1	40,7	44,4	40,7	40,7	44,4	48,1	44,4	44,4	44,4

Table 3 The effective countries using the DEA model (1) from 2011 to 2022

The total factor **efficiency for each country** in individual years from 2011 to 2022 presents the coefficient β (*beta*) and individual efficiency scores for each input and output factor $\beta^{empl}, \beta^{fast}, \beta^{fcind}, \beta^{fctra}, \beta^{gdp}, \beta^{srene}, \beta^{apgg}$ which are marked in the graphs as $beta, b\_empl, b\_fast, b\_fcind, b\_fctra, b\_gdp, b\_srene, b\_apgg$.
The evaluation of the **development trend of the average efficiency** for 27 EU countries is shown in Figure 2. *The average total efficiency beta* was 0,227 in 2011 and slightly improved to 0,203 in 2013, followed by an oscillation to 0,225 in 2022. The cause of inefficiency is mainly the share of the renewable energy output factor *in the final energy consumption* (*srene*). The corresponding efficiency β^{srene} improved from 0,594 in 2011 to 0,489 in 2015 and then deteriorated to 0,524 in 2022. The second critical factor is *emissions* (*appgg*). The average β^{apgg} fluctuated slowly between 0,242 and 0,282 in 2020. Other critical factors are *employment (empl),* followed by *energy consumption in the transport sector (fctra) and then also in the industry sector (fcind).* The average β^{fctra} is constantly slowly increasing from 0,136 in 2011 to 0,210 in 2022, the average β^{fcind} occurs in the interval of 0,098 to 0,152 in 2021, which is lower, and less inefficient compared to the transport sector. Therefore, it is a legitimate question to reassess the energy mix used in transport but also in industry. The input factor capital (*fast*) and the output factor *gdp* have a slight influence on the average total efficiency.
The results of the **total inefficiency analysis** for the period 2011-2022 are presented in Figure 3. The results obtained by optimizing the DEA model (1) and confirmed by the clustering analysis document that Czechia (median $\beta = 0,696$), Slovakia (0,666), Hungary (0,657), Bulgaria (0,599), and Belgium (0,554) belong to the lowest efficiency countries. The cluster profile for these countries points to a key problem with *the share of renewable energy in final energy consumption* (mean β^{screne} = 2,199) i.e., a significant increase in the level of that output factor is necessary), another critical factor is the necessary *reduction of air pollutants and greenhouse gases* (mean $\beta^{apgg} = 0,548$). The third critical factor to improve total efficiency is *employment* (mean $\beta^{empl} = 0,377$) and a reduction of thousand hours worked is necessary. Based on the results in this group, a *reduction in energy consumption in the transport sector* is also desirable.

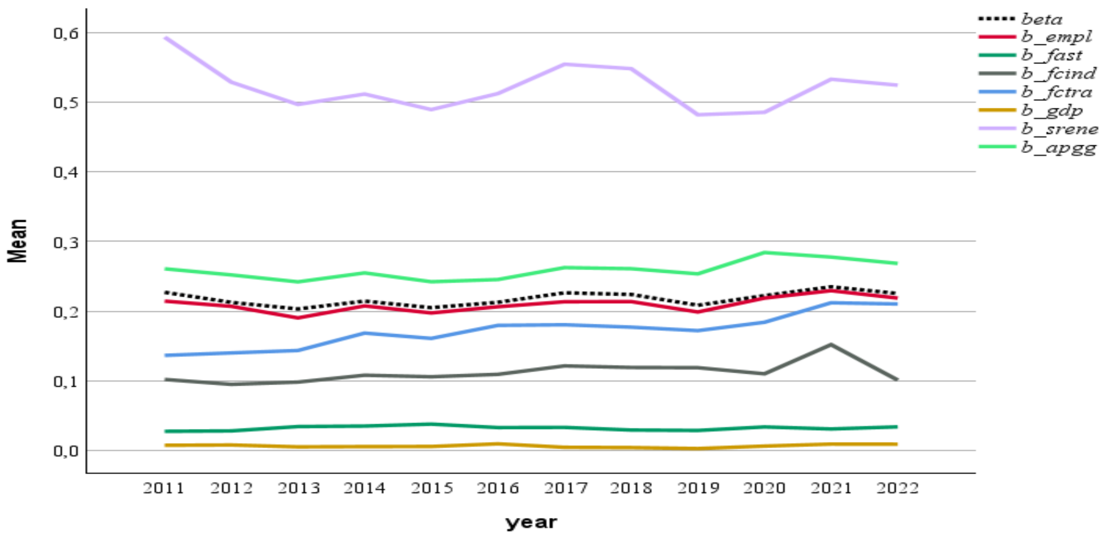


Figure 2 Trends in the development of average efficiency

In the next part of the paper, we will focus on the evaluation **of energy and emission efficiency in 27 EU countries**. Figure 2 shows that the average $\beta^{fcind}, \beta^{fctra}$ measures the *efficiency of energy consumption in the industry and transport sectors*, while the most significant cause of inefficiency is the high energy consumption in transport and even has a worsening trend of the average level from 0,136 to 0,210 in the years 2011-2022, especially after 2019. The trend of improvement in average efficiency in the industry starts at 0,102 in 2011 and returns to almost the same level in 2022, with a higher peak in 2021. In terms of country classification, it is confirmed that *Finland* is much higher β^{fcind} (mean 0,515) than β^{fctra} (0,119). Moreover, this phenomenon is different compared to most inefficient countries. The second country, which has significant problems with energy consumption compared to other countries, both in transport (0,557) and industry (0,478) and especially with a strong increase in 2021 is Bulgaria. The obtained results also agree with other literature (Wang et al. (2014), Zhou et al. (2012), Cámara-Aceituno et al. (2024)) that economically well-developed countries have higher energy and emission efficiency. If we evaluate *emission efficiency* using β^{apgg} the worst efficiency compared to best practice, Romania (average 0,778), followed by Bulgaria (0,771) and Czechia (0,758) show the results. |These countries should fundamentally reduce air pollutants and greenhouse gases. Only Romania sees a slow decline from 0,798 in 2011 to 0,758 in 2022 but the level is still high.

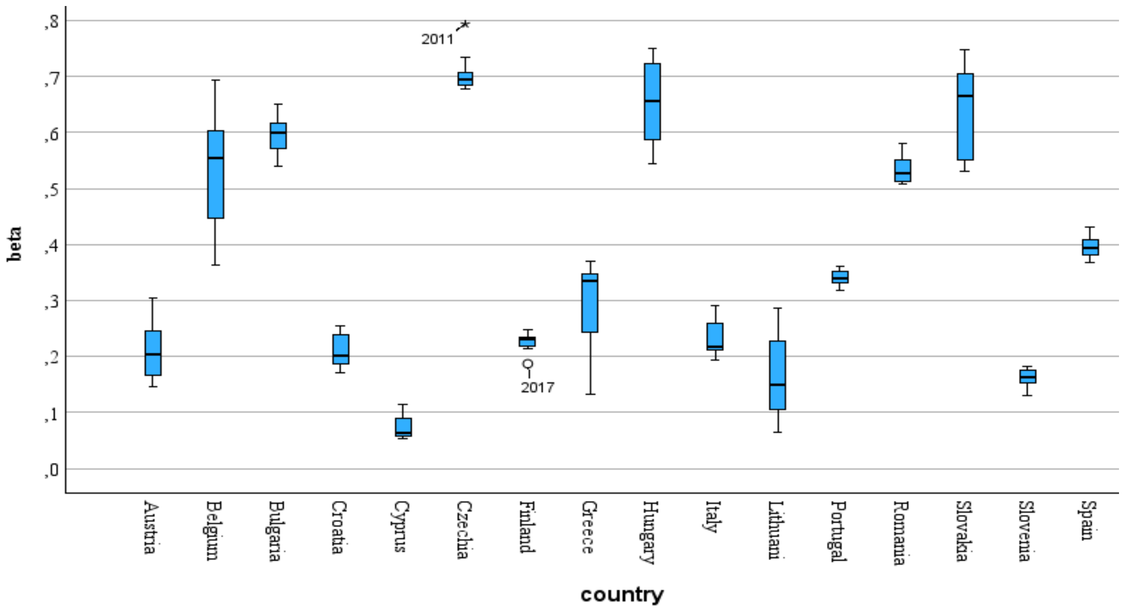


Figure 3 Total efficiency *beta* for inefficient countries in 2011-2022

5 Conclusions

In this study, we evaluated the energy and emission efficiency of EU countries during 2011-2022. This research highlights significant disparities in energy and emission efficiency in EU countries. The study confirmed the usual number of efficient EU countries (41%) in optimizing the extended and nonorientated DEA model with DDF. The main cause of inefficiency was the low share of renewable energy in final energy consumption, and another critical factor was the high level of undesirable output of air pollutants and greenhouse gases. When examining energy efficiency in two sectors, it was found that the problem for most inefficient countries is the high consumption of energy in the transport sector compared to the industry sector. The exception was Finland. By identifying best practices and areas that need improvement, this study provides a roadmap for policymakers to improve energy sustainability and reduce environmental impacts in the EU.

Further research should include more variables, such as energy storage capacities and technological advancements, and apply dynamic DEA models to account for changes over time. Additionally, extending the analysis to non-EU countries could provide a broader comparative perspective.

Acknowledgments

This article has been produced with the financial support of the European Union under the REFRESH – Research Excellence for Region Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition and by the SGS project (SP2024/083) of the Faculty of Economics, VŠB-Technical University of Ostrava.

References

Borozan, D., & Starcevic, D. P. (2019). European energy industry: Managing operations on the edge of efficiency. *Renewable and Sustainable Energy Reviews*, *116*. https://doi.org/10.1016/j.rser.2019.109401

Cámara-Aceituno, J., Hermoso-Orzáez, M. J., Terrados-Cepeda, J., Mena-Nieto, Á., & García-Ramos, J. E. (2024). Application of the data envelopment analysis technique to measure the environmental efficiency of the 27 countries of the European Union during the period 2012–2020. *Clean Technologies and Environmental Policy*, *26*(5), 1477-1505. https://doi.org/10.1007/s10098-023-02553-9

Eurostat (2024a). *Eurostat database.* [Data set]. https://ec.europa.eu/eurostat/databrowser/explore/all/all_themes

Eurostat. (2024b). Energy consumption decomposition analysis on EU-27 and EU-27 members - 2023 edition. https://ec.europa.eu/eurostat/en/web/products-statistical-working-papers/w/ks-tc-23-006

Gökgöz, F., & Yalçın, E. (2023). An environmental, energy, and economic efficiency analysis for the energy market in the European Union. *Environmental Progress & Sustainable Energy*, *42*(4). https://doi.org/10.1002/ep.14068

Meng, F., Su, B., Thomson, E., Zhou, D., & Zhou, P. (2016). Measuring China's regional energy and carbon emission efficiency with DEA models: A survey. *Applied Energy*, *183*, 1-21. https://doi.org/10.1016/j.apenergy.2016.08.158

Sueyoshi, T., Yuan, Y., & Goto, M. (2017). A literature study for DEA applied to energy and environment. *Energy Economics*, *62*, 104-124. https://doi.org/10.1016/j.eneco.2016.11.006

Wang, K., & Wei, Y. -M. (2014). China's regional industrial energy efficiency and carbon emissions abatement costs. *Applied Energy*, *130*, 617-631. https://doi.org/10.1016/j.apenergy.2014.03.010

Wang, S., Chu, C., Chen, G., Peng, Z., & Li, F. (2016). Efficiency and reduction cost of carbon emissions in China: a non-radial directional distance function method. *Journal of Cleaner Production*, *113*, 624-634. https://doi.org/10.1016/j.jclepro.2015.11.079

Zhou, P., Ang, B. W., & Poh, K. L. (2008). A survey of data envelopment analysis in energy and environmental studies. *European Journal of Operational Research*, *189*(1), 1-18. https://doi.org/10.1016/j.ejor.2007.04.042

Zhou, P., Ang, B. W., & Wang, H. (2012). Energy and CO2 emission performance in electricity generation: A non-radial directional distance function approach. *European Journal of Operational Research*, *221*(3), 625-635. https://doi.org/10.1016/j.ejor.2012.04.022

Designing optimal transportation patterns for radiation accident recovery

Robert Hlavatý[1], Helena Brožová[2], Anna Selivanova[3], Tereza Sedlářová-Nehézová[4]

Abstract. In case of a radiation accident in a nuclear power plant (NPP), there is a possibility of radionuclide releases that would affect the vicinity of the NPP. Such an event subsequently requires responsible authorities to decontaminate the affected areas and remove the produced contaminated waste to the designated interim storage sites. Given the extent of the incident, the recovery of the areas may turn into long-term logistic operation involving considerable machinery and personnel. We propose an optimal routing methodology to approach this situation and employ linear optimization to deliver an effective solution to the problem. This specific routing problem is constrained not only by the vehicle capacities but also by the doses the personnel can take during the process. We seek a solution that minimizes the distance travelled by vehicles and, thus, the time for which the personnel is exposed to radiation. The presented methodology is based on the real-world measures that would be taken in case of such an incident. The model results would allow the authorities to plan sufficient vehicle and personnel availability, estimate the time needed to clear contaminated areas and estimate the capacity of the interim storage sites. We demonstrate our methodology on a small example.

Keywords: dose rate, linear optimization, nuclear power plant, radiation accident, robust optimization, routing problem, waste

1 Introduction

Our motivation is based on cooperation with the National Radiation Protection Institute (SÚRO). Within the cooperation, we assessed the consequences of radiation accidents at nuclear power plants (NPP). These accidents can take many forms, and the one discussed in our paper considers that the radiation release affects the vicinity of NPP. In such cases, released radionuclides contaminate the land around NPP. The contamination is not homogeneous, and its levels are determined by many factors, such as wind direction and atmospheric conditions, that play a significant role in the spread of radionuclides (Wang et al., 2011). Also, the geographical layout, including topography and proximity to water bodies, also affects the dispersion pattern of radiation. Additionally, the type of radioactive material released, its quantity, and the duration of the release are essential parameters for simulation accuracy (Wang et al., 2011). Based on the simulation, it is possible to predict the contamination levels in different sectors around the NPP. This may vary depending on the various simulation scenarios. Disposing of contaminated material is essential to reduce the long-term environmental and health risks associated with radioactive substances (Linet et al., 2018). The experience is known in the decontamination of radioactive soil at the National Research Center Kurchatov Institute (Volkov et al., 2011). The study highlights the complexities of managing contaminated soil, including assessing the composition of contamination, soil volume, decontamination processes, and monitoring radiation conditions during disposal.

In order to dispose of the contaminated soil, it is necessary to initiate a significant logistical operation requiring careful planning to minimize risks to personnel exposed to radiation (Tashlykov et al., 2022). Tashlykov et al. (2022) employ routing algorithms to develop an optimal strategy for dosimetrists moving around in inhomogeneous radiation fields. This approach is suitable in the environments that are a priori known. In the case of the radiation accident we discuss here, such knowledge is unavailable, and the personnel operates in an environment where it is hard to avoid radiation. Under ideal conditions, there is a negligible radiological impact on the neighbourhood of the transported radioactive material (Calleja and Gutiérrez, 2011). However, in this case, the contaminated material would be transported by ordinary trucks without safety containers, which would cause the truck crew to receive radiation doses that need to be considered. Doses from transporting radioactive materials under different

[1] CZU Prague, Department of Systems Engineering, Kamýcká 129, 16500 Prague-Suchdol, hlavaty@pef.czu.cz
[2] CZU Prague, Department of Systems Engineering, Kamýcká 129, 16500 Prague-Suchdol, brozova@pef.czu.cz
[3] National Radiation Protection Institute (SÚRO), Bartoškova 1450/28, 14000, Prague-Nusle, anna.selivanova@suro.cz
[4] CZU Prague, Department of Systems Engineering, Kamýcká 129, 16500 Prague-Suchdol, nehezova@pef.czu.cz

scenarios were studied by Chun, Park and Cheong (2020). Under normal conditions, there are designated quantitative procedures that minimize both transportation costs and risk, as described by Alumur and Kara (2005). In case of an anticipated radiation accident, there is no information about the availability of such countermeasures, and the transportation procedures solely focus on carrying away the contaminated soil to designated storage sites. In the aftermath of the Fukushima Daiichi accident, it has been observed that there is a problem with large amounts of contaminated waste, which disposal workers must handle. These workers must be accordingly protected from the radiation exposure (Yasui, 2014). In our contribution, we develop a methodology based on the multi-index transportation problem, a classic approach introduced by Haley (1963). The multi-index transportation problem has many practical uses, as Junginger (1993) shows. We present our variant of the multi-index transportation problem, which is, in our case, incorporated in a two-stage optimization procedure. This approach looks suitable for the problem in question, i.e. searching for an optimal way of transporting the contaminated waste to storage sites while considering the radiation doses received during that process. We accompany our methodology with a small example based on real but anonymized data.

2 Methods

The following chapter describes the quantitative approach to solving the problem in question. The model is based on the factual circumstances that arise in case of a radiation accident. The logical framework was provided by SÚRO and corresponds to the real reaction of the authorities to a radiation accident. The model is invariant to the location or national policies of different countries that are activated in case of such an accident.

2.1 Routing model

NPP's emergency planning zone (EPZ) is divided into several segments. The radioactive waste is gathered and concentrated on each segment at one designated loading area. At each of the I loading areas, the radioactive waste is loaded onto a vehicle with a limited capacity and transported to one of the J interim waste storages. Each vehicle is operated by a crew which is exposed to radiation during transportation. The general problem statement is the following: *we seek the optimal way to transport the radioactive waste from all segments to designated interim waste storages while minimizing the time travelled by all vehicles and thus minimizing the dose the crew receives during the transport.*

First, it is necessary to set up all variables and parameters that will be subject to optimization. Then, the mathematical two-stage optimization model is constructed. During the first stage, we determine the necessary workforce (minimum number of crews needed to manage the whole problem). As a *crew*, we understand a single truck with one driver and one assistant. The driver and assistant can be mutually rotated during the manipulation and transportation of the waste. In the second stage, we compute the optimal workforce distribution among the crews with respect to minimization of travelling time, i.e., from which loading area to which waste storage the radioactive waste should be carried and by which crew.

Input parameters

The optimization problem (in the second stage) resembles a three-dimensional traditional transportation problem, and the individual dimensions are represented by the *loading areas*, *waste storages* and *crews*. We consider I loading areas, J waste storages and K crews and establish the fundamental variable of our problem $x_{ijk} \in \mathbb{R}^{I\times J\times K}$ representing the number of tours from loading area $i = 1, \dots, I$ to waste storage $j = 1, \dots, J$ by crew $k = 1, \dots, K$. Next, we define a set of all parameters that enter the optimization models in the first or the second stage.

Parameter	Description	Unit
t_{ij}	Travel time between the loading area i and waste storage j	h
a_i	Amount of radioactive waste gathered at the loading area i	t
MA_i	Caesium massic activity of the radioactive waste gathered in the loading area i	Bq/kg
α	Conversion coefficient for radioactive waste handling	(mSv·kg)/(Bq·h)
b_j	The capacity of the waste storage j	t
DR_i	Radioactive waste dose rate at the loading area i	mSv/h
T_i	Number of tours needed to carry away all radioactive waste from the loading area i	-

ATC	Average truck capacity	t
MT_i	Manipulation time for loading at loading area i	h
MT_j	Manipulation time for unloading at waste storage j	h
$TMTD$	Total manipulation time dose received during the whole problem	mSv
$TTTD$	Total transportation time dose received during the whole problem	mSv
DL	Maximum dose limit for all crews $k = 1, \dots, K$	mSv

First-stage optimization

In the first stage, we compute the minimum necessary number of crews needed to manage the problem. The number of required crews is given by the dose limit DL that a crew can take. Once this limit is reached, the crew can no longer participate in the transportation. It is expected that the crews receive the doses over time spent by a) manipulation, i.e. loading and unloading the waste; b) transporting the waste from the loading area to the waste storage. The total manipulation time dose ($TMTD$) can be easily computed without using an optimization model as follows:

$$TMTD = \sum_{i=1}^{I}\sum_{j=1}^{J}\left(\frac{\left(MT_i + MT_j\right) * ATC * DR_i}{a_i}\right) \tag{1}$$
$$DR_i = \alpha * MA_i, \forall i$$

The first term sums up the manipulation time $\left(MT_i + MT_j\right)$ which is multiplied by the dose rate DR_i normalized per ton and the number of tons contained in the truck (ATC). The total dose rate of the radioactive material at the loading area i is $DR_i = \alpha * MA_i, \forall i$ and the dose per 1t is DR_i/a_i.

However, the total transportation time dose received during the whole problem is not a priori known because it is not apparent which waste storage will be used for which loading areas and, thus, how much time the transportation would take. Therefore, it is necessary to solve the optimization problem that would arrange the transportation in a way that minimizes the total transportation time dose ($TTTD$) and thus minimizing the objective function value of the problem (2), which shows the $TTTD$ for the problem. We propose the following optimization model with the variable x_{ij} representing the number of tours from loading area $i = 1, \dots, I$ to waste storage $j = 1, \dots, J$

$$TTTD = \min \sum_{i=1}^{I}\sum_{j=1}^{J}\left(\frac{t_{ij} * ATC * DR_i}{a_i}\right) x_{ij}$$

subject to

$$\sum_{j=1}^{J} x_{ij} = T_i, i = 1, \dots, I \tag{2}$$
$$\sum_{i=1}^{I} x_{ij} \le \frac{b_j}{ATC}, j = 1, \dots, J$$
$$x_{ij} \ge 0, \forall i, j$$

The expression in the brackets in the objective expresses the dose rate DR_i normalized per ton (divided by a_i) received per transportation time t_{ij} and transported amount ATC. The first constraint ensures that the required number of tours from each loading area i is delivered. The value of the parameter T_i is computed as $T_i = \frac{a_i}{ATC}$. The second constraint ensures that the waste storage capacity is not exceeded. It is expected that the total capacity of the waste storage is sufficient to store all radioactive waste, i.e. $\sum_{j=1}^{J} b_j \ge \sum_{i=1}^{I} a_i$. Then, the number of required crews to manage the problem is computed as $K = \left\lceil \frac{TMTD+TTTD}{DL} \right\rceil$.

Second-stage optimization

Completing the first stage reveals the number K of crews that will be needed to cover the problem in the ideal case. In the second stage, it is necessary to assign the crews to transport the radioactive waste from loading areas to waste storages. Once again, we seek to minimize the total dose received by the crews while distributing the transportation among crews.

$$\min \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(\frac{t_{ij} * ATC * DR_i}{a_i} + \frac{(MT_i + MT_j) * ATC * DR_i}{a_i} \right) x_{ijk}$$

subject to

$$\sum_{j=1}^{J} \sum_{k=1}^{K} x_{ijk} = T_i, i = 1, \dots, I$$

$$\sum_{i=1}^{I} \sum_{k=1}^{K} x_{ijk} \leq \frac{b_j}{ATC}, j = 1, \dots, J \tag{3}$$

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \left(\frac{t_{ij} * ATC * DR_i}{a_i} + \frac{(MT_i + MT_j) * ATC * DR_i}{a_i} \right) x_{ijk} \leq DL, k = 1, \dots, k$$

$$x_{ijk} \geq 0, \forall i, j, k$$

The objective function minimizes the total dose received on transportation and manipulation for each crew on all tours. Same as in the first-stage model, the first constraint secures that the required number of tours from each loading area i is delivered, and the second constraint secures that the waste storage capacity is not exceeded. The third constraint ensures that no crew receives a higher dose than DL during the overall transportation.

3 Results

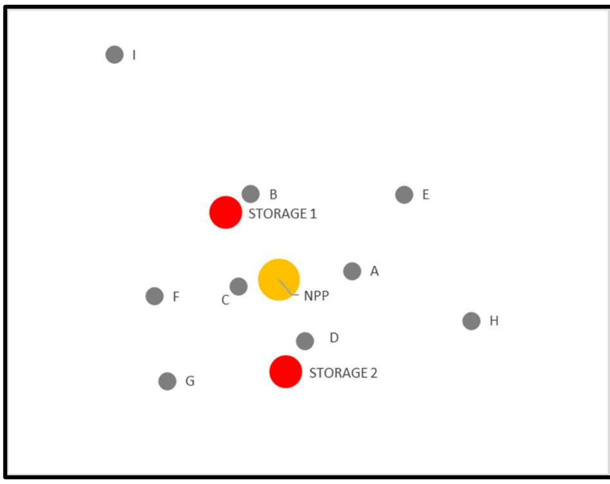
To demonstrate the applicability of our approach, we proceed with a small, real-based example. Due to data sensitivity, it is not possible to disclose detailed information. We have experimented with our approach to the situation with an existing EPZ divided into several sectors that are subject to simulation. Based on the simulations carried out by SÚRO, it is possible to set up different scenarios of land contamination for these sectors. Each sector has a designated loading area, and at the same time, there are two designated storage sites that would act as a destination for contaminated material disposal. For simplicity and due to the extent of the conference paper, we assume only nine sectors $i = 1, \dots, 9$ named A…I for the demonstration purposes. Note that these sectors are not topologically identical and depend on the land characteristics.

Parameter	Value
MT_i	0.5 h $\forall i$
MT_j	0.5 h $\forall j$
ATC	8 t
a^T	(27, 49, 71.8, 10.5, 21, 20, 20, 40, 85) 10^3 t
α	0.0000015 (mSv·kg)/(Bq·h)
MA^T	(1.12, 25, 4.5, 0.5, 0.5, 1.5, 0.15, 0.5, 5) 10^6 Bq/kg
b_j	∞ t, $j = 1,2$ (unlimited capacity)
DL	0.5 mSv

Travel time matrix (min)	Storage 1	Storage 2
A	15.45	9.61
B	2.73	13.26
C	5.21	7.36
D	12.82	3.12
E	17.87	19.58
F	11.64	16.54
G	17.19	13.49
H	27.85	17.27
I	14.88	24.86

Table 1 Model input

We begin with the 1st stage of the optimization problem with the input described in Table 1. The numbers were provided by SÚRO except for the travel times between destinations and storages, which can be easily computed using any available map tool. We proceed with computing terms (1) and (2), and after doing so, it occurs that $TMTD + TTTD = 1.27$mSv and thus the number of crews needed to secure the transportation is $K = 3$, considering the given dose limit 0.5mSv per worker. Figure 1 shows the topology of the problem in terms of locations, and it was created as a projection into the plane using their earth coordinates. Figure 2 shows the optimal distribution of contaminated material to both storage sites regardless of the crew distribution. Figure 2 only serves as an informative byproduct of the computing model (2).



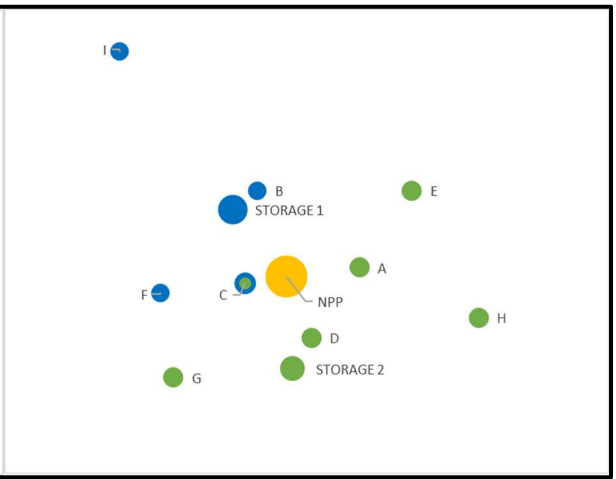


Figure 1 Problem topology **Figure 2** Distribution to storages

In Figure 2, the blue loading sites are affiliated with *Storage 1,* and the green loading sites are affiliated with *Storage 2*. The material at loading site C is split and distributed to both storage sites in the optimal case. Next, we proceed with the 2nd stage of the optimization problem and compute problem (3) for $K = 3$. Figure 3 shows the optimal distribution among all three transportation crews.

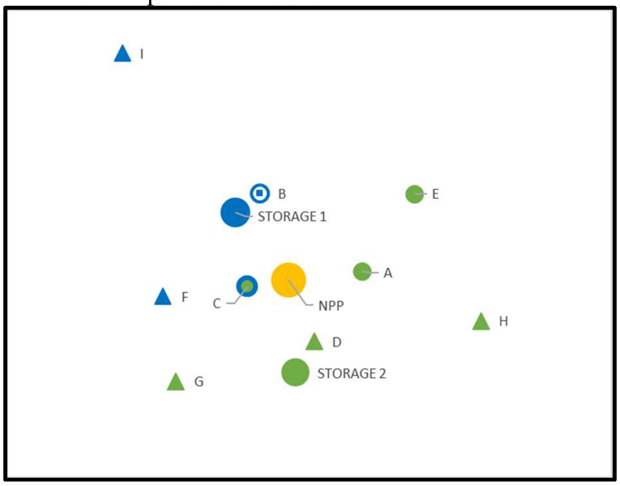


Figure 3 Optimal distribution among crews

Figure 3 has the following interpretation: Material from blue loading sites will be transported to *Storage 1*. Material from green loading sites will be transported to *Storage 2*. Material from *loading site C* will be split and transported to both storage sites. This distribution is no different from the one in Figure 2. The loading sites marked as circles will be served by *Crew 1*, loading sites marked as triangles will be served by *Crew 2*, and loading sites marked as squares will be served by *Crew 3*. Note that the only square is *loading site B* together with a circle because it needs to be served by two crews.

Figure 3 demonstrates the distribution among three crews such that a single crew does not receive a dose higher than 0.5mSv. The total distribution of doses among crews is 0.49mSv, 0.5mSv, and 0.28mSv, respectively. Knowing the exact distribution and the number of tours needed to carry all the contaminated material to storage sites, it is simple to evaluate the total time required for such a job. In our demonstrative case, it would be 7760 hours of driving and considerably more hours of manipulation, being 43037 hours, if one hour is needed for every tour for loading and unloading. This parameter may also be subject to different scenarios. Calculating the total decontamination time for our scenario, it takes approximately five years and ten months of working time for three crews. It is not utterly important to employ only three crews in such a case; it is only the lower limit for such a number. Any larger number of crews would be acceptable and probably desirable, considering the total time needed for transportation. This is why our approach is rather a decision-making support than a strict planning tool.

Conclusion

The illustrative example is only a simple variant of the actual real problem. We intend to provide a complex analysis of the real-based situation with considerably larger numbers of loading areas. There are many possibilities for different transportation scenarios. This is mainly influenced by the amount of contaminated material at the loading sites, its massic activity, the number and capacity of storage sites and the availability of transportation crews. Especially the amount of the material and the intensity of massic activity are only predictable by simulation procedures but the reality in case of such an accident may as well differ a lot. This is why our further attempts will aim to work with uncertain parameter values and apply the robust optimization approach to the problem in question.

Acknowledgements

This paper was supported by the Czech Science Foundation (GAČR) project No. P403-22-11117S.

References

[1] Alumur, S., Kara, B. (2007). A new model for the hazardous waste location-routing problem. Computers & Operations Research, 34(5), 1406-1423. https://doi.org/10.1016/j.cor.2005.06.012

[2] Calleja, J., Gutiérrez, F. (2011). Radiological impact associated with the transport by road of radioactive material in Spain. Radioprotection, 46(3), 331-343. https://doi.org/10.1051/radiopro/2011109

[3] Chun, G., Park, J., Cheong, J. (2020). Calculation of potential radiation doses associated with predisposal management of dismantled steam generators from nuclear power plants. Sustainability, 12(12), 5149. https://doi.org/10.3390/su12125149

[4] Haley, K. (1963). The multi-index problem. Operations Research, 11(3), 368-379. https://doi.org/10.1287/opre.11.3.368

[5] Junginger, W. (1993). On representatives of multi-index transportation problems. European Journal of Operational Research, 66(3), 353-371. https://doi.org/10.1016/0377-2217(93)90223-a

[6] Linet, M., Kazzi, Z., Paulson, J. (2018). Pediatric considerations before, during, and after radiological or nuclear emergencies. Pediatrics, 142(6), e20183001. https://doi.org/10.1542/peds.2018-3001

[7] Tashlykov, O., Sesekin, A., Chentsov, A., & Chentsov, A. (2022). Development of methods for route optimization of work in inhomogeneous radiation fields to minimize the dose load of personnel. Energies, 15(13), 4788.

[8] Volkov, V., Danilovich, A., Zverkov, Y., Ivanov, O., Koltyshev, S., Lemus, A., Potapov, V., Semenov, G., Stepanov, V., Chesnokov, A., Shisha, A. (2011). Experience in decontamination of radioactive soil on the grounds of the national research center Kurchatov Institute. Atomic Energy, 110(2), 123-130. https://doi.org/10.1007/s10512-011-9400-x

[9] Wang, G., Zhao, W., Zhao, J., Dai, D., Song, Y., Song, Z. (2011). Predicting the spread of nuclear radiation from the damaged Fukushima nuclear power plant. Chinese Science Bulletin, 56(18), 1890-1896. https://doi.org/10.1007/s11434-011-4513-0

[10] Yasui, S. (2014). New regulations for radiation protection for work involving radioactive fallout emitted by the Tepco Fukushima Daiichi app accident—disposal of contaminated soil and wastes. Journal of Occupational and Environmental Hygiene, 11(10), 147-156. https://doi.org/10.1080/15459624.2014.922688

Application of chooser options to valuation of investment opportunities

Jiří Hozman[1], Tomáš Tichý[2]

Abstract. Real options approach can be applied to a variety of investment opportunities to help investors improve risk management and make more informed decisions. Since a chooser option is a contract written on the maximum of the set of individual real options providing certain (mutually different) investment opportunities, this flexibility makes chooser options particularly relevant for valuing investment opportunities that involve uncertainty and the need for strategic decision-making. In this paper, we present an incorporation of chooser options into the valuation process and examine their valuation via contingent claims framework under single-factor uncertainty. The resulting PDE problems are of the Black-Scholes type and are solved using a discontinuous Galerkin approach. Finally, we provide a simple conceptual example of how chooser options can be used in investment scenarios facing a choice between expansion, contraction, or total abandonment.

Keywords: real options pricing, flexibility value, chooser option, Black-Scholes model, discontinuous Galerkin method

JEL Classification: C44, G13
AMS Classification: 65M60, 35Q91, 91G60

1 Introduction

Real options valuation is one of the key decision-making tools that applies conventional financial options theory to evaluate strategic investments and takes into account irreversibility of investments, choice of timing and uncertainty over the future rewards from investments, see Dixit and Pindyck (1994). These attributes are more evident for situations where choosing among several investment opportunities is considered, see, e.g. Trigeorgis (1993). Therefore, the proper integration of chooser options, a type of exotic option, into investment valuation within the context of real options valuation establishes a framework for assessing the value of flexibility, embedded in a portfolio of investment opportunities. Regarding valuation techniques related to real options approach, the contingent claims analysis, leading to formulations via partial differential equations (PDEs) and their subsequent numerical solutions, provides complex results and comprehensive information on modern investment issues, usable in further post-processing.

The concept of this contribution arises from recent conference papers due to Hozman and Tichý (2021) and Hozman and Tichý (2022), where expansion and contraction real options under output price uncertainty were studied in a separate way. In contrast, the aim of this paper is to incorporate the concept of chooser options into valuation of the managerial flexibility allowing to choose among several investment strategies. Inspired by the numerical pricing schemes for conventional financial options by Hozman and Tichý (2020), we here extend the discontinuous Galerkin (DG) approach to the case of chooser options. We proceed as follows: in Section 2 the portfolio of selected investment opportunities is formulated and relevant PDE pricing models are introduced, while in Section 3 a numerical valuation scheme is presented. Finally, in Section 4 a simple comprehensible application to a multi-scenario decision problem is provided.

2 Chooser option as a portfolio of investment opportunities

In this paper we provide a flexibility valuation of a one-stage investment project with the following opportunities to

(E) *expand* the production rate by factor $\kappa_1 > 1$ requiring the implementation cost $\mathcal{K}_1 > 0$;
(C) *contract* the production rate by factor $\kappa_2 \in (0, 1)$ requiring the implementation cost $\mathcal{K}_2 > 0$;
(A) *abandon* the project for its salvage price $\mathcal{K}_3 > 0$.

Making a single decision above is related to any time between the present time and the date of expiration $T > 0$. These opportunities can be considered separately (i.e., E, C, A) or in combination as a portfolio of opportunities (i.e.,

[1] Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic, jiri.hozman@tul.cz
[2] Department of Finance, VSB-TU Ostrava, 17. listopadu 15, 708 00, Ostrava, Czech Republic, tomas.tichy@vsb.cz

E&C, E&A, C&A, E&C&A). In terms of conventional financial options, the separated opportunity is described by a call option (on expansion) or a put option (on contraction or to abandon) under American exercise right with the given strikes $\mathcal{K}_i$ and maturity date T. On the other hand, the combined value of two or more individual options may be seen as an American chooser option, written on the maximum of the set of these individual options providing certain single investment opportunities (e.g., expansion, contraction, abandonment). Note that if the realization of the investment opportunity is allowed only at the time instant T, we speak about the European exercise right.

In what follows, we briefly recall PDE models by Li and Wang (2019) to valuate investment project flexibility. The standard approach is based on backward induction, starting from determination of project values and ending with evaluation of desired embedded option values. In line with Li and Wang (2019), we assume that the value of the project as well as the flexibility of the investment opportunity are possible to express as a function of output price P and actual time t. We denote by $V_0(P,t)$ the value of the project having no investment opportunities. In contrast, the functions $V_1(P,t)$, $V_2(P,t)$ and $V_3(P,t)$ stand for the values of an investment project with the embedded option (E), (C) and (A), respectively. The key issue is to find the true value of the flexibility to invest $F(P,t)$, $t \in [0,T)$, within the selected combination of opportunities under consideration. Without loss of generality, assume the combination E&C&A, then on the expiration date (i.e., $t = T$), the values of this investment opportunity and the relevant projects are interconnected through the payoff of the corresponding chooser option, i.e.,

$$F(P,T) = \Pi(P) \equiv \max\Big(\underbrace{V_1(P,T) - V_0(P,T) - \mathcal{K}_1}_{\text{flexibility (E)}}, \underbrace{V_2(P,T) - V_0(P,T) - \mathcal{K}_2}_{\text{flexibility (C)}}, \underbrace{\mathcal{K}_3 + V_3(P,T) - V_0(P,T)}_{\text{flexibility (A)}}, 0\Big), \quad (1)$$

for $P > 0$. The modification of (1) for cases E&C, E&A, C&A or separated opportunities is straightforward by removing the specific arguments of the *max*-function.

In order to evaluate $\Pi(P)$ we are faced with the task of determining values of all investment projects included in (1) at $t = T$. Obviously, $V_3(P,T) = 0$ due to the decision (A), but in the remaining possibilities we proceed as follows. We assume that fluctuations in project/flexibility values are tracked back to uncertainty via the output price P, driven by a geometric Brownian motion with a drift factor $r - \delta$ (i.e., risk-free interest rate $r > 0$ and mean convenience yield $\delta > 0$) and volatility $\sigma > 0$. This concept introduces a certain degree of uncertainty in the investment opportunity reflecting the fluctuations in values of the underlying variable P in real-world scenarios. We follow the contingent claims framework as in Haque et al. (2014). Incorporating the cash flow rate function $\varphi_i(P,t)$, $i = 0,1,2$, within a finite time horizon (e.g., limited resources), which are associated with the corresponding investment project, we conclude that value functions V_i, $i = 0,1,2$, satisfy the following deterministic backward PDEs of the Black-Scholes (BS) type, i.e.,

$$\frac{\partial V_i}{\partial t} + \frac{1}{2}\sigma^2 P^2 \frac{\partial^2 V_i}{\partial P^2} + (r-\delta)P\frac{\partial V_i}{\partial P} - rV_i = -\varphi_i \quad \text{in } (0,\infty) \times [T, T^*),\ i = 0,1,2, \quad (2)$$

where $T^* \gg T$ stands for the maximal life-time of all considered investment projects. At that time, we assume that all projects are already worthless, i.e., $V_i(P,T^*) = 0$, $i = 0,1,2$, which poses the terminal conditions for the problem (2).

The last task is to determine the flexibility (i.e., option) values of the embedded investment opportunity. Repeating the same steps as above also for the value function $F(P,t)$, we conclude that flexibility as an option premium is driven by the governing equation with the same BS differential operator as in (2), but with zero right-hand side. The possibility to realize the embedded flexibility known as European constraint is incorporated into this equation by terminal condition (1). Moreover, the American option style is partially interesting, allowing exercise any time in $[0,T]$. This early exercise imposes an additional constraint that $F(P,t) \geq \Pi(P)$ for any $t \in [0,T)$.

There are several approaches how to handle the early exercise feature, among the widely used ones, just penalty techniques by Zvan et al. (1998) allow us to formulate the option (i.e., flexibility) valuation problem as follows

$$\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 P^2 \frac{\partial^2 F}{\partial P^2} + (r-\delta)P\frac{\partial F}{\partial P} - rF + q_F = 0 \quad \text{in } (0,\infty) \times [0,T), \quad (3)$$

where the penalty term q_F is defined to ensure American constraint by using conditions

$$q_F(P,t) = 0, \text{ if } F(P,t) > \Pi(P), \qquad q_F(P,t) > 0, \text{ if } F(P,t) = \Pi(P), \quad t \in [0,T). \quad (4)$$

Note, if we put $q_F(P,t) = 0$, for all $P > 0$ and $t \in [0,T)$, in the case of European exercise right, the penalty approach can be unified for both exercise features considered.

In summary, we deal with several terminal value problems for different investment projects and one for the embedded option, the formulations of which differ only slightly. Nevertheless, the localization of the governing equations (2) and (3) to a bounded interval $\Omega = (0, P_{\max})$ is necessary for the subsequent numerical treatment. Therefore we have to impose project as well as option values for zero and sufficiently large (i.e., $P_{\max}$) output prices. The project values are estimated by the net present value approach for the given cash flow rates as follows

$$V_i(z,t) = \int_t^{T^*} \varphi_i(z,\xi) e^{-r(\xi-t)} \mathrm{d}\xi, \qquad z \in \{0, P_{\max}\},\ t \in [T, T^*),\ i = 0,1,2. \tag{5}$$

Considering flexibility values, the exercise rights lead to a couple of Dirichlet boundary conditions in the form

$$F(z,t) = \begin{cases} e^{-r(T-t)}\Pi(z), & \text{(European)} \\ \Pi(z), & \text{(American)} \end{cases}, \qquad z \in \{0, P_{\max}\},\ t \in [0, T). \tag{6}$$

3 Numerical valuation

Regarding the real options valuation problems, one can apply various numerical approaches to propose efficient pricing algorithms, arising from pricing of financial option counterparts. Within this contribution, the proposed valuation methodology is based on the DG approach by Rivière (2008) to cope with the convection-diffusion character of the governing equations, determined by ratio $(r-\delta)/\sigma^2$. Accordingly, the numerical solutions, representing approximate project as well as real option values, are constructed as the piecewise polynomial functions of the p-th order defined on the spatial grid $0 = P_0 < P_1 < \ldots < P_N = P_{\max}$ (with the assigned mesh size h) without necessitating continuity across these partition nodes. Let the space of such functions be denoted by $S_h^p(\Omega)$, for a detailed description the reader is referred to Rivière (2008).

One of the crucial steps of the valuation procedure is the treatment of the American early exercise feature in (3). More precisely, to meet the conditions (4), we introduce the variational form of the penalty function q_F as follows

$$(q_F(\cdot,t), v_h) = \underbrace{c_{\mathrm{p}} \int_\Omega \chi_{\mathrm{exe}}(t)\Pi(\cdot) v_h \,\mathrm{d}P}_{Q_{\mathrm{R}}(v_h)} - \underbrace{c_{\mathrm{p}} \int_\Omega \chi_{\mathrm{exe}}(t) F(\cdot,t) v_h \,\mathrm{d}P}_{Q_{\mathrm{L}}(F, v_h)}, \quad v_h \in S_h^p(\Omega), \tag{7}$$

where $(\cdot,\cdot)$ denotes in fact the inner product in $L^2(\Omega)$. The function $\chi_{\mathrm{exe}}(t)$ in (7) is defined as an indicator function of the exercise region at time instant t (see Hozman and Tichý, 2020) and $c_{\mathrm{p}} > 0$ represents a weight to enforce the early exercise. To increase the clarity of the numerical valuation procedure, we split the form (7) into linear functional Q_{R} and bilinear form Q_{L}.

With respect to the space-time domain of governing equations, the numerical valuation consists of two consecutive phases: spatial semi-discretization and temporal discretization. Within the first phase, at each time instant, we construct the semi-discrete solution, represented as the solution to the system of ordinary differential equations (ODEs) with the given terminal condition, see Hozman and Tichý (2020). The second phase is then devoted to the discretization of these ODE systems using a semi-implicit Euler scheme and results into a sequence of linear algebraic problems (with sparse matrices) related to a time partition $T^* = t_0 > t_1 > \ldots t_R = T > t_{R+1} > \ldots > t_M = 0$ with fixed time step $\tau = T^*/M$. Let us denote $u_{h,m}^{(i)} \in S_h^p(\Omega)$, $i = 0,1,2$, the approximation of the corresponding project value functions V_i from (2) at time level $t_m \in [T, T^*]$. Similarly, we put $w_h^m \approx F(\cdot, t_m)$, $t_m \in [0, T]$ from (3). Assuming the portfolio of investment opportunities E&C&A, the sought present value of such flexibility w_h^M is computed using the backward induction procedure in the following steps (note that $t_{m+1} - t_m = -\tau$):

(S1) Set homogeneous initial states $u_{h,0}^{(i)} = 0$, $i = 0,1,2$.

(S2) The discrete project value functions $u_{h,m}^{(i)}$ for $m = 1,2,\ldots,R$ and $i = 0,1,2$ are defined recursively by the following scheme

$$\left(u_{h,m}^{(i)}, v_h\right) - \tau \mathcal{D}_h\left(u_{h,m}^{(i)}, v_h\right) = \left(u_{h,m-1}^{(i)}, v_h\right) - \tau \ell_h^{(i)}(v_h)(t_m) + \tau\left(\varphi_i(t_m), v_h\right) \quad \forall\, v_h \in S_h^p(\Omega), \tag{8}$$

where the bilinear form $\mathcal{D}_h(\cdot,\cdot)$ stands for the DG discrete variant of the BS differential operator and the linear form $\ell_h^{(i)}(\cdot)(t)$, $i = 0,1,2$, enforces the boundary conditions (5).

(S3) Set the terminal condition for the flexibility value function as relevant chooser (real) option by S_h^p-projection of $\Pi(\cdot)$ from (1), i.e.,

$$\left(w_h^R, v_h\right) = \left(\max\left(u_{h,R}^{(1)} - u_{h,R}^{(0)} - \mathcal{K}_1, u_{h,R}^{(2)} - u_{h,R}^{(0)} - \mathcal{K}_2, \mathcal{K}_3 - u_{h,R}^{(0)}, 0\right), v_h\right) \quad \forall\, v_h \in S_h^p(\Omega). \tag{9}$$

(S4) Similarly to (S2), the discrete flexibility value function w_h^m for $m = R+1, \dots, M$ is defined as follows

$$\left(w_h^m, v_h\right) - \tau \mathcal{D}_h\left(w_h^m, v_h\right) + \tau \mathcal{Q}_h\left(w_h^m, v_h\right) = \left(w_h^{m-1}, v_h\right) - \tau \ell_h(v_h)(t_m) + \tau q_h(v_h)(t_m) \quad \forall v_h \in S_h^p(\Omega), \tag{10}$$

where the linear form $\ell_h(\cdot)(t)$ balances boundary conditions (6) and forms $\mathcal{Q}_h(\cdot,\cdot)$ and $q_h(\cdot)(t)$ stands for the DG discrete variants of $\mathcal{Q}_\mathrm{L}$ and $\mathcal{Q}_\mathrm{R}$ from (7).

For a detailed description of forms in (8) and (10), we refer the interested reader to Hozman and Tichý (2020). The modification of the valuation scheme (S1)–(S4) for different set of investment opportunities is again straightforward.

4 Application to conceptual case study

The presented conceptual study reflects real-world scenarios based on idealized basis from the iron ore mining industry as in Li and Wang (2019). The goal of the application is to comprehensively interpret and analyze the outcomes of the valuation procedure (S1)–(S4), which represent managerial flexibility considered in the form of various combinations of selected individual options. Specifically for a mining company, it encompasses four different combinations (i.e., E&C, E&A, C&A and E&C&A) under the European as well as American exercise rights constructed from the following actions:

(E) to double iron ore production (i.e., $\kappa_1 = 2$) requiring the additional investment cost $\mathcal{K}_1 = 10$;
(C) to halve the production of an iron ore (i.e., $\kappa_2 = 0.5$) requiring the implementation cost $\mathcal{K}_2 = 0.5$;
(A) to abandon the mine (as a mining project) for its salvage price $\mathcal{K}_3 = 20$.

The values of $\mathcal{K}_1$, $\mathcal{K}_2$ and $\mathcal{K}_3$ are given in thousands of million USD and the realization of all investment opportunities is related to the time span of one year, i.e., $T = 1$.

Similarly to Li and Wang (2019), we consider a mining company with the given mining plan associated with iron ore mining production rate $g_0(t)$, expressed in thousands of million dry metric tonnes (dmt) of iron ore per year. The value of such a mining company with no embedded investment opportunity is described by the function $V_0(P,t)$, that depends on commodity price P, reported in USD per dmt of iron ore. Further, we consider two mining investment projects of values $V_1(P,t)$ and $V_2(P,t)$ that adopt the individual options to scale up and scale down iron ore production (as described above), respectively. These investment opportunities are consistent with how the mine is operated, i.e., with mining production rates $g_1(t)$ and $g_2(t)$. Accordingly, we define

$$g_0(t) = \begin{cases} s(t), & t \in [0, T_0^*), \\ 0, & t \in [T_0^*, T^*], \end{cases} \quad g_1(t) = \begin{cases} s(t), & t \in [0, T), \\ \kappa_1 \cdot s(t), & t \in [T, T_1^*), \\ 0, & t \in [T_1^*, T^*], \end{cases} \quad g_2(t) = \begin{cases} s(t), & t \in [0, T), \\ \kappa_2 \cdot s(t), & t \in [T, T_2^*), \end{cases} \tag{11}$$

where the function $s(t)$ represents a hypothetical mining production rate with unlimited resources and T_i^*, $i = 0, 1, 2$, denotes lifetime (in years) of the particular investment project. Intuitively, the contraction feature implies that $T^* = T_2^*$. Let Q denote the total reserve of the iron ore mine (in thousands of million dmt), then $\int_0^{T^*} g_i(\xi)\mathrm{d}\xi = Q$ for $i = 0, 1, 2$. Subsequently, in order to determine the particular (after-tax) cash flow rate (in thousands of million USD per year), we follow the definition by Haque et al. (2014) and prescribe

$$\varphi_i(P,t) = g_i(t)\Big((1-D)P - c(t)\Big)(1-B), \qquad P \in [0, P_{\max}],\ t \in [0, T^*],\ i = 0, 1, 2, \tag{12}$$

where $c(t)$ is the average cash cost per 1 dmt of iron ore, D the rate of state royalties and B the income tax rate.

Referring to Li and Wang (2019), the forthcoming numerical experiments are performed with the following project and market data: $Q = 10$, $s(t) = 0.1e^{0.007t}$, $D = 0.05$, $B = 0.30$, $c(t) = 35e^{0.005t}$, $\sigma = 0.30$, $r = 0.06$, $\delta = 0.02$, which are the representatives of parameter values of practical significance. More precisely, $c(0) = 35$ USD per dmt corresponds to prices from 01/2007. From the market data and (11) we get $T_0^* \doteq 75.80$, $T_1^* \doteq 43.24$ and $T_2^* \doteq 124.64$. Concerning the discretization aspects, we approximate project/flexibility value functions as piecewise quadratic (i.e., $p = 2$) on the fixed uniformly partitioned grid of domain Ω with $h = 0.25$ and $P_{\max} = 100$. The time step is set to $\tau = 0.02$ and in order to handle numerically the American early exercise feature, we choose $c_\mathrm{p} = 10^5/\tau$ in (7). The whole implementation is done in the solver FreeFem++, see Hecht (2012).

First, we consider European exercise rights, i.e., the realization of the investment opportunity is allowed only at time instant corresponding to the length of one year. Figure 1 captures the DG approximations of flexibility values in the whole spatial domain Ω for particular single-investment opportunity. More precisely, the terminal states corresponding to payoffs of relevant single options (as in the BS model) are depicted on the left graph. In contrast,

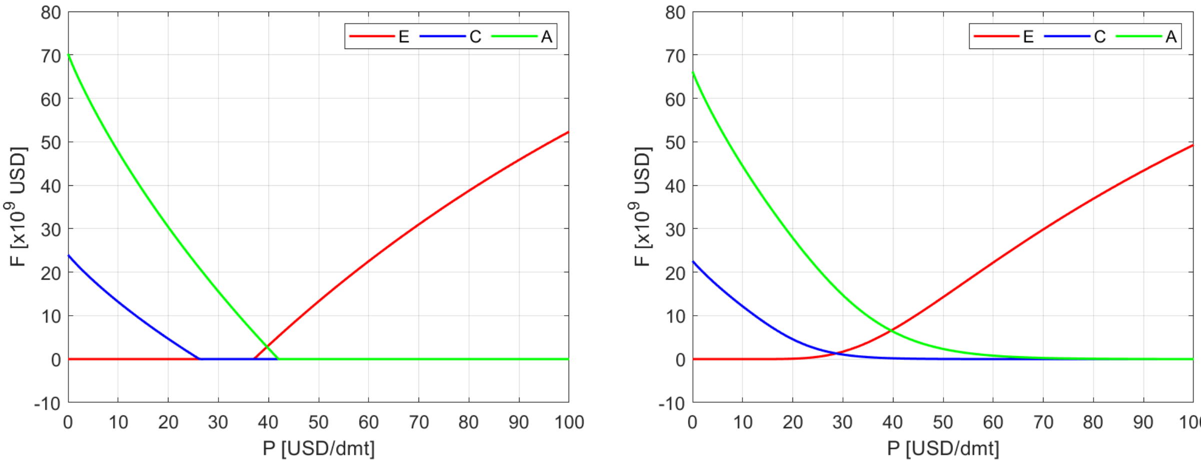


Figure 1 Flexibility values (under the European exercise right) of individual investment opportunities at realization time $t = T$ (left) and at present time $t = 0$ (right)

the right graph illustrates the smoothing of singularities from terminal conditions as time runs backward from $t = T$ to $t = 0$. In terms of financial options, one can easy recognize that shapes of flexibility value functions for (E) coincide with the character of financial call options and vice versa scenarios (C) and (A) match the put options.

Secondly, we focus on portfolios combined from the given investment opportunities. Figure 2 shows these four scenarios in comparison with the single-investment opportunities and a sum of their individual values. At first glance, one of the intuitive expectation is proven, namely that the flexibility value of the portfolio of investment opportunities is always more valuable than the value of each individual opportunity. Regarding their sum, if the two real options (both realized at the same time) are of opposite type (i.e., scenarios E&C and E&A), then they would in fact have no interaction and be purely additive, see top graphs. On the other hand, for real options of the same type, their interaction is not negligible. Specifically, the portfolio of two puts has a sub-additive property, see scenarios C&A and E&C&A. These observations are fully in line with the findings by Trigeorgis (1993).

As the last attribute, we investigate the effect of early exercise constraint (i.e., American exercise right) on flexibility values. Apart from that American real options cost more than their European counterparts, the smooth pasting condition (linking payoff and flexibility functions) allows us to indicate the optimal exercise prices for the realization of particular investment opportunities in a single-option case as well as within the portfolio of possibilities. The optimal exercise prices, related to the present time, are listed in Table 1 for all considered scenarios. From the point of view of decision-makers, this type of information is essential for setting a price barrier in which a given investment opportunity is significantly valuable.

flexibility	E	C	A	E&C	E&A	C&A	E&C&A
optimal exercise price	66.0	17.75	28.25	17.25 (C) 66.0 (E)	24.75 (A) 72.0 (E)	28.25 (A)	24.75 (A) 72.0 (E)

Table 1 Optimal exercise prices relevant to the present time for various investment scenarios

In a brief conclusion, as the leitmotif of this contribution, it can be stated that the utilization of chooser options improves the evaluation of investment projects as well as the analysis of potential outcomes of various investment strategies and allow investors to select the most advantageous path with respect to the considered uncertainties.

Acknowledgements

Both authors were supported through the Czech Science Foundation (GAČR) under project 22-17028S. The support is greatly acknowledged. Furthermore, the second author also acknowledges the support provided within SP2024/003, a research project of VSB-TU Ostrava.

References

Dixit A, Pindyck R (1994) Investment Under Uncertainty. Princeton University Press, Princeton

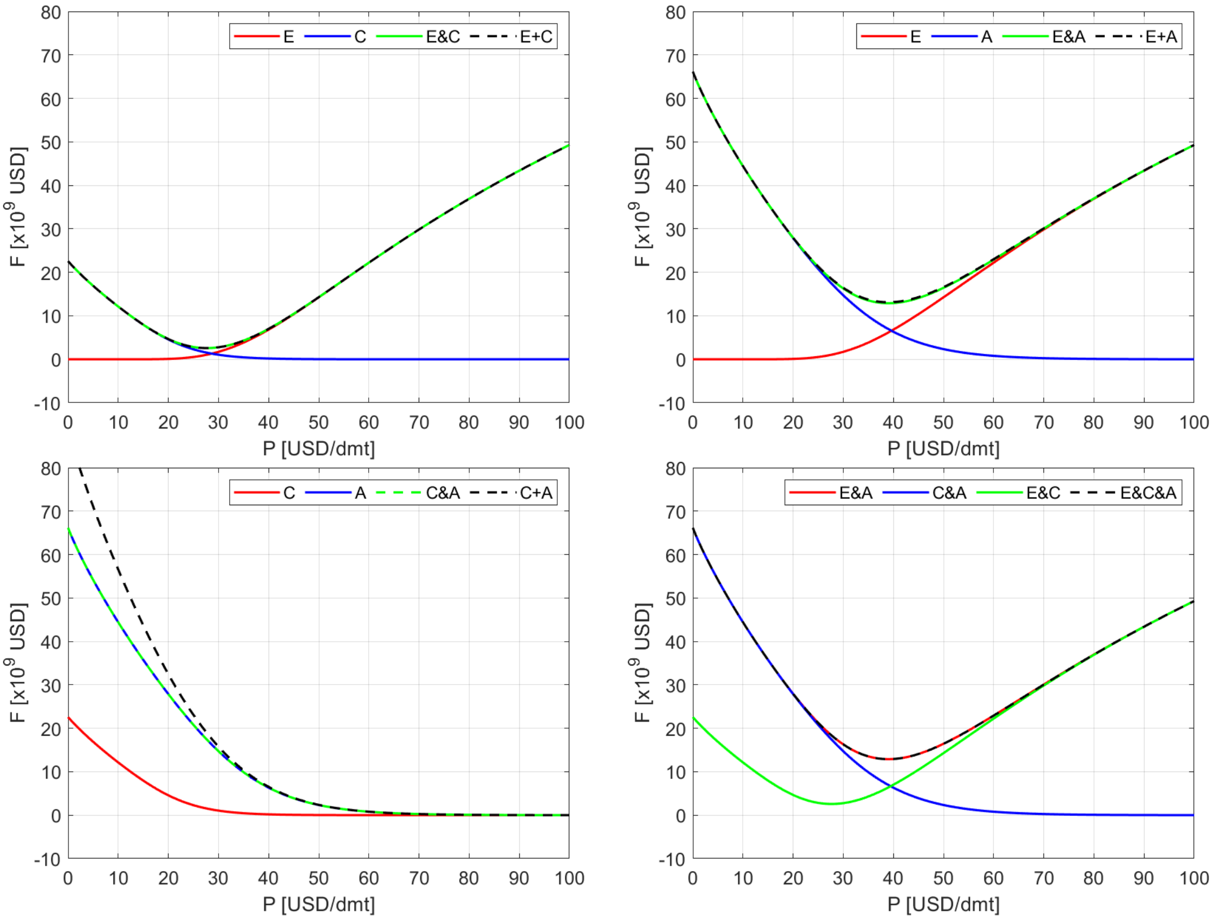


Figure 2 Flexibility values (under the European exercise right) at present time $t = 0$ for various scenarios: interaction between investment opportunities and comparison with the purely additive case

Haque M, Topal E, Lilford E (2014) A numerical study for a mining project using real options valuation under commodity price uncertainty. Resources Policy 39:115–123. `https://doi.org/10.1016/j.resourpol.2013.12.004`

Hecht F (2012) New development in FreeFem++. Journal of Numerical Mathematics 20:251–265. `https://doi.org/10.1515/jnum-2012-0013`

Hozman J, Tichý T (2020) The discontinuous Galerkin method for discretely observed Asian options. Mathematical Methods in the Applied Sciences 43:7726–7746. `https://doi.org/10.1002/mma.6160`

Hozman J, Tichý T (2021) Numerical Valuation of the Investment Project with Expansion Options Based on the PDE Approach. In: 39th International Conference on Mathematical Methods in Economics 2021. Conference Proceedings. Czech University of Life Sciences Prague, Prague, pp 185–190

Hozman J, Tichý T (2022) Numerical Valuation of the Investment Project Flexibility Based on the PDE Approach: An Option to Contract. In: 40th International Conference on Mathematical Methods in Economics 2022. Conference Proceedings. College of Polytechnics Jihlava, Jihlava, pp 122–128

Li N, Wang S (2019) Pricing options on investment project expansions under commodity price uncertainty. Journal of Industrial & Management Optimization 15:261–273. `https://doi.org/10.3934/jimo.2018042`

Rivière B (2008) Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation. SIAM, Philadelphia

Trigeorgis L (1993) The Nature of Option Interactions and the Valuation of Investments with Multiple Real Options. Journal of Financial and Quantitative Analysis 28:1–20. `https://doi.org/10.2307/2331148`

Zvan R, Forsyth PA, Vetzal KR (1998) Penalty methods for American options with stochastic volatility. Journal of Computational and Applied Mathematics 91:199–218. `https://doi.org/10.1016/S0377-0427(98)00037-5`

On-line Learning Process for Setting of Heuristic Parameters

Jaroslav Janáček[1], Marek Kvet[2]

Abstract. Tuning of sophisticated optimization heuristics represents a substantial part of the heuristic application and it decides on final success or fail of the application. Tuning of a heuristic is usually based on proper setting of heuristic parameters at such values, which ensure the most efficient run of the heuristic. The admissible values of the parameters are known in advance only in rear cases. Mostly, they must be determined for each individual case separately. It can be performed by previous research during the phase of heuristic tuning or by a self-learning process, which is a part of regular heuristic performance. Within this paper, an on-line learning process applied to swap heuristic parameter setting is studied. The swap heuristic is run in the frame of the gradual refinement process assigned to the problem of Pareto front approximation. The heuristic environment assures frequent repeating of the heuristic run and thus, the learning process may lead to significant results. The issue of the learning process may be either a recommendable parameter value or it can find that the parameter belongs to the class of sensitive parameters and no recommendable value exists.

Keywords: Location problems, Pareto front approximation, heuristics, online learning process

JEL Classification: C44, C61
AMS Classification: 90C05, 90C06, 90C10, 90C27

1 Introduction

Mathematical models of different location problem are frequently used as a support tool when making strategic decisions connected with public service system designing. Recently, there are several efficient solving approaches that enable us to obtain the optimal or at last suboptimal solution in a reasonable time window [1, 2, 3, 4, 5]. The problem may get extra ordinarily challenging if the model structure deviates from the usual standards. Another example of extra ordinary modeling and solving challenge consists in the necessity of optimizing two different quality criteria at the same time. In such a case, the Pareto front of feasible solutions is usually provided to the decision makers as a suitable output. Since the completion of the Pareto front requests for unpredictable computational time, plenty of heuristic approaches have been developed [8, 9, 10, 13].

Efficiency of many heuristic algorithms applied to optimization location problems depends on setting of parameters, which control run of the optimization process. In some cases, a good value of a parameter can be obtained by experiments performed with a given type of problem instances. Nevertheless, when the algorithm with the fixed parameter is applied to different instance of the problem, the obtained result may be very unsatisfactory. In this paper, we face the problem of finding a good approximation of the Pareto front of p-location problem solutions, where each solution is evaluated according to two objective functions. The approximating set of non-dominated solutions is subsequently created by applying the neighborhood search heuristics to solutions of the current approximation and adding the suitable inspected solutions to the approximation. The approach is characterized by repeated usage of the heuristics on very similar problems, which differ in the starting solution only and in slope of the comprehensive objective function, which is represented by the weighted sum of the two original objective functions. This situation represents a suitable environment for applying self-learning process to setting up the control parameters of the heuristic algorithm.

The remaining part of this study is structured as follows: The following chapter contains a description of the problem together with mathematical formulations of the objective functions. Furthermore, we discuss here the notion of Pareto front and its approximation. We also suggest a metric for evaluating the quality and accuracy of Pareto front approximation. In section 3, the main theoretical part of the paper is presented. We explain here the

[1] University of Žilina, Faculty of Management Science and Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovakia, jaroslav.janacek@fri.uniza.sk
[2] University of Žilina, Faculty of Management Science and Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovakia, marek.kvet@fri.uniza.sk

parametrized neighborhood search algorithm and gradual refinement approach. The fourth section describes th self-learning process. Section 5 summarizes the results of performed computational experiments and finally, the last section concludes presented research and suggests possible directions for future development.

2 The Problem Description

This paper examines the p-location problem, which involves contradicting systems and fair criteria. The problem can be expressed using the following notations. Let I represent the set of m applicants, denoted as $I = \{1, 2, \ldots, m\}$, from which a total of p elements will be selected to build a service center. Therefore, a straightforward solution to any p-location problem can be expressed in two distinct manners: either as a collection of indices from I that precisely includes p elements, or as a vector of location variables. The initial approach involved defining the set of all possible solutions. The second approach utilizing a vector is more appropriate for implementation and practical application, as it directly aligns with the mathematical model. The decision on service center locating at a location $i \in I$ is usually modelled by a binary variable $y_i \in \{0, 1\}$, which takes the value of one if a center should be located at i and it takes the value of zero in the opposite case. This way, the following expression (1) can define the set $\boldsymbol{Y}$ of all feasible solutions.

$$\boldsymbol{Y} = \left\{ \boldsymbol{y} \in \{0,1\}^m : \sum_{i=1}^{m} y_i = p \right\} \tag{1}$$

In regard to the system and fair criteria, their mathematical formulations necessitate the introduction of various novel symbols. The symbol J represents the collection of n network nodes, denoted as $J = \{1, 2, \ldots, n\}$, where the system users are grouped together. It is possible for the sets I and J to be identical, and there is no issue if the values of m and n equal. Each member j in the set J is associated with a weight coefficient b_j. The value of b_j can be interpreted in various ways, such as the number of users in the system who share location j or the predicted frequency of service demands at this node. Moreover, let the constant t_{ij} denote the time it takes to travel from the service center at position i to the user's location j. Based on the idea of generalized disutility as described in references [12], the system criterion assumes that a maximum of r centers placed closest to each client location can be involved in providing the service to clients. The term q_k in the criterion description represents the likelihood that the k-th nearest center is the closest one, which is currently reachable and free for service providing. The operation $min_k\{\}$ retrieves the k-th minimal value from the given list of values. Concrete values of mentioned coefficients can be found in [8, 9, 10, 11, 13]. Following these assumptions, the system criterion can be defined by (2).

$$f_1(\boldsymbol{y}) = \sum_{j \in J} b_j \sum_{k=1}^{r} q_k \min_k \{ t_{ij} : i \in I;\ y_i = 1 \} \tag{2}$$

The fair criterion (3) expresses the number of users, whose distance from the nearest located service center expressed by time exceeds the given limit T.

$$f_2(\boldsymbol{y}) = \sum_{j \in J} b_j \max \left\{ 0,\ sign\left(\min\{ t_{ij} : i = 1, \ldots, m; y_i = 1 \} - T \right) \right\} \tag{3}$$

No doubts that if a middle-sized or large instance of the associated location problem is solved, then the complete set of feasible solution cannot be processed due to its high cardinality. Nevertheless, the original large set can be presented by a special subset called *Pareto front*. It can be understood as a collection of solutions, in which the property of so-called *non-dominance* holds for each pair of its elements. [6, 7]. Thus, the Pareto front is formed by non-dominated solutions, which cannot be improved in one objective unless the other objective is worsened. As the process of the complete Pareto front obtaining is very demanding, we focus on heuristic search of a good approximation of the Pareto front for the bi-criterial problem. In other words, we seek for such set of non-dominated solutions *NDSS* that could serve as a good approximation of the former Pareto front. The approximating collection *NDSS* will be represented by a list of *noNDSS* solutions $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^{noNDSS}$ ordered according to increasing values of f_2. Here, the symbol *noNDSS* represents the cardinality of the *NDSS* set and it is assumed to be non-negative integer. To obtain a relevant approximation, the bordering solutions $\boldsymbol{y}^1$ and $\boldsymbol{y}^{noNDSS}$ must be determined to be very close to the most left and the most right solutions of the Pareto front as concerns the values of f_1 and f_2. Under these assumptions, the quality of the approximation *NDSS* can be measured by $A(NDSS)$ computed according to the expression (4).

$$A(NDSS)=\sum_{k=1}^{noNDSS-1}\left(f_1\left(\mathbf{y}^k\right)-f_1\left(\mathbf{y}^{noNDSS}\right)\right)\left(f_2\left(\mathbf{y}^{k+1}\right)-f_2\left(\mathbf{y}^k\right)\right) \quad (4)$$

3 Parametrized Neighborhood Search Algorithm and Gradual Refinement Approach

The *p*-location problem is formulated as a task of selection of *p* locations from *m* potential ones to minimize an objective function.

Swap operation performed with a solution *P* consists in selection of two locations $i \in P$ and $j \notin P$ and adjusting the solution *P* so that $\underline{P} = (P\text{-}\{i\})\cup\{j\}$. The resulting $\underline{P}$ is further denoted by *swap*(*P*, *i*, *j*). The set of all solutions, which can be obtained from *P* by exactly one swap operation, is called the neighborhood of *P*.

The standard swap algorithm starts with an input solution *P* and inspects the solutions from its neighborhood. If a better solution $\underline{P}$ is found, then substitution $P = \underline{P}$ is performed and the neighborhood search is repeated with new *P*. If no better solution is found, the swap algorithm terminates.

The determination of *P* can be performed according the generalized strategy depending on parameters *aTime*, *maxNos* and *thr*. The studied algorithm is described as follows.

NeighborhoodSearchAlgorithm(*P*, *aTime*, *maxNos, thr*)

0. Set *OK* = *true.*
1. If $CPU \le aTime$ and *OK*, then set $Nos = 0, f^* = f(P)$, $i^* = 0$, $P = \{i\in\{1, \ldots, m\}, y^c_i = 1\}$, $Q = \{1, \ldots, m\}$-*P* and continue with step 2. Otherwise return *P* and terminate.
2. While *Nos* < *maxNos* and *OK* perform the following commands for each pair (*i*, *j*), $i\in P$, $j \in Q$.
   If $f(P) - f(\text{swap}(P, i, j)) \ge thr$, then do $Nos = Nos+1$ and if $f(\text{swap}(P, i, j)) < f^*$, then set $f^* = f(swap(P, i, j))$, $i^* = i$, $j^* = j$.
3. If $i^* > 0$, then substitute $P = swap(P, i^*, j^*)$, else set *OK* = *false.*
   Continue with step 1.

This neighborhood search algorithm is used in the frame of approximation approach called the gradual refinement [8, 9, 10, 13].

4 Parametrized Neighborhood Search Algorithm and Gradual Refinement Approach

To avoid handmade setting of algorithm parameters, we embedded a self-learning procedure into the algorithm approximating the Pareto front. The procedure controls the parameter setting process based on information about previous successes or fails of the individual settings and evaluation of the last used setting in comparison with the last but one. Information on past of the process is condensed in state component values $E_1, \ldots, E_q$, where *q* is the number of parameters. If $E_i > 0$, then increasing of the parameter par_i is supported and in the opposite case, the parameter decreasing is recommended. Parameter par_i is changed randomly up or down by the value of $delt_i$ depending on the random trial with probability $prob_i$ in favor of increasing. Probability $prob_i$ is computed according to (**).

$$\begin{aligned} prob_i &= 0 \ if \ E_i < -1 \\ prob_i &= 1 \ if \ E_i > -1 \\ prob_i &= (1 + E_i)/2 \quad otherwise \end{aligned} \quad (5)$$

Let us consider that the last but one parameter setting denoted by $\boldsymbol{par}^{pr}$ cave the value of $f(\boldsymbol{par}^{pr})$ of the objective function *f*. The new setting obtained as the result of random trials and associated updating gave the objective function value $f(\boldsymbol{par}^{next})$. Then the sate value E_i is updated according to (6).

$$E_i = \alpha E_i + \beta sign\left(\left(f(\boldsymbol{par}^{pr}) - f(\boldsymbol{par}^{next})\right)\left(par_i^{next} - par_i^{pr}\right)\right) \quad (6)$$

Parameter α is known as parameter of forgetting and parameter β is called intensity of learning.

The self-learning procedure performs according to following steps.

0. {Initialization of the learning process}

Set the parameters $\boldsymbol{par}^{pr}$ at starting values and compute $f(\boldsymbol{par}^{pr})$ and set $E_i = 0$ for $i = 1, \ldots, q$. Define $\boldsymbol{par}^{next} = \boldsymbol{par}^{pr}$ and $f(\boldsymbol{par}^{next}) = f(\boldsymbol{par}^{pr})$

1. {State update and next parameter generating}

Update E_i according to (6) and determine $prob_i$ according to (5) for $i = 1, \ldots, q$. Set $\boldsymbol{par}^{next} = \boldsymbol{par}^{pr}$. Perform the random trials and obtain new $\boldsymbol{par}^{next}$ and compute $f(\boldsymbol{par}^{next})$.

2. {Termination rule}

If a given computational time elapses, then terminate, otherwise go to 1.

5 Numerical Experiments

The goal of the numerical experiments is to verify influence of self-learning way of parameter setting on quality or Pareto front approximation. The set parameters of the neighborhood search heuristics were *aTime*, *maxNos* and *thr* mentioned is section 3 together with algorithm description. Quality of the Pareto front approximation was evaluated by the resulting gap between area determined by the approximation and the exact Pareto front. *Gap* is given in percentage, and it can be calculated by the formula (7), where *A*(*NDSS*) stands for area determined by the approximation and *A*(*PF*) stands for area determined by exact Pareto front.

$$gap = 100 * \frac{A(NDSS) - A(PF)}{A(PF)} \tag{7}$$

The experiments were carried out on eight benchmarks derived from existing emergency medical aid systems in the individual higher territorial units - HTU of the Slovak Republic. The exact Pareto fronts were presented in [6, 7] and the corresponding characteristic are reported in Table 1, where each benchmark corresponds to one row. The first two columns denoted by m and p contain the problem size. The *NoS* stands for number of solutions forming the Pareto front (*PF*) and the *A*(*PF*) denotes associated *Area*. The list of problem instances contains the HTU of Bratislava (BA), Banská Bystrica (BB), Košice (KE), Nitra (NR), Prešov (PO), Trenčín (TN), Trnava (TT) and Žilina (ZA). In the used input data, all inhabited network nodes represent the set of possible candidate locations of service centers and the possible demand locations as well.

Table 1 Benchmarks and the exact Pareto fronts characteristics

HTU	*m*	*p*	*NoS*	*A*(*PF*)
BA	87	14	34	569039
BB	515	36	229	1002681
KE	460	32	262	1295594
NR	350	27	106	736846
PO	664	32	271	956103
TN	276	21	98	829155
TT	249	18	64	814351
ZA	315	29	97	407293

As concerns the objectives defining the location problem, the system objective is sum of expected traversing times between demand locations and the nearest available service center. The associated $r = 3$ probabilities q_k were $q_1 = 0.77$, $q_2 = 0.16$ and $q_3 = 100 - q_1 - q_2$. Parameter T used in the fair criterion was set to the value of 10 minutes [6, 7, 13].

To make the algorithms comparable, the whole approximation process was restricted to five minutes of computation time and then, the resulting *NDSS* sets were compared. Because the learning process depends on results of random trial, the computational process was run ten times and the average results were presented in Table 2.

Computational time of one gradual refinement applied on one instance was limited by 300 seconds.

The experiments reported in this study were performed on a common PC equipped with the Intel® Core™ i7-3610QM CPU@2.30 GHz processor and 8 GB RAM. The algorithms were implemented in Java programming language in the NetBeans IDE 8.2 software. The self-learning process described in section 4 was performed for parameters $\alpha = 0.8$ and $\beta = 0.2$.

We started with determination of the parameter ranges, *aTime* obtained its values from interval [0.01 s, 1.0 s], *maxNos* took integer values from interval [1, 100] and *thr* ranges in interval [0, V^{max}]. The value of V^{max} was determined proportionally to the area of rectangle determined by bordering solutions x^0 and x^{no} of Pareto front associated with the individual instance. The solution x^0 is the member of Pareto front with minimal value of the objective function f_2 and x^{no} is the member with maximal value of the objective function f_2. Then, the value V^{max} is determined according to (8).

$$V^{\max} = (f_2(\mathbf{x}^{no}) - f_2(\mathbf{x}^0))(f_1(\mathbf{x}^0) - f_1(\mathbf{x}^{no}))/100 \quad (8)$$

The starting values of the parameters were set at center of their ranges, and increment or decrement of the parameters were 0.05, 1 and $V^{max}/30$ respectively.

In Table 2, average *gaps* and computational times are reported in the first two columns. The next three columns denoted by *aTime*, *maxNos* and *thr* contain the resulting average values of these parameters at the end of the self-learning process. The last three columns denoted by E[0], E[1] and E[2] contain the final values of the state components connected with the individual parameters.

Table 2 Average results of numerical experiments

	gap	CT[s]	*aTime*	*maxNos*	*thr*	E[0]	E[1]	E[2]
ZA	0,73	66,6	0,560	31,1	0	0,005	-0,165	0
TT	0,00	10,7	0,405	41,2	0	0,035	-0,610	0
TN	0,73	32,8	0,325	34,5	0	-0,200	0,455	0
PO	0,14	524,3	0,645	60,7	0	0,030	0,280	0
NR	0,67	98,3	0,415	55,5	0	0,085	-0,530	0
KE	1,40	372,7	0,495	55,3	0	-0,135	0,545	0
BB	0,31	325,6	0,535	55,5	0	0,075	-0,355	0
BA	2,19	0,5	0,420	45,6	0	0,085	0,200	0

It can be noticed that the self-learning algorithm achieved excellent results as concerns the accuracy of the approximation and furthermore, it offers good starting values of the parameters *aTime*, *maxNos* and *thr*. The positive values of *thr*= were excluded from the range of recommended setting.

6 Conclusions

The research conducted in this work aimed to optimize the public service system by addressing two conflicting objectives that cannot be concurrently achieved. The Pareto front appears to provide an adequate foundation for responsible authorities to determine the deployment of service centers in such situations. The primary scientific objective of this work was to focus on the methodologies used for approximating the Pareto front. Due to the multiple ways in which the set of non-dominated system designs can be formed and the sensitivity of most common heuristics to different parameters, focus was given to the on-line learning process for determining heuristic parameters.

The fine-tuning of advanced optimization heuristics is a significant aspect of applying heuristics, as it ultimately determines the success or failure of the application. The optimization of a heuristic typically involves adjusting the heuristic parameters to certain values that guarantee the best effective execution of the heuristic. The permissible values of the parameters are only known in advance in rare instances. Most of the time, they have to be decided for every single case independently. It can be conducted by prior study in the heuristic tuning phase or through a self-learning process that is a component of regular heuristic performance. This research examined the application of an online learning method to adjust heuristic parameter settings. The swap heuristic was executed within the context of the iterative refining process applied to the problem of approximating the Pareto front. The heuristic environment ensures that the heuristic run is repeated frequently, which can lead to considerable results in the learning process. Based on achieved results we can coclude that the scientific goal was fulfilled at a good level of satisfaction.

Future study in the subject of bi-criteria optimization should focus on the advancement of alternative methods to achieve accurate Pareto front approximation. Another potential study area may explore the generalization of existing approaches and their adaptation for the design of multi-objective systems.

Acknowledgements

This work was supported by the following grants: VEGA 1/0077/22 "Innovative prediction methods for optimization of public service systems", and VEGA 1/0654/22 "Cost-effective design of combined charging infrastructure and efficient operation of electric vehicles in public transport in sustainable cities and regions". This paper was supported by the Slovak Research and Development Agency under the Contract no. APVV-19-0441.

References

[1] Arroyo, J. E. C., dos Santos, P. M., Soares, M. S. and Santos, A. G. (2010). A Multi-Objective Genetic Algorithm with Path Relinking for the p-Median Problem. In: Proceedings of the 12th Ibero-American Conference on Advances in Artificial Intelligence, 2010, pp. 70–79.

[2] Bertsimas, D., Farias, V. F., Trichakis, N. (2011). The Price of Fairness. In Operations Research, 59, 2011, pp. 17-31.

[3] Brotcorne, L, Laporte, G, Semet, F. (2003). Ambulance location and relocation models. Eur. Journal of Oper.Research, 147, pp. 451-463.

[4] Buzna, Ľ., Koháni, M., Janáček, J. (2013). Proportionally Fairer Public Service Systems Design. In: Communications - Scientific Letters of the University of Žilina 15(1), pp. 14-18.

[5] Current, J., Daskin, M. and Schilling, D. (2002). Discrete network location models, Drezner Z. et al. (ed) Facility location: Applications and theory, Springer, pp. 81-118.

[6] Grygar, D., Fabricius, R. (2019). An efficient adjustment of genetic algorithm for Pareto front determination. In: TRANSCOM 2019: conference proceedings, Amsterdam: Elsevier Science, pp. 1335-1342.

[7] Janáček, J., Fabricius, R. (2021). Public service system design with conflicting criteria. In: IEEE Access: practical innovations, open solutions, ISSN 2169-3536, Vol. 9, pp. 130665-130679.

[8] Janáček, J., Kvet, M. (2023). Adaptive Parameter Setting for Public Service System Design. In: Strategic management and its support by information systems 2023, in print

[9] Janáček, J., Kvet, M. (2021). Quality measure of Pareto front approximation. In: Quantitative methods in economics: Multiple criteria decision making 21: Proceedings of the international scientific conference, Bratislava: Letra Edu, pp. 95-103.

[10] Janáček, J., Kvet, M. (2022). Repeated refinement approach to Bi-objective p-location problems. In_: INES 2022: 26th IEEE International Conference on Intelligent Engineering Systems, New York: Institute of Electrical and Electronics Engineers, pp. 41-45.

[11] Jankovič, P. (2016). Calculating Reduction Coefficients for Optimization of Emergency Service System Using Microscopic Simulation Model. In: *17th International Symposium on Computational Intelligence and Informatics*, pp. 163-167.

[12] Kvet, M. (2014). Computational Study of Radial Approach to Public Service System Design with Generalized Utility. In: *Proceedings of International Conference DT 2014*, Žilina, Slovakia, pp. 198-208.

[13] Kvet, M., Janáček, J. (2023). Hyperheuristic as Tuning Tool of Generalized Swap Strategy. In: META 2023: The 9th International Conference on Metaheuristics and Nature Inspired Computing, 1.-4.11. 2023, in print

Almost Stochastic Dominance Analysis of Mean-Variance Efficient Portfolios

Jana Junová[1], Miloš Kopa[2]

Abstract.
Stochastic dominance is a tool that allows the comparison of random variables, representing the random returns of investments under very general assumptions. However, the generality of these assumptions can lead to situations where dominance between two random variables does not exist, even though the majority of investors evidently prefer one. For this reason, a relaxation of stochastic dominance called almost stochastic dominance was introduced. While the definition of almost first-order stochastic dominance is widely accepted, the definitions of almost second-order stochastic dominance (ASSD) vary.
This article aims to describe the different approaches to ASSD and analyze the relationships between them. Using data regarding the monthly returns of 49 industry representative portfolios, we find the mean-variance efficient portfolios and analyze their ASSD relationship to the minimum variance portfolio, employing and comparing different definitions of ASSD.
Keywords: stochastic dominance, almost stochastic dominance, portfolio optimization

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Stochastic dominance (SD) allows the comparison of random variables, which may represent unknown returns of investments. It allows the evaluation of investment options without detailed knowledge of individual investor's preferences based solely on the assumption that the investor's true utility function belongs to a specific class. In this context, a random variable X is said to dominate a random variable Y if X has a higher expected utility than Y for all utility functions from that class. The broadest class is composed of all non-decreasing utility functions, corresponding to first-order stochastic dominance (FSD), a concept introduced by Quirk and Saposnik (1962). Second-order stochastic dominance (SSD), originally studied in Hadar and Russell (1969), Hanoch and Levy (1969), Rothschild and Stiglitz (1970), and Rothschild and Stiglitz (1971), further assumes that investors are risk-averse. The corresponding class of utility functions includes all non-decreasing concave functions. Stochastic dominance is widely studied to this day, such as in Domínguez et al. (2021); Kopa et al. (2021); Vitali and Moriggia (2021); Kopa et al. (2024).

Leshno and Levy (2002) described that the classes of utility functions considered for both described orders of stochastic dominance might be unnecessarily wide, possibly making the stochastic rules too strict. They introduced almost stochastic dominance (ASD), which is based on narrower sets of utility functions. How much the set is narrowed or how much the SD is relaxed is described by a tolerance parameter ε. To achieve certain desired properties, further definitions of the almost second-order stochastic dominance (ASSD) were introduced in Tzeng et al. (2012) and Chang et al. (2019). A different approach to relax the SD rules called tractable almost stochastic dominance, was introduced in Lizyayev and Ruszczyński (2012).

Another way to study investments is the mean-risk analysis. Markowitz (1952) and Markowitz (1959) introduced portfolio optimization problems with respect to mean and variance. Malavasi et al. (2021) analyze the relationships between mean-variance and SSD optimality. We expand it by considering the almost second-order stochastic dominance. We analyze the size of tolerance parameters necessary to achieve ASSD for portfolios with different means and variances. We do this considering all of the above-stated approaches to ASSD.

The structure of this work is the following. We briefly introduce stochastic dominance and its properties in Section 2. Section 3 presents four different definitions of almost second-order stochastic dominance, and some of their

[1] Charles University/Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 49/83, 186 75 Praha 8, junova@karlin.mff.cuni.cz
[2] Charles University/Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 49/83, 186 75 Praha 8, kopa@karlin.mff.cuni.cz

properties. The empirical analysis of mean-variance efficient portfolios and their almost stochastic dominance relationships to the minimum variance portfolio is performed in Section 4. Section 5 summarizes the conclusions.

2 Stochastic dominance

Throughout this work, we consider random variables X and Y with the support in $[a, b]$. The distribution function of X will be denoted by F and the distribution function of Y by G. The first- and the second-order stochastic dominance is defined as follows.

Definition 1. X dominates Y by the first-order stochastic dominance ($X \succeq_{(1)} Y$) if $F(x) \leq G(x)$ for all $x \in [a, b]$. Define the integrated distribution functions $F^{(2)}(x) = \int_a^x F(t)dt$ and $G^{(2)}(x) = \int_a^x G(t)dt$. Then X dominates Y by the second-order stochastic dominance ($X \succeq_{(2)} Y$) if $F^{(2)}(x) \leq G^{(2)}(x)$ for all $x \in [a, b]$.

It is well known and proved, for example, in Levy (2006) that stochastic dominance has the following properties.

Proposition 1. *(i) $X \succeq_{(1)} Y$ ($X \succeq_{(2)} Y$) if and only if $\mathbb{E}\,u(X) \geq \mathbb{E}\,u(Y)$ for all utility functions $u \in U_1$ (U_2) where $U_1 = \{u : \mathbb{R} \to \mathbb{R}, u' \geq 0\}$ and $U_2 = \{u : \mathbb{R} \to \mathbb{R}, u' \geq 0\, u' \leq 0\}$.*

(ii) $X \succeq_{(1)} Y \implies X \succeq_{(2)} Y$.

Analogical properties are also desirable in the case of almost stochastic dominance, as will be discussed later.

3 Almost stochastic dominance

Denote $\| F - G \| = \int_a^b | F(t) - G(t) | \, dt$. Almost stochastic dominance was first introduced in Leshno and Levy (2002) and their definition of almost first-order and almost second-order SD is the following.

Definition 2. For $\varepsilon \in (0, 0.5)$, we say that:

(i) X dominates Y by ε-Almost FSD ($X \succeq_{(1)}^{\varepsilon} Y$) if

$$\int_{S_1} F(t) - G(t)dt \leq \varepsilon \, \| F - G \|,$$

where

$$S_1 = \{t \in [a, b] : G(t) < F(t)\}.$$

(ii) X dominates Y by ε-Almost SSD ($X \succeq_{(2)}^{\varepsilon} Y$) if

$$\int_{S_2} F(t) - G(t)dt \leq \varepsilon \, \| F - G \|,$$

where

$$S_2 = \{t \in S_1 : \int_a^t G(x)dx < \int_a^t F(x)dx\}, \text{ and } E(X) \geq E(Y).$$

We will denote the above-defined almost stochastic dominance as AFSD^{LL} and ASSD^{LL}.

Analogous properties to those described in Proposition 1 are also desired in the case of almost stochastic dominance. Following Guo et al. (2013), we call them expected utility maximization and hierarchy property.

Definition 3 (Expected utility maximization)**.** The AFSD (resp. ASSD) relationship $X \succeq_{(1)}^{\varepsilon} Y$ (resp. $X \succeq_{(2)}^{\varepsilon} Y$) satisfies the expected utility maximization property if

$$\mathbb{E}\,u(X) \geq \mathbb{E}\,u(Y) \; \forall u \in U_1^*(\varepsilon) (\text{resp. } U_2^*(\varepsilon))$$

where

$$U_1^*(\varepsilon) = \left\{ u \in U_1 : u'(x) \leq \inf_{y \in [a,b]} \{u'(y)\} \left[\frac{1}{\varepsilon} - 1 \right], \forall x \in [a, b] \right\}$$

and

$$U_2^*(\varepsilon) = \{u \in U_2 : -u''(x) \leq \inf_{y \in [a,b]} \{-u''(y)\} \left[\frac{1}{\varepsilon} - 1 \right] \forall x \in [a, b]\}.$$

Definition 4 (Hierarchy)**.** The AFSD and ASSD ($X \succeq^{\varepsilon}_{(1)} Y$ and $X \succeq^{\varepsilon}_{(2)} Y$) relationship satisfies the hierarchy property if

$$X \succeq^{\varepsilon}_{(1)} Y \Rightarrow X \succeq^{\varepsilon}_{(2)} Y.$$

Unlike the definition of AFSD$^{\mathrm{LL}}$, the definition of ASSD$^{\mathrm{LL}}$ was modified several times to find one that satisfies the hierarchy property, as well as the expected utility maximization property.

Tzeng et al. (2012) presented an example that shows that the ASSD$^{\mathrm{LL}}$ does not satisfy the expected utility maximization property and suggested the following definition, which we will call ASSD$^{\mathrm{Tz}}$. Denote $\| F^{(2)} - G^{(2)} \| = \int_a^b \mid F^{(2)}(t) - G^{(2)}(t) \mid dt$.

Definition 5. X dominates Y by ε-ASSD$^{\mathrm{Tz}}$ ($X \succeq^{\varepsilon,Tz}_{(2)} Y$) if

$$\int_{\hat{S}_2} F^{(2)}(t) - G^{(2)}(t) dt \leq \varepsilon \parallel F^{(2)} - G^{(2)} \parallel,$$

where

$$\hat{S}_2 = \{t \in [a,b] : \int_a^t G(x)dx < \int_a^t F(x)dx\}.$$

However, Guo et al. (2013) showed that the ASSD$^{\mathrm{Tz}}$ does not satisfy the hierarchy property, i.e. AFSD$^{\mathrm{LL}} \not\Rightarrow$ ASSD$^{\mathrm{Tz}}$.

A further modification of the definition of ASSD was introduced in Chang et al. (2019). They argue that it satisfies both the hierarchy and the expected utility maximization properties. However, it is also the most complicated one. We denote it ASSD$^{\mathrm{Ch}}$.

Definition 6. X dominates Y by ε-ASSD$^{\mathrm{Ch}}$ ($X \succeq^{\varepsilon,Ch}_{(2)} Y$) if

$$\int_{S_2} F(t) - G(t) dt \leq \varepsilon \Big(\int_{S_2} F(t) - G(t) dt + \int_{\overline{S_2}} G(t) - F(t) dt \Big),$$

where

$$\overline{S_2} = \{t \in [a,b] : \int_a^t G(x)dx \geq \int_a^t F(x)dx\}.$$

In addition to the three above-described approaches to ASD, there is also a completely different one, which was introduced in Lizyayev and Ruszczyński (2012) and is called tractable almost stochastic dominance. It is also a relaxation of stochastic dominance. However, the main motivation for this definition was the tractability of portfolio optimization problems with ASD constraints.

Definition 7. A random variable X ε-almost dominates a random variable Y by ε-AnSD$^{\mathrm{LR}}$ ($X \succeq^{\varepsilon,LR}_{(n)} Y$) if there exists a non-negative random variable Z such that $\mathbb{E}\, Z \leq \varepsilon$ and $X + Z \succeq_{(n)} Y$.

Junová and Kopa (2024) proved that the minimum necessary ε for ASSD$^{\mathrm{LR}}$ is equal to the general measure of second-order stochastic non-dominance using the first-order Wasserstein distance GND_1^2 and showed how it can be computed. This result was used for the computations in the empirical part.

4 Empirical results

We use data capturing monthly returns of 49 industry representative portfolios from the Fama/French data library. The data set spans from January 1970 until December 2023. We construct 21 mean-variance efficient portfolios. We denote their return $R_0, \dots R_{20}$ (the higher the index, the higher the mean return). A portfolio with minimum variance and returns R_0 is used as a benchmark. The remaining portfolios have higher means and variance. The average monthly mean return of the benchmark is 0.93 percent, and the average monthly return of R_{20} is 1.35 percent. The required returns of the remaining portfolios are spread equidistantly in-between.

None of the returns of the 20 mean-variance efficient portfolios, $R_1, \dots, R_{20}$, dominates the benchmark R_0 with respect to SSD. Nevertheless, they dominate it by any of the described ASSD for a sufficiently large tolerance parameter ε. We compute the minimum necessary ε's for R_i to dominate R_0 with respect to each type of ASSD. To do so, we use a numerical technique for computation of the integrals.

The results are presented in Table 1 where the minimum tolerance parameter needed to dominate R_0 by each of the considered types of ASSD[1] is shown. The ε's, which were the lowest according to each definition, are highlighted in bold.

	LL	Tz	Ch	LR
R_1	0.03	0.01	0.073	**0.004**
R_2	0.024	0.009	0.043	0.009
R_3	**0.02**	**0.007**	**0.031**	0.012
R_4	0.032	0.01	0.045	0.019
R_5	0.04	0.017	0.051	0.031
R_6	0.08	0.027	0.095	0.051
R_7	0.139	0.058	0.162	0.172
R_8	0.168	0.098	0.202	0.307
R_9	0.189	0.14	0.233	0.466
R_{10}	0.205	0.176	0.257	0.637
R_{11}	0.223	0.219	0.288	0.867
R_{12}	0.243	0.262	0.321	1.14
R_{13}	0.26	0.303	0.352	1.439
R_{14}	0.276	0.338	0.381	1.764
R_{15}	0.29	0.365	0.409	2.126
R_{16}	0.303	0.392	0.435	2.536
R_{17}	0.315	0.415	0.46	2.987
R_{18}	0.327	0.437	0.485	3.502
R_{19}	0.345	0.497	0.528	4.416
R_{20}	0.398	0.497	0.66	8.048

Table 1 Comparison of minimal ε's necessary for R_i to dominate R_0 with respect to each type of ASSD.

Figure 1 is based on the numbers presented in Table 1 and provides a better perspective of how the minimum necessary ε's evolved as the required mean of the mean-variance efficient portfolios increased.

In the case of the first three approaches presented here, the minimum necessary ε is a ratio of the amount of violation of SSD (which is measured in different ways) and the overall difference of the distributions (also measured in different ways). Both the numerator and denominator are very low numbers in this empirical study. Unlike the other approaches, ASSD^{LR} accounts only for the amount of violation of SSD (measured in another way). It is not divided by the very low number representing the overall difference in distributions, so the necessary ε's are much lower - the values are approximately 10 000 times lower for portfolios with low mean and low variance, and approximately 100 times lower for portfolios with high mean and high variance.

We can see from both Table 1 and Figure 1 that according to ASSD^{LR}, the tolerance parameter ε monotonically increases as the mean and variance increase, indicating increasing violation of SSD rules. On the other hand, according to the first three presented approaches by Leshno and Levy (2002), Tzeng et al. (2012) and Chang et al. (2019), the lowest tolerance parameter is not actually achieved for the portfolio with the lowest mean and variance but for R_3, which has slightly higher mean and variance. The greater violation of SSD rules is compensated for by the overall greater difference between the distributions of the benchmark and R_3. The minimum tolerance parameter then monotonically increases with increasing mean and variance of the returns of the portfolios in all of these three approaches. However, the slope of the increase differed between them. It can be seen in Figure 1 that the lines representing minimum necessary ε's for ASSD^{LL} and ASSD^{Tz} even intersect. The ε's necessary for ASSD^{Ch} are the greatest of these three in all cases

5 Conclusion

We have described four different approaches to ASSD and used them to analyze how far the returns of mean-variance efficient portfolios are from dominating the returns of benchmark (minimum variance portfolio) by SSD.

[1] The values of ε were multiplied by 1000 in case of ASSD^{LR} to be comparable to other kinds of ASSD.

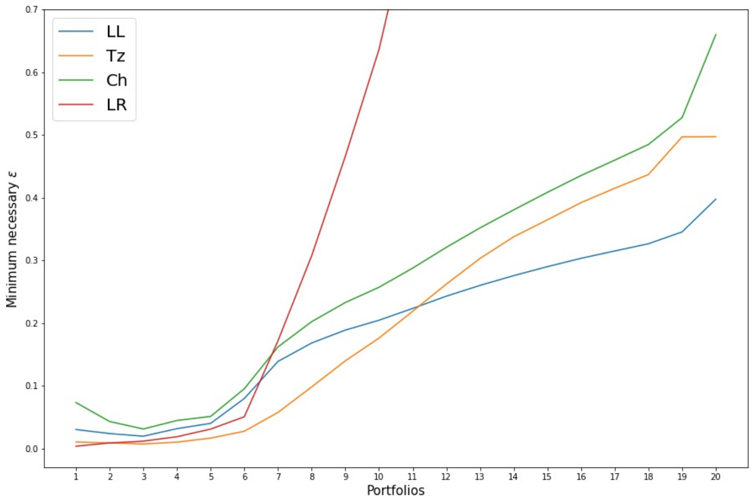


Figure 1 Minimum ε necessary in the compared types of ASSD.

ASSD^{LL} and ASSD^{Ch} compute the violation of SSD in the same way using the set S_2, and they both compare it to the overall difference in distribution functions, which they however measure differently. While ASSD^{LL} computes it by integrating the absolute value of the difference in the distribution functions, ASSD^{Ch} computes it by integrating the difference (not the absolute value) over two different sets. As a result, the overall difference measured by the approach of Chang et al. (2019) can be smaller, resulting in the need for greater ε's to achieve ASSD^{Ch}. This can be seen also in the results of our empirical study. The violation area in case of ASSD^{Tz} is computed using a wider set $\hat{S_2}$ ($S_2 \in \hat{S_2}$). They compute the size of the violation and the size of the overall difference using the integrated distribution functions. ASSD^{LR} accounts only for the violation of SSD by considering a non-negative random variable Z, whose addition to the originally considered random variable leads to domination of the benchmark by SSD.

Considering these definitions, we computed the lowest tolerance parameter ε that is necessary for each of the considered mean-variance optimal portfolios and each of these approaches to ASSD to almost dominate the benchmark. We showed that according to the approach by Lizyayev and Ruszczyński (2012), the necessary tolerance parameter increases as the variance increases, which complies to our expectations - the distribution of R_1 is closest to the distribution of the benchmark R_0. On the other hand, in the remaining three approaches, the lowest necessary tolerance parameters were achieved for the portfolio with slightly higher variance and mean. They account not only for the violation of SSD but also for the overall difference in distributions, showing that the greater absolute violation of SSD can be compensated for by the greater mean.

These findings could be further extended to the case of higher orders of almost stochastic dominance (Tzeng et al. (2012); Tsetlin et al. (2015); Chang et al. (2019)) or coherent risk measures used instead of variance (Artzner et al. (1999)). Moreover, the principles of almost stochastic dominance could be applied to multistage stochastic decision-making (Vitali et al. (2017)). Finally, these techniques could be employed even in the presence of endogenous randomness (Kopa and Rusý (2021, 2023)).

Acknowledgements

The research was supported by the Czech Science Foundation (Grant 22-10878K), GAUK (Grant number 190123), and the Martina Roeselová Memorial Fellowship granted by the IOCB Tech Foundation.

References

Artzner P., Delbaen F., Eber J., Heath D. (1999) Coherent measures of risk. Mathematical Finance, 9(3), 203–228

Chang J.-R., Lju W.-H., Hung M.-W. (2019) Revisiting generalized almost stochastic dominance. Annals of Operations Research, 281(1), 175–192

Domínguez R., Vitali S., Carrión M., Moriggia V. (2021) Analysing decarbonizing strategies in the european power system applying stochastic dominance constraints. Energy Economics, 101, 105438

Guo X., Zhu X., Wong W.-K., Zhu L. (2013) A note on almost stochastic dominance. Economics Letters, 121(2), 252–256

Hadar J., Russell W. R. (1969) Rules for ordering uncertain prospects. American Economic Review, 59(1), 25–34

Hanoch G., Levy H. (1969) The efficiency analysis of choices involving risk. Review of Economic Studies, 36(3), 335–346

Junová J., Kopa M. (2024) Measures of stochastic non-dominance in portfolio optimization. Submitted to Europan Journal of Operational Research

Kopa M., Kabašinskas A., Sutiene K. (2021) A stochastic dominance approach to pension-fund selection. IMA Journal of Management Mathematics, 33(1), 139–160

Kopa M., Moriggia V., Vitali S. (2024) Multistage stochastic dominance: an application to pension fund management. Annals of Operations Research. https://doi.org/10.1007/s10479-023-05658-y

Kopa M., Rusý T. (2021) A decision-dependent randomness stochastic program for asset–liability management model with a pricing decision. Annals of Operations Research, 299, 241–271

Kopa M., Rusý T. (2023) Robustness of stochastic programs with endogenous randomness via contamination. European Journal of Operational Research, 305(3), 1259–1272

Leshno M., Levy H. (2002) Preferred by "all" and preferred by "most" decision makers: Almost stochastic dominance. Management Science, 48(8), 1074–1085

Levy H. (2006) Stochastic Dominance: Investment Decision Making under Uncertainty. 2nd Edition. Springer

Lizyayev A., Ruszczyński A. (2012) Tractable almost stochastic dominance. European Journal of Operational Research, 218, 448–455

Malavasi M., Ortobelli Lozza S., Trück S. (2021) Second order of stochastic dominance efficiency vs mean variance efficiency. European Journal of Operational Research, 290(3), 1192–1206

Markowitz H. M. (1952) Portfolio selection. The Journal of Finance, 7(1), 77–91

Markowitz H. M. (1959) Portfolio Selection: Efficient Diversification of Investments. Yale University Press

Quirk J. P., Saposnik R. (1962) Admissibility and measurable utility functions. Review of Economic Studies, 29(2), 140–146

Rothschild M., Stiglitz J. (1970) Increasing risk: I. a definition. Journal of Economic Theory, 2(3), 225–243

Rothschild M., Stiglitz J. (1971) Increasing risk ii: Its economic consequences. Journal of Economic Theory, 3(1), 66–84

Tsetlin I., Winkler R. L., Huang R., Tzeng L. Y. (2015) Generalized almost stochastic dominance. Operations Research, 63(2), 363–377

Tzeng L. Y., Huang R. J., Shih P.-T. (2012) Revisiting almost second-degree stochastic dominance. Management Science, 59, 1250–1254

Vitali S., Moriggia V. (2021) Pension fund management with investment certificates and stochastic dominance. Annals of Operations Research, 299(1), 273–292

Vitali S., Moriggia V., Kopa M. (2017) Optimal pension fund composition for an italian private pension plan sponsor. Computational Management Science, 14, 135–160

A Comparative Analysis of SVAR and Traditional Filtering Methods in Output Gap Estimation

Dominik Kavřík[1]

Abstract. The Taylor curve, illustrating the trade-off between inflation variability and output gap fluctuations, is pivotal in shaping monetary policy decisions. Traditional analyses often utilize the Hodrick-Prescott filter to estimate the output gap, which may introduce distortions affecting policy interpretation. This paper proposes an alternative approach by employing the Structural Vector Autoregression (SVAR) model to filter the output gap and compares its effectiveness against standard methods like Hodrick-Prescott filter. This comparative analysis aims to uncover how different filtration methods influence the stability and accuracy of the Taylor curve estimates. The results suggest that the choice of filtering technique not only significantly alters the perceived efficacy of monetary policy but also necessitates a reassessment of methodological preferences in macroeconomic analysis. This study underscores the importance of selecting robust filtering tools in the empirical evaluation of key macroeconomic relationships.

Keywords: monetary policy, output gap, Taylor curve

JEL Classification: E31, E51, C13
AMS Classification: 91B62

1 Introduction

The Taylor curve, introduced by John B. Taylor (1979), illustrates the inverse relationship between the volatilities of the output gap and inflation, reflecting the efficiency of central bank monetary policy aimed at minimizing welfare loss. For an optimal monetary policy, this relationship would manifest as a negative correlation between these variances, serving as an indicator of sound policy (Olson et al., 2012). Understanding the significance of accurately estimating the output gap, this paper delves into various methodologies for this purpose.

Common methods traditionally employed include the Hodrick-Prescott (HP) filter, which has been widely used despite its controversial aspects, such as end-point bias and spurious cycles (Hamilton, 2018). Known for its atheoretical nature—lacking integration with macroeconomic theory—the HP filter has prompted the exploration of alternative methods such as band-pass filters like Christiano-Fitzgerald (CF) filter (Christiano & Fitzgerald, 2003), which target business cycle frequencies and provide a theoretically grounded approach.

Diving deeper into more robust methods, the structural vector autoregressive (SVAR) model with Blanchard-Quah decomposition is used to estimate the output gap. Offering a sophisticated technique within the SVAR framework, the Blanchard-Quah decomposition identifies structural shocks by imposing long-run restrictions based on economic theory. Understanding that demand shocks have no long-term effect on output levels, while supply shocks do, this methodology separates transitory components from permanent ones, providing a clearer picture of the output gap.

Critical to this analysis is the comparative evaluation of these methods. Extending this comparison will enhance the robustness of output gap estimation and improve the interpretability of monetary policy assessments.

2 Methodology

The goal of this study is to see the different implication of the monetary policy assessment using filtering methods for output gap that rely not only on statistical filtering but on economic theory. For this, the SVAR model is used, e.g. in [7] or [14], to obtain the output gap measure.

The primitive system assumed in this analysis is SVAR(1) of unemployment and output, as follows:

$$\boldsymbol{A}\begin{bmatrix}\Delta y_t\\ u_t\end{bmatrix} = \boldsymbol{\alpha} + \boldsymbol{\phi}\begin{bmatrix}u_{t-1}\\ y_{t-1}\end{bmatrix} + \begin{bmatrix}\epsilon_{1,t}\\ \epsilon_{2,t},\end{bmatrix} \tag{1}$$

[1] Prague University of Economics and Business, Department of Econometrics, Winston Churchill Square 4, 13067 Prague, Czech Republic, kavd00@vse.cz

where $\epsilon_{1,t} \sim$ i.i.d. $(0, \sigma_u^2)$ and $\epsilon_{2,t} \sim$ i.i.d. $(0, \sigma_y^2)$ are the pure innovations. By this structure, it is possible to account for a feedback or contemporaneous effect of the unemployment shock to GDP and vice versa. Matrix $\boldsymbol{A} = \begin{bmatrix} 1 & a_{12} \\ a_{21} & 1 \end{bmatrix}$, $\begin{bmatrix} \alpha_{0,1} \\ \alpha_{0,2}, \end{bmatrix}$, $\boldsymbol{\phi}$ are the structural parameters and $\begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t}, \end{bmatrix} \sim i.i.d(\mathbf{0}, \mathbf{\Sigma_\epsilon})$, where $\mathbf{\Sigma_\epsilon} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is the variance-covariance matrix. Following [2], if u_t and Δy_t do not have a unit-root, it is possible to rewrite the reduced form of (1), which is:

$$\begin{bmatrix} \Delta y_t \\ u_t \end{bmatrix} = \begin{bmatrix} c_1^* \\ c_2^* \end{bmatrix} + \begin{bmatrix} \phi_{11}^* & \phi_{12}^* \\ \phi_{21}^* & \phi_{22}^* \end{bmatrix} \begin{bmatrix} u_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix}, \tag{2}$$

into VMA(∞), that allows to place a long-run restriction on of the demand-side shock ϵ_1, t.

$$\begin{bmatrix} \Delta y_t \\ u_t \end{bmatrix} = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \Theta_i \boldsymbol{\epsilon}_{t-i}, \tag{3}$$

where the long-run restriciton is implemented as $\sum_{i=0}^{\infty} \Theta_i = \begin{bmatrix} 0 & \theta_{11}(i) \\ \theta_{21}(i) & \theta_{22}(i) \end{bmatrix}$, meaning that a demand shock does not affect the real GDP in the long-run. The output gap is then obtained by using estimated matrix $\mathbf{A}$ as follows: $\hat{\boldsymbol{y}} = y_0 + \sum_{i=0}^{\infty} \epsilon_{2,t-i}$. The resulting output gap is shown in the following chapter.

Since the goal of this study is to assess the impact of usage of theoretical SVAR model on the Time varying parameter (TVP) model result, such as the one in [17] or [12] the HP-filtered output gap will be estimated using the following formula based on [9] assuming it is possible to decompose real GDP into growth (g_t) and cycle (c_t) with a smoothing parameter λ, which is set to 1600. This filter solves the follwing minimization problem: $\min \left\{ \sum_{t-1}^{T} c_t^2 + \lambda \sum_{t=1}^{T} [\Delta g_t - \Delta g_{t-1}]^2 \right\}$.

Given that BQ output gap is available, next the conditional volatility of inflation and output gap for both filters are estimated. Inflation is additionally tested for unit root test robust towards breakpoint using zivot-andrews unit root test [20]. The conditional volatiltiy is estimated in two steps, first the mean equation is estimated by ARIMA(p,d,q) model with orders chosen by Bayesian information criterion (BIC). The residuals are then passed into GARCH model to obtain conditional volatility of inflation and both measures of output gap, using Blanchard-Quah decomposition and HP filter.

The estimated variances then enter the time-varying parameter (TVP) model to empirically test the time-varying Taylor Curve relationship. The TVP model defined as follows:

$$h_{\pi,t} = \beta_{0,\,t} + \beta_{1,\,t} h_{\tilde{y}\;t} + u_t \tag{4}$$

$$\boldsymbol{\beta}_{t+1} = \boldsymbol{I}\boldsymbol{\beta}_t + \boldsymbol{v}_{t+1} \tag{5}$$

where $\boldsymbol{v}_{t+1} \overset{i.i.d.}{\sim} \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$,, $h_{\pi,t}$. $h_{\tilde{y},t}$ are the conditional volatilities of inflation and output gap, respectively, and the evolution of $\beta_{1,t}$ is of particular interest because it measures the direction of a relationship between the output gap variance and inflation variance.

3 Data

Time series data of the real GDP, unemployment and Consumer Price Index (CPI) of the United Kingdom come from the FRED database. It is quarterly frequency and inflation and the output gap. The data set used in this study ranges from the first quarter of 1981 through the third quarter of 2018. The output gap is obtained using BQ SVAR model described above. For the resulting estimate of the output gap of United Kingdom, see the figure below.

Figure 1 BQ output gap of United kingdom.

Time series of inflation was tested for a unit root in the presence of a structural break using Zivot-Andrews [20] unit root test. The null hypothesis of a unit root was rejected on the 5% level of significance.

4 Results and Discussion

First, the estimated conditional volatilites of both output gaps are plotted below.

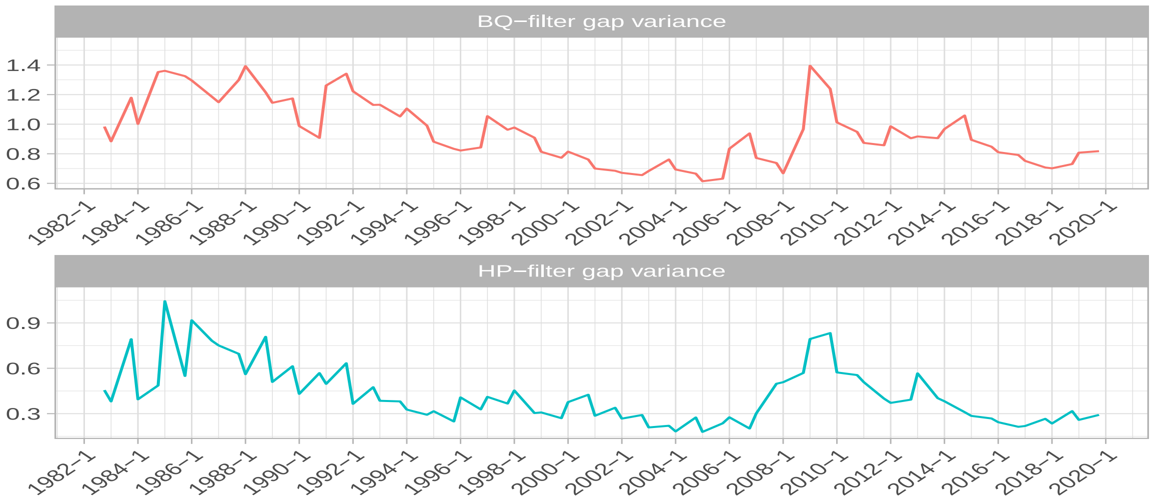


Figure 2 Comparison of conditional volatility of BQ output gap (above) and HP filtered output gap (below).

The model with time-varying parameters, presented in equations 4 and 5, is estimated via the Kalman filter as described by [11]. This study primarily focuses on the time-varying parameter $\beta_{1,t}$ because its absolute value has significant implications for the effectiveness of monetary policy. A negative β_1 suggests a trade-off relationship, implying monetary policy efficiency according to the Taylor curve theory [19]. This theory has been applied for this purpose in studies such as [17, 13, 4], and is also examined from a historical perspective in [16]. Therefore, $\beta_{1,t}$ is crucial for evaluating the sensitivity of Taylor curve estimations. The following figure depicts the estimated trajectory of this parameter $\beta_{1,t}$ as specified in equation 4.

The analysis of the time-varying parameter $\beta_{1,t}$ from the Taylor curve estimations for the UK provides insights into the effectiveness of monetary policy over various periods, using both the Blanchard-Quah (BQ) and Hodrick-Prescott (HP) filters to estimate the output gap. A negative β_1 indicates effective monetary policy, suggesting a beneficial trade-off between inflation volatility and output gap variability, while a positive β_1 implies inefficiency. From 1985 Q1 to 1987 Q2, both filters consistently showed negative β_1 values, reflecting effective monetary policy during a period of economic recovery following the early 1980s recession.

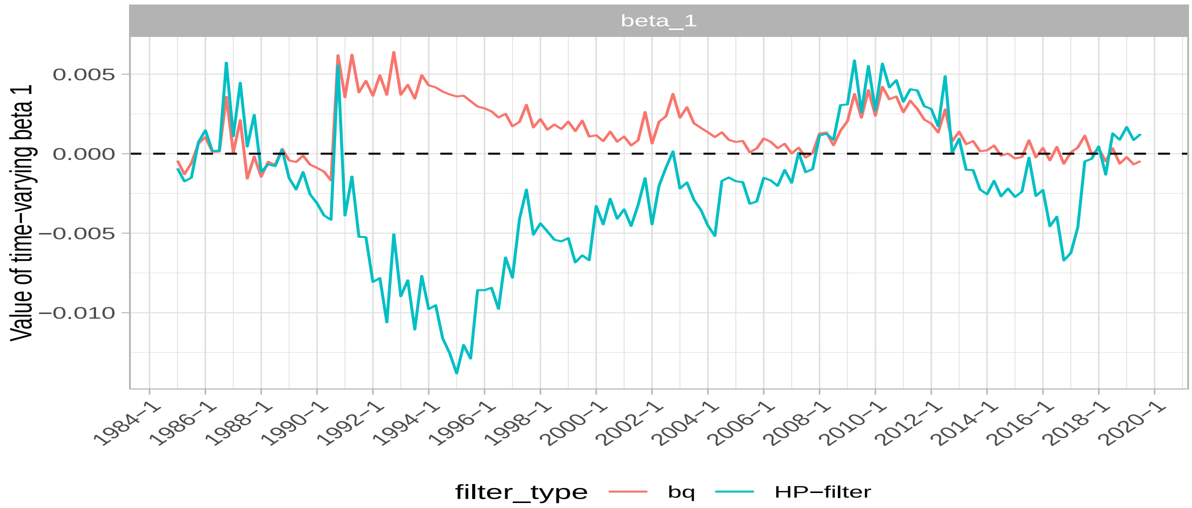


Figure 3 Time-varying parameter beta 1 obtained for both BQ SVAR output gap (red line) and HP filter output gap (blue line).

However, from late 1987 to 1988 Q4, there was notable divergence, with the BQ decomposition showing positive β_1 values, likely due to the economic impacts of the 1987 stock market crash. This crash, known as "Black Monday," occurred on October 19, 1987, and led to a sudden and severe global market decline, which significantly affected investor confidence and economic stability in the UK and worldwide [18]. The filters' differing responses to this shock highlight the sensitivity of monetary policy assessments to sudden economic disruptions. During the early 1990s recession (1990 Q1 - 1993 Q2), both filters again indicated policy effectiveness with negative β_1 values, aligning with the tight monetary policy measures of that time [10]. In contrast, the late 1990s (1997 Q1 - 2000 Q4) saw mostly negative β_1 values from the HP filter, suggesting continued policy efficiency, while the BQ decomposition occasionally showed positive values. This period was characterized by economic stability and growth [15].

Post-2008 financial crisis data (2008 Q1 - 2010 Q4) revealed divergence, with the BQ decomposition showing positive β_1, indicating inefficiencies in the face of severe economic disruptions. Finally, the recovery period from 2011 Q1 to 2019 Q3 saw both filters mostly yielding positive β_1 values, reflecting the challenges of managing the trade-off between inflation and output gap during a time of sustained low interest rates and unconventional monetary policies [3].

5 Conclusion

This study investigated the impact of different output gap filtration techniques on the estimation of the Taylor curve within a time-varying parameter model framework. By comparing the Hodrick-Prescott (HP) filter, and the Blanchard-Quah (BQ) decomposition, it was found that the choice of filtration method significantly influences the perceived effectiveness of monetary policy as measured by the Taylor curve.

The results indicate that while the HP filter has been traditionally favored for its simplicity, it suffers from several drawbacks, including potential end-point bias and lack of theoretical grounding. The BQ decomposition offer more theoretically robust alternatives that better align with economic theory and business cycle considerations. The BQ decomposition, in particular, provides a clearer distinction between temporary and permanent components of the output gap, which can lead to a more accurate assessment of monetary policy.

The analysis showed that the time-varying parameter $\beta_{1,t}$, which indicates the relationship between the variances of inflation and the output gap, varied significantly depending on the filter used. The HP filter tended to show a more stable policy environment, while the BQ decomposition highlighted periods of policy inefficiency, especially during economic disruptions like the 1987 stock market crash and the 2008 financial crisis. This suggests that relying solely on a single filtering method may provide an incomplete picture of monetary policy effectiveness.

Overall, the findings highlight the importance of selecting an appropriate filtering technique for output gap estimation. Using multiple methods can offer a more comprehensive understanding of the macroeconomic dynamics and the impact of monetary policy. Future research should continue to explore the implications of different filtering techniques to enhance the robustness and reliability of economic policy evaluations perhaps using simulations methods by generating counterfactuals using macroecnomic models.

Acknowledgements

This research was supported by the Internal Grant Agency of Prague University of Economics and Business under Project F4/52/2023.

References

[1] Beveridge, S. & Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary economics*, 7(2):151–174.

[2] Blanchard, O. J. & Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79(4):655–673.

[3] Bank of England (2013). *The UK's Economic Recovery: The Beginning of the End, or the End of the Beginning?*. Bank of England Publications.

[4] Chatterjee, S. (2002). The Taylor Curve and the Unemployment-Inflation Tradeoff. *Federal Reserve Bank of Philadelphia Business Review*, pages 26–33.

[5] Christiano, L. J. & Fitzgerald, T. J. (2003). The band pass filter. *International economic review*, 44(2):435–465.

[6] Cochrane, J. H. (1988). How big is the random walk in gnp? *Journal of political economy*, 96(5):893–920.

[7] Chen, Jiaqian and Górnicka, Lucyna. *Measuring Output Gap: Is It Worth Your Time?*. International Monetary Fund, 2020.

[8] Hamilton, J. D. (2018). Why you should never use the hodrick-prescott filter. *Review of Economics and Statistics*, 100(5):831–843.

[9] Hodrick R. J. & Prescott, E. C. "Postwar US business cycles: an empirical investigation," *Journal of Money, Credit, and Banking*, pp. 1–16, 1997, JSTOR.

[10] International Monetary Fund (1993). *World Economic Outlook: A Survey by the Staff of the International Monetary Fund*. IMF Publications.

[11] Kalman R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45.

[12] Kavřík D. "Sensitivity Analysis of Taylor Curve Estimation: A Comparison of GARCH and Stochastic Volatility Models," in *International Conference on Applied Economics and Business*, 2023, pp. 17–24, Springer.

[13] Lee, J. (1999). The inflation and output variability tradeoff: evidence from a GARCH model. *Economics Letters*, 62(1):63–67.

[14] Mitchell, James, Mazzi, Gian Luigi, and Moauro, Filippo. *Structural VAR based estimates of the euro area output gap: Theoretical considerations and empirical evidences*. In *28th Symposium of Forecasters, Nice, 17 June*, 2008.

[15] Organisation for Economic Co-operation and Development (2001). *OECD Economic Outlook No. 69*. OECD Publishing.

[16] Olson, E. & Enders, W. (2012). A Historical Analysis of the Taylor Curve. *Journal of Money, Credit, and Banking*, 44(7):1285–1299.

[17] Olson, E., Enders, W. &, Wohar M. E. (2012). An Empirical Investigation of the Taylor Curve. *Journal of Macroeconomics*, 34(2):380–390.

[18] Shiller, R. J. (1989). *Market Volatility*. MIT Press.

[19] Taylor, J. B. (1979). Estimation and control of a macroeconomic model with rational expectations. *Econometrica: Journal of the econometric society*, pages 1267–1286.

[20] Zivot, E. & Andrews, D. W. K. (2002). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of business & economic statistics*, 20(1):25–44.

Sentiment analysis in press releases about the Czech economy by institutional stakeholders

Nikola Kaspříková[1]

Abstract. The paper provides a natural language processing analysis based on press releases of three established institutional stakeholders in Czech economy who are traditionally getting coverage in Czech media. The press releases by the Czech Chamber of Commerce, Czech-Moravian Confederation of Trade Unions and by the Confederation of Industry of the Czech Republic are analysed in this paper. The aim of the paper is to learn what is the sentiment polarity in recent press releases of these institutions and if there is a major difference in sentiment between the institutions. The results of basic analysis and polarity sentiment scores are reported.
The results of a dictionary based sentiment analysis show that all three institutions produced documents of mixed polarity within the period selected for the analysis. Most documents have positive sentiment polarity, but there were documents with negative polarity too. The Czech-Moravian Confederation of Trade Unions produced press releases with the lowest mean polarity score, but there is no major difference in sentiment polarity score in documents with respect to the institution which has produced them.

Keywords: text mining, sentiment analysis, polarity score, Czech economy

JEL Classification: E60
AMS Classification: 91C99

1 Introduction

There are man stakeholders in the Czech economy, some of them are organized in institutions. There is a couple of well-established institutional stakeholders in the Czech economy. These are regularly getting coverage in the Czech media. The traditional institutional stakeholders in the Czech Republic include the Czech Chamber of Commerce (in Czech Hospodářská komora), Czech-Moravian Confederation of Trade Unions (in Czech Českomoravská konfederace odborových svazů) and the Confederation of Industry of the Czech Republic (in Czech Svaz průmyslu České Republiky).

The attitudes, goals and requirements of the stakeholders may be reflected in the communication of the representatives of these institutions. The aim of the research reported in this paper is to learn what is the sentiment polarity in recent communication of these institutions and if there is a major difference in the sentiment in communication of the institutions.

The press releases of the institution may provide a valuable source for the analysis of goals and attitudes of the subject. The tools for natural language processing (NLP) may serve for a quick automated, maybe even unbiased, formal analysis. The NLP methods include tools which may be used for automated keywords detection, sentiment polarity evaluation (to find out whether the text is positive of negative) or for the identification of the topics which are behind the documents. The NLP methods have been used many times in research in social sciences, see e. g. (DiMaggio et al, 2013) or (Fowobaje et al, 2022). For a detailed description of the methods used to analyse text see (Jurafsky and Martin, 2023), (Blei et al, 2003) and (Grün and Hornik, 2011).

The paper provides a natural language processing analysis based on recent press releases of major institutional stakeholders in Czech economy. The press releases by the Czech Chamber of Commerce, Czech-Moravian Confederation of Trade Unions and by the Confederation of Industry of the Czech Republic are analyzed in this paper. The results of basic analysis and polarity sentiment scores are reported.

The choice of the institutions to be included in the analysis has been driven by long enough history, so that established institutions are included. Another requirement was an economy-wide operation of the institution and sufficient coverage in the Czech media. The aim of the paper is to analyze recent communication, related to the same set of events, within rather short period. At the same time reasonably large collection of documents has to

[1] Prague University of Economics and Business, Department of Mathematics, nám. W. Churchilla 1938/4, Praha, Czech Republic, nikola.kasprikova@vse.cz

be collected. The collection of the documents has been produced using manual retrieval, no automated tools have been used. Six most recent press releases by the Czech Chamber of Commerce and ten most recent press releases by the Confederation of Industry of the Czech Republic released in March and April 2023 are included in the analysis. The Czech-Moravian Confederation of Trade Unions publishes its statements often in forms of videos, such are not included in this analysis. The five most recent press releases of the Czech-Moravian Confederation of Trade Unions published in the text form in April 2023 or earlier are included in the collection of documents for the analysis, this includes four documents published between January and April 2023 and one document published even in 2022. The complete collection of documents for the analysis consists of 21 documents obtained from the websites of the institutions of interest.

2 Most frequent nouns and co-occurrences

The most frequent nouns (lemmas) are shown in Table 1 in descending order for each of the sources

Trade Unions	Confederation of Industry	Chamber of Commerce
mzda (wage)	firma (company)	komora (chamber)
vláda (government)	průmysl (industry)	rok (year)
rok (year)	svaz (confederation)	firma (company)
zaměstnanec (employee)	doprava (transport)	zaměstnanec (employee)
inflace (inflation)	spolupráce (cooperation)	podnikatel (entrepreneur)
svaz (confederation)	návrh (proposal)	země (country)
práce (labour)	rok (year)	republika (republic)
energie (energy)	mise (mission)	práce (labour)
člověk (man)	sněmovna (chamber of deputies)	stát (state)
zvýšení (increase)	partner (partner)	průmysl (industry)

Table 1 Most frequent nouns (descending) by source

The co-occurrences of words, i. e. if a word is closely followed by another one in the text, may be even more informative then the frequencies of single words (or lemmas).

Trade Unions	Confederation of Industry	Chamber of Commerce
minimum wage	Confederation of Industry	Chamber of Commerce
trade union	industry transport	Czech Republic
real wage	Chamber of Deputies	Czech economy
challenge the government	Czech company	vicepresident chamber
Czech Republic	president confederation	Environment
pension reform	president industry	data box
real decline	Taiwan partner	president chamber

Table 2 Most frequent co-occurrences nouns, verbs and adjectives (descending) by source

Some co-occurrences in Table 2 reflect actual topics such as data box or Taiwan partner. The term Environment (in Czech a co-occurrence "životní prostředí") is interesting too. The co-occurrences found in the documents released by the Czech-Moravian Confederation of Trade Unions are visualized in Figure 1.

3 Sentiment polarity analysis

The sentiment analysis has been done using the `udpipe` (Wijffels, 2023) package for R statistical software (R Foundation for Statistical Computing, 2024). The Czech SubLex dictionary (Veselovská and Bojar, 2013) was used for the sentiment analysis.

All three institutions have produced both documents with positive sentiment polarity score and documents with negative polarity score, see Table 3 and Figure 2. The distribution of the sentiment polarity score does not seem much different, see also the plot of the probability density estimates for the sentiment polarity score by source in

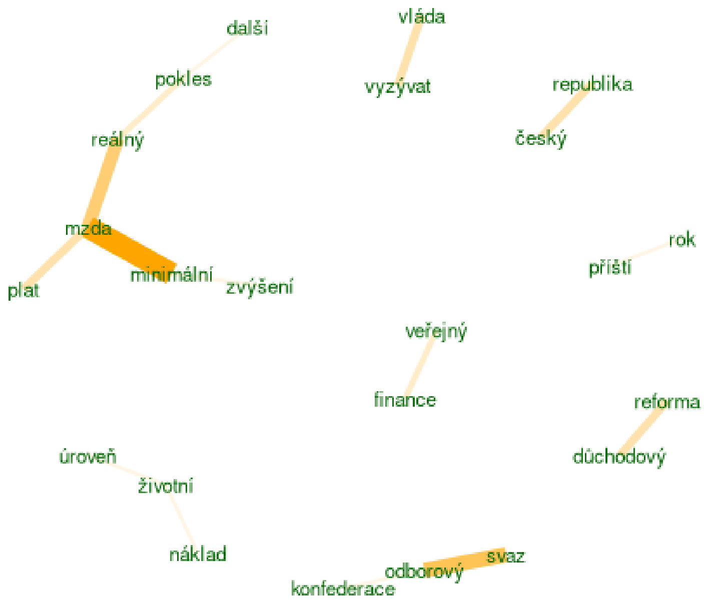


Figure 1 Co-occurrences in Trade Unions press releases

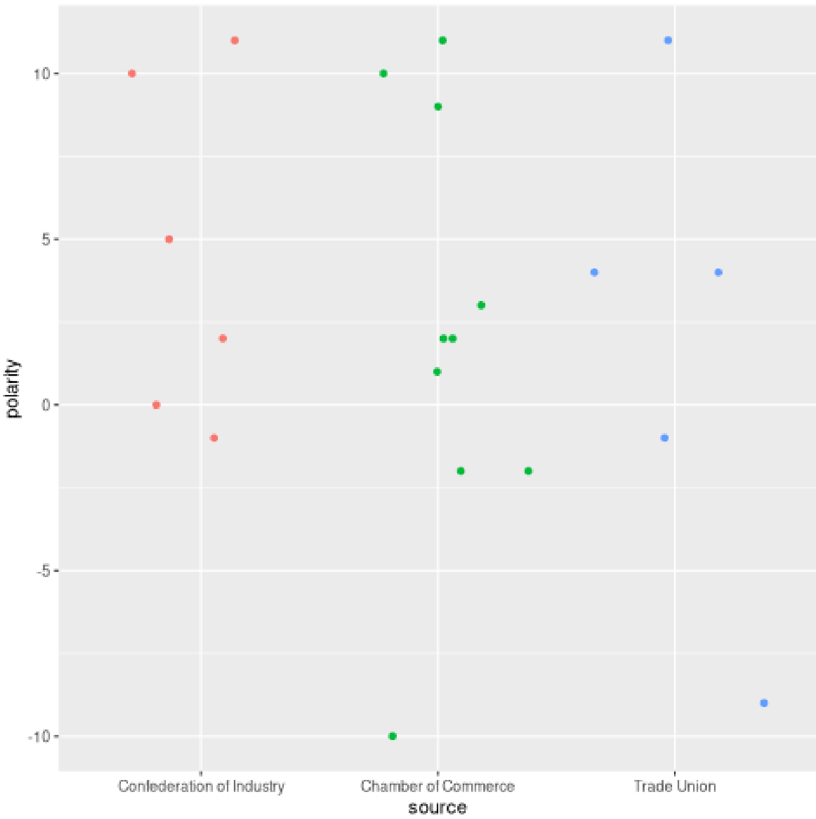


Figure 2 Document sentiment polarity score by source of document

Figure 3. The similarity of sentiment by source of documents may be a little surprising, considering the different topics of the communication, reflected in the most frequent terms used, see Table 1, Table 2 and Figure 1. The Czech-Moravian Confederation of Trade Unions press releases have the lowest mean sentiment polarity score, but overall the communication of this institution still has mostly positive polarity score, even if the messages are referring to terms such as inflation or further decline of real wages. This suggest that the institution presents mostly just a mild level of criticism of the government steps.

Source	Minimum	Median	Mean	Maximum
Confederation of Industry	-1	3.5	4.5	11
Chamber of Commerce	-10	2	2.4	11
Trade Union	-9	4	1.8	11

Table 3 Sentiment polarity score statistic by source

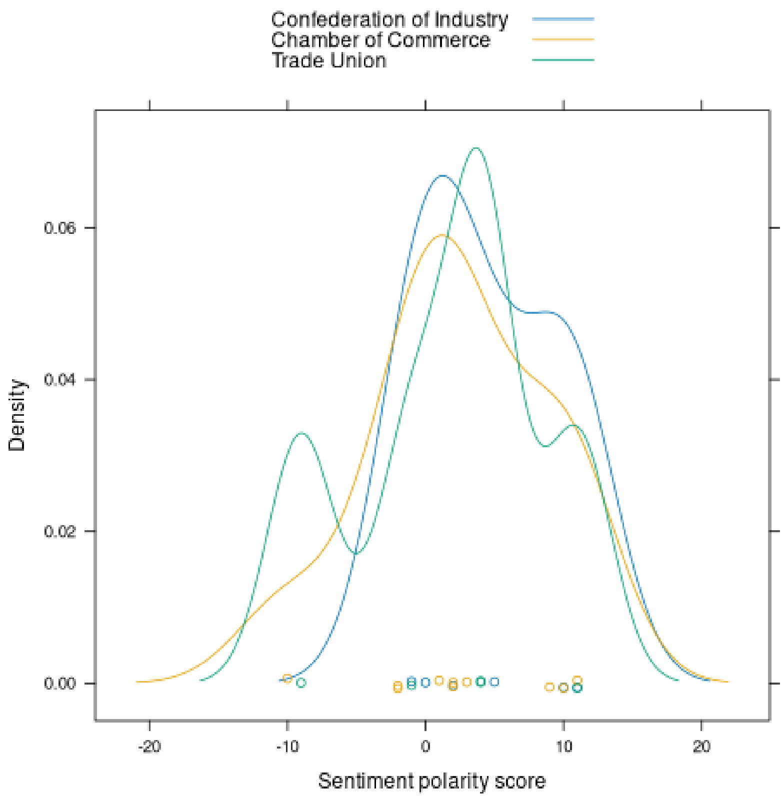


Figure 3 Probability density estimate for document sentiment polarity score by source of document

4 Conclusion

The results of a dictionary based sentiment analysis show that all three institutional stakeholders produced documents of mixed polarity within the period selected for the analysis. Most documents have positive sentiment polarity, but there were documents with negative polarity too. There is no clear difference in sentiment polarity score in documents with respect to the institution which has produced them. This result is a little surprising considering the results of frequency analysis of words and co-occurrences in the documents. The results of frequency analysis suggest that the core topics of communication of the Czech-Moravian Confederation of Trade Unions include labour and the remuneration for it, accompanied by the government and inflation. On the contrary the Confederation of Industry of the Czech Republic often reports on the companies and their cooperation with partners in its communication. The Czech Chamber of Commerce seems being the most general in the communication, referring to companies, employees, labour, industry and state. Nevertheless there have not been major differences between the three institutional stakeholders in the sentiment polarity scores, all three institutions have at least slightly positive mean polarity score. None of the institutions seems to produce clearly negative views on the current state of the Czech economy or present a sharp criticism of the government.

A detailed topic analysis may reveal more details about the documents and about possible differences in communication of the three institutions. It could be also beneficial to include the video press releases by the Czech-Moravian Confederation of Trade Unions in the analysis, after converting the video messages to text.

Acknowledgements

The paper has been produced using the institutional support for long term development at Prague University of Economics and Business.

References

Blei D, Ng A, Jordan M (2003) Latent dirichlet allocation. J Mach Learn Res 3:993–1022

DiMaggio P, Nag M, Blei D (2013) Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of u.s. government arts funding. Poetics 3:570–606. https://doi.org/10.1016/j.poetic.2013.08.004

Fowobaje KR, Mashood LO, Ekholuenetale M, et al (2022) Qualitative content analysis of nigerian heads-of-state and presidents' inaugural addresses: Text mining, topic modeling and sentiment analysis. SN Social Sciences 2

Grün B, Hornik K (2011) topicmodels: An r package for fitting topic models. Journal of Statistical Software 13. https://doi.org/10.18637/jss.v040.i13

Jurafsky D, Martin JH (2023) Speech and language processing: An introduction to natural language processing computational linguistics, and speech recognition. https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf

R Foundation for Statistical Computing (2024) R: A language and environment for statistical computing. https://www.R-project.org

Veselovská K, Bojar O (2013) Czech sublex 1.0, lindat/clariah-cz digital library at the institute of formal and applied linguistics (Úfal). https://hdl.handle.net/11858/00-097C-0000-0022-FF60-B

Wijffels J (2023) udpipe: Tokenization, parts of speech tagging, lemmatization and dependency parsing with the 'udpipe' 'nlp' toolkit . r package version 0.8.11. https://CRAN.R-project.org/package=udpipe

A Panel Analysis of the Economic Determinants of Military Spending

Michael Kejmar[1], Jiří Neubauer[2], Jakub Odehnal[3]

Abstract. The development of military spending in NATO countries is characterized by an increase in military spending caused mainly by the changing security situation in Europe. The security situation and the economic environment are considered as factors (determinants) influencing the size of military spending. The aim of the article is to present the possible use of the dynamic panel data model (GMM) to identify military spending determinants of selected NATO countries. To analyze the determinants of military spending of 23 countries, the following economic variables describing the economic, fiscal development of a country were selected: the size of government expenditures, the size of the government budget surplus (deficit), the size of the country's indebtedness, the economic development measured by GDP, and the size of government revenues. The results of the dynamic panel data model confirm the positive effect of government expenditures, GDP on military spending and negative effect of government debt on military spending. The results of the model confirm the expected hypotheses about the impact of selected economic variables on the military spending of 23 NATO countries in the period 1998-2022.

Keywords: GMM model, Economic Determinants, Military Spending
JEL Classification: C33, E69
AMS Classification: 62J05

1 Introduction

NATO military spending is one of the parts of public finances with turbulent developments nowadays. The economic environment can be observed as one of the significant factors influencing the size of military spending. Within NATO, the ambition of member countries is to spend at least 2% of gross domestic product on defense, which for a long period of time has been an unachievable target for most Alliance countries. The change in the perception of its own security can be observed mainly as a result of the NATO Summit in Wales in 2014. In the current period, a significant increase can be observed in military spending, especially by European countries. From an economic point of view, these countries can be characterized as suffering from a state budget deficit with increasing debt. This paper presents an analysis of the economic determinants of military spending. To analyze the link between the size of a country's military spending and its economic and fiscal characteristics, the following variables were selected: the size of government expenditures, the size of the government budget surplus (deficit), the size of the country's indebtedness, the economic development measured by GDP, and the size of government revenues.

2 Fiscal Determinants of Military Spending

The author (Nikolaidou, 2008) analyses the determinants of military spending in the form of economic, political and military factors on the examples of 15 European economies (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Portugal, Spain, Sweden, the Netherlands and the United Kingdom) in the period 1961 to 2005. The estimated ARDL model is based on an analysis of the link between military expenditures and selected variables including: GDP, population size, government expenditure size, trade balance, aggregate military expenditure size of European NATO member countries, US military expenditure, Greek military expenditure and Turkish military expenditure. In terms of the economic determinants analyzed, the findings confirm the positive impact of GDP on the size of military expenditure in the case of Greece, Portugal, the Netherlands, the United Kingdom, Italy, Spain and Austria. The relationship between military spending and debt size has been analyzed by the authors (Khan et al., 2021) who identify the conclusion that an increase in military spending leads to an increase in external debt for the 35 economies analyzed. Similarly, the authors

[1] University of Defence, Brno, Department of Resource Management, Kounicova 65, Brno, Czech Republic, michael.kejmar@unob.cz
[2] University of Defence, Brno, Department of Quantitative Methods, Kounicova 65, Brno, Czech Republic, jiri.neubauer@unob.cz
[3] University of Defence, Brno, Department of Resource Management, Kounicova 65, Brno, Czech Republic, jakub.odehnal@unob.cz

(Dudzevičiūtė et al, 2021) examined the relationship between the size of military spending and debt levels for selected European Union countries over the period 2005 to 2019. The results of the model did not reveal a consensus in the conclusions about the existence of a link between the analyzed variables, where in the case of Luxembourg, Latvia and Denmark the link was not confirmed, in the case of Lithuania, Slovenia and Slovakia a negative relationship was revealed, i.e. that at the same time as military spending increased, indebtedness decreased, and only in the case of Estonia the authors revealed a positive relationship between the size of military spending and indebtedness. Economic determinants as a factory influencing the size of military spending have been further investigated by authors such as (Dedebek and Meriç, 2015; Özsoy, 2008; Eita and Mbazima, 2008; Iiyambo and Kaulihowa, 2020; Odehnal and Neubauer, 2020). The economic variables used in each study are shown in Table 1.

authors	**economic determinants**
Nikolaidou, 2008	GDP, size of government spending, balance of payments, population
Khan et al., 2019	external debt, GDP growth rate
Dudzevičiūtė, 2021	debt
Dedebek and Meriç, 2015	government revenue, government expenditure, external debt
Özsoy, 2008	non-military government expenditure, government budget deficit
Eita a Mbazima, 2008	government revenue
Iiyambo and Kaulihowa, 2020	government revenue, debt
Odehnal and Neubauer, 2020	risk of Budget Balance, risk of Foreign Debt, risk of Inflation, risk of GDP per Capita, risk of GDP Growth

Table 1 Economic determinants of military spending

3 Data and Methods

To analyze the relationship between the size of military spending of 23 NATO countries (Czech Republic, Slovakia, Hungary, Poland, Germany, France, Belgium, Denmark, Netherlands, Luxembourg, Norway, Finland, Lithuania, Latvia, Estonia, Greece, Croatia, Bulgaria, Romania, Slovenia, Italy, Spain and Portugal) and economic determinants (the size of government expenditures, the size of the government budget surplus (deficit), the size of the country's indebtedness, the economic development measured by GDP, and the size of government revenues) were selected variables from the SIPRI database (military spending, billions Euro, constant prices) and the Eurostat database (economic variables, billions Euro). The development of the average values of military spending of the countries analyzed between 1998 and 2022 is shown in Figure 1.

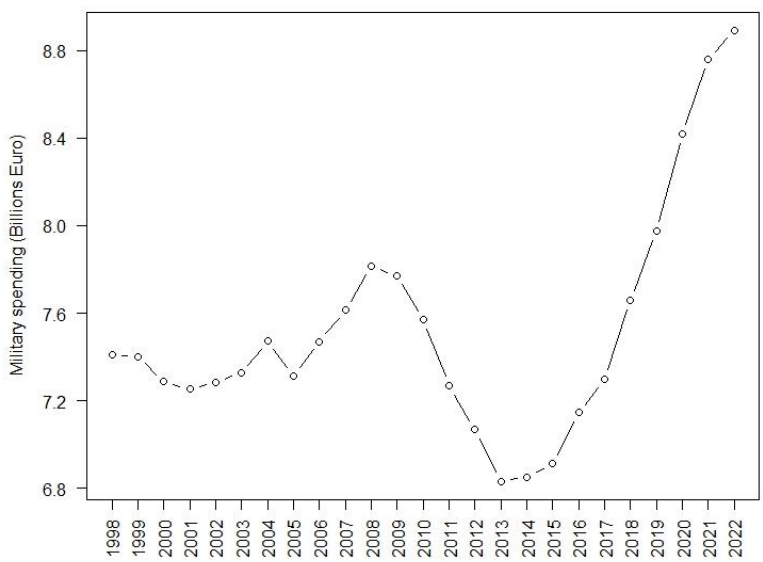


Figure 1 Average military spending selected NATO countries

The evolution of the average military spending of the selected NATO countries shows a significant increase in military spending after 2014, mainly due to the change in the security environment manifested by a significant increase in military spending, especially in European countries (Poland, Lithuania, Latvia, Estonia). However, the increasing trend is evident in most of the NATO countries with the goal of achieving a political consensus to spend at least 2 percent of the country's gross domestic product on defense. At the same time, the figure shows the effect of the economic crisis resulting in a significant drop in military spending after 2009. Thus, the development of a country's GDP was as one of the economic determinants of military spending by the alliance countries. In general, an increase in military expenditures can be expected in periods of economic growth, accompanied by increased in government revenues. In terms of fiscal determinants, a decline in military expenditures can be expected, especially during periods of increasing government deficits and economic indebtedness. However, in terms of the current evolution of military expenditure in NATO countries, it should be emphasized that its current evolution is mainly influenced by the change in the security environment, which is a different situation compared to, for example, the post-2009 period, when military spending was mainly influenced by the evolution of the economic environment.

For the purpose of analyzing the economic determinants of military spending, we use a dynamic model for panel data. A basic dynamic linear panel data model can be written in the form

$$y_{it} = \rho y_{i,t-1} + \beta' X_{it} + \mu_i + \varepsilon_{it}, \tag{1}$$

where $i = 1,2,\dots,n$ is the individual index (group, country, …), $t = 1,2,\dots,T$ is the time index and ε_{it} is an error term, X_{it} is a $k \times 1$ vector of regressors, β_{it} is a $k \times 1$ vector of parameters. The model is first differenced to get rid of the individual effect. First differencing (1) yields

$$\Delta y_{it} = \rho \Delta y_{i,t-1} + \beta' \Delta X_{it} + \Delta\varepsilon_{it}. \tag{2}$$

The error term $\Delta\varepsilon_{it}$ is autocorrelated and also correlated with lagged dependent variable $\Delta y_{i,t-1}$. Generalized method of moments (GMM) approach is used to get estimates of equation (2), see Arellano and Bond (1991).

4 Results

The results of the dynamic panel data model describing the link between military spending and selected economic determinants are shown in Table 2. The results show the relationship between military spending and the size of military spending of the previous year (lagged value), the size of government expenditures, the size of the country's indebtedness and the size of the government budget surplus (deficit). A two-step GMM approach was employed to estimate the parameters of (2). We used R software 4.3.2, package plm 2.6-4. The robust inference was performed. Autocorrelation of residuals was tested for the lag 1 and 2, p-values of the autocorrelation test of errors are 0.074 for lag 1 and 0.859 for lag 2. The statistical insignificance of government revenues may be due to the high correlation with GDP.

variables	**coefficient**	**S.E.**	**p-value**
military spending (t–1)	0.6683	0.0659	<0.001
government expenditures	0.0085	0.0036	0.020
country's indebtedness	−0.0017	0.0005	<0.001
government budget surplus (deficit)	−0.0008	0.0014	0.580
GDP	0.0023	0.0014	0.098
government revenues	0.0003	0.0063	0.951

Table 2 Parameter estimates of GMM model – all variables

Table 3 describes the results of the model after reduction of the variables confirming the identical relationships between the analyzed variables respecting at the same time the expected links between the development of the economic environment and the size of military spending. P-values of the autocorrelation test of errors are 0.074 for lag 1 and 0.874 for lag 2. According the Sargan test, the instrument are valid. Test results are consistent with the model assumptions.

variables	coefficient	S.E.	p-value
military spending (t–1)	0.6723	0.0708	<0.001
government expenditures	0.0088	0.0020	<0.001
country's indebtedness	−0.0016	0.0005	<0.001
GDP	0.0022	0.0013	0.086

Table 3 Parameter estimates of GMM model – final

In terms of expected links, the model confirms the expected relationship for all statistically significant variables. In the case of government expenditures, there is an increase in all analyzed countries (except Greece, where there is a decline after 2012). In the case of the indebtedness, it can be seen that most analyzed NATO countries have been facing a long-term government budget debts. The increase is evident as a consequence of the economic crisis after 2009 and as a consequence of the pandemic crisis after 2020. Thus, the significant increase in military spending evident especially after 2014 is one of the factors increasing the size of government spending, which, if financed in a period of government budget deficit, also appears as a factor increasing the level of indebtedness. In the case of the analysis of the relationship between GDP and the size of military expenditures, it is evident that the growth of the product generates state budget revenues that allow the growth of military expenditures.

5 Conclusions

Military spending is one of the parts of government spending that has seen a significant increase in NATO countries, mainly due to the change in the security environment after 2014. In general, factors influencing the size of military spending include the security environment and the evolution of the economic environment. The economic determinants of military spending include, in particular, the development of a country's GDP, where an increase in military spending can be expected in a period characterized by economic growth. Thus, the decline in military expenditure of NATO countries was particularly evident after 2009 as a consequence of the economic crisis leading to a significant decrease in military expenditure of most NATO countries. A consequence of the decline in military spending was that only a small group of member countries were able to meet their political commitment to spend 2% of GDP on defense. Beyond GDP, fiscal variables such as the government budget deficit can be considered as economic determinants. Other fiscal determinants can be considered, e.g. the indebtedness of the economy, where an increase in military spending can act as one of the factors increasing the indebtedness of the economy.

The paper analyses the relationship between selected economic determinants of military spending (the size of government expenditures, the size of the government budget surplus (deficit), the size of the country's indebtedness, the economic development measured by GDP, and the size of government revenues) and the size of military spending in selected NATO countries. The results of the dynamic panel data model show that for the 23 NATO countries analyzed in the period 1998–2002, military spending was mainly influenced by the size of the previous year's military spending, the size of government spending, level of national GDP and indebtedness. Thus, the results confirmed the assumptions that military spending can be expected to increase in the case of increasing GDP. In the case of increasing indebtedness, military expenditure would decrease.

However, the current development of military expenditure of NATO countries is mainly influenced by security determinants, which is evident especially from the significant increase in military expenditure of European NATO countries.

Acknowledgements

The paper was supported by institutional funding aimed at the development of the research organization (Long-Term Organization Development Goals at the FML).

References

Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economic Studies 58: 277–297.

Dedebek E, Meric M (2015) Avrupa Birliği ülkelerinde savunma harcamalarını belirleyen ekonomik ve sosyal faktörler: Panel veri analizi. Sayıştay Dergisi. 97:89-104. https://dergipark.org.tr/en/pub/sayistay/issue/61552/919188.

Dudzevičiūtė G, Česnuitytė, V, Prakapienė, D (2021) Defence Expenditure–Government Debt Nexus in the Context of Sustainability in Selected Small European Union Countries. Sustainability, 12. https://doi.org/10.3390/su13126669.

Eita J H, Mbazima D (2008) The causal relationship between government revenue and expenditure in Namibia. Journal of Economic and Financial Science 2:175-186. https://journals.co.za/doi/abs/10.10520/EJC139131.

EUROSTAT (2024) General and Regional Statistics. https://ec.europa.eu/eurostat/data/database.

Iiyambo H, Kaulihowa T (2020) An assessment of the relationship between public debt, government expenditure and revenue in Namibia. Public Sector Economics. 3:331-353. https://doi.org/10.3326/pse.44.3.3.

Khan L, Arif I, Waqar S (2021) The Impact of Military Expenditure on External Debt: The Case of 35 Arms Importing Countries. Defence and Peace Economics 5:588-599. https://doi.org/10.1080/10242694.2020.1723239.

Nikolaidou E (2005). The demand for military expenditure: evidence from the eu15 (1961–2005). Defence and Peace Economics 4:273-292. https://doi.org/10.1080/10242690802166533.

Odehnal J, Neubauer J (2020). Economic, Security, and Political Determinants of Military Spending in NATO Countries. Defence and Peace Economics 5:517-531. https://doi.org/10.1080/10242694.2018.1544440.

Özsoy O (2008) Government budget deficits, defence expenditure and income distribution: the case of Turkey. Defence and Peace Economics 1:61-75.https://doi.org/10.1080/10242690701347689.

SIPRI (2024) Military Expenditure Database. https://www.sipri.org/databases/milex.

Regional Intensity of the Freight: Functional Analysis of Variance

Peter Knížat[1]

Abstract. New digital data sources provide a great opportunity for econometrists to study the application of more complex statistical models, of which the corresponding empirical results can lead to prompt decision making since its early availability. In this paper, we introduce a theoretical framework of analysis of variance that is extended to a functional space. An estimation procedure uses a modified method of least squares that minimizes a set of mathematical objects in its criterion. The estimated functional parameters are observed on a continuous domain rather than discrete pointwise estimates as its classical counterparts. In the empirical analysis, we use a dataset of electronic records that is collected from the satellite-based toll system in Slovakia. Each record refers to a passage of the vehicle through a section of the monitored road. We aggregate data into weekly time series, i.e. a number of passing vehicles per week for each district. The weekly time series are transposed into a functional space through an expansion by basis splines. The observed mathematical objects that correspond to each district are categorized within a particular region and its co-variability is analysed within the concept of functional analysis of variance. The main objective of the paper is to carry out the assessment of the intensity of the regional freight in Slovakia. The empirical results show different seasonal patterns of variations and significant differences of the freight intensity between regions.

Keywords: regional freight, analysis of variance, functional parameters

JEL Classification: C49
AMS Classification: 62P20, 62R10

1 Introduction to Econometric Problem

In this paper, an econometric problem stems from using a new digital data source which can be used for flash estimates of the freight intensity or the economic activity. An electronic toll system, which is based on satellite technology, is in operation to collect toll from vehicles that use satellite-covered sections of the Slovak road network[2]. Data are collected in an automated way as an inherent function of on-board units, installed in vehicles[3].

The research related to study German toll data for nowcasting of the industrial economic performance and the identification of economic cycles is carried out by [3]. The paper states that there is a close relationship between the economic activity and the freight since the economic activity requires transport services. In Slovakia, the first study of toll data in the context of forecasting the macroeconomic output was done by [6]. The empirical results show that toll data have a high potential to capture the overall macroeconomic trend.

The main objective of this paper is to study the regional intensity of the freight during the year 2020 in Slovakia. The results of the empirical research can be useful for the management of the road infrastructure and the control of the freight during the specific time periods. Moreover, the paper aims to encourage further research that would conceptualise more complex econometric models for the early estimation of the regional economic activity.

In the empirical analysis, we use the functional analysis of variance, where the estimated parameters are functions that enable the assessment of the intensity of the freight across the entire observed time period, e.g. for analysis of seasonal variations. The research of the application of statistical models on functional data is abundant and the state-of-the-art manuscripts that cover both the theory and the practice are [4] and [5], respectively. The

---

[1] University of Economics in Bratislava, Department of Operational Research and Econometrics, Dolnozemská cesta 1, 852 35 Bratislava, peter.knizat@euba.sk. The author is a PhD candidate.

[2] Skytoll, electronic collection of toll, Slovak Republic. Available at: https://www.skytoll.com/elektronicky-mytny-system-sr/.

[3] In the Slovak Republic, all vehicles with the total weight of over 3.5 tons or vehicles with the total weight of over 3.5 tons listed in § 4 par. 2 letters b) and c) of Act no. 106/2018 Coll. on the operation of vehicles in road traffic (vehicles of category M and N) except for motor vehicles of category M1 and except for vehicle sets consisting of motor vehicles of category M1 and N1, are obligated to install the on-board unit.

most recent advances in functional data analysis are discussed in [1]. To the author's knowledge, no research of toll data within the concept of the functional space exist.

1.1 Data

An observed dataset contains electronic records for all vehicles that passed through monitored sections of roads during the year 2020. The record contains date and time, identification of the road section, vehicle identification and information related to the vehicle. A section of the road is generally defined from the border of the intersection that forms its starting point, with an end point defined by the subsequent intersection. Each border of the intersection falls within one of the 79 districts that are categorized in 8 regions[4].

In the initial step, we filter out vehicles that are not used for the transportation of goods and commodities, e.g. buses. We then aggregate daily records into weekly time series for each district, i.e. a number of passages for each vehicle is counted at each intersection. In the next step, the observed weekly time series are transposed into mathematical objects for each district. This results in a set of 715 observed mathematical objects, also called curves, which represents a functional variable.

Table 1 shows the aggregated dataset of weekly time series that represents a set of data points for each district, which are determined as a number of passages of vehicles through monitored (inter)sections of roads. Each district is categorised in a particular region that is shown in the last column.

District	**Week 1**	**Week 2**	**…**	**Week 52**	**Week 53**	**Region**
Bánovce	40.1	126.2	…	68.3	33.5	B. Bystrica
B. Bystrica	68.0	257.0	…	163.6	63.6	B. Bystrica
…	…	…	…	…	…	…
Michalovce	47.0	166.0	...	118.0	67.4	Košice
…	…	…	…	…	…	…
Zvolen	199.1	675.0	…	418.6	220.0	Žilina

Table 1: Weekly time series of passages of vehicles by district (in thousands), 2020 (abridged).

Table 1 shows that the number of passages of vehicles in the first and the last week rapidly decreases, which is caused by the incomplete weeks, i.e. a number of days is less than seven. We further note that differences in the volume of the freight among districts can be very large. The following table shows a number of districts by region.

Category	Region	Number of districts
1	B. Bystrica	13
2	Bratislava	4
3	Košice	7
4	Nitra	7
5	Prešov	13
6	Trenčín	9
7	Trnava	7
8	Žilina	11

Table 2: A number of districts by region, 2020.

1.2 Methodology

To formalise the notation, we denote the matrix of real-valued observations in Table 1 as $\boldsymbol{X} = x_{ij}$, where $j = 1, \dots, p$ refers to a number of weeks and $i = 1, \dots, n$ to a number of districts. In the initial step, a function is fitted into a matrix of real numbers x_{ij} across each district i. A notation for such a functional variable is $\chi_i(t)$, where $i = 1, \dots, n$ now refers to a number of curves for each district that are observed on a continuous domain $t \in T$. Theoretical concepts of functional spaces are provided by [4].

[4] https://en.wikipedia.org/wiki/Districts_of_Slovakia.

[5] Note that 5 districts of Bratislava and 4 districts of Košice are merged into one district for Bratislava and one district for Košice, respectively.

We use the basis spline expansion, which is adaptable to any complicated data structure, to construct continuous functions $\chi_i(t): R \to R$ for each district i. Hence, the functional variable $\chi_i(t)$ can be expressed in terms of the basis spline expansion [4] as

$$\chi_i(t) = \sum_{k=1}^{K} c_{ik} B_k(t) \tag{1}$$

The basis spline functions defined in the $k \times 1$ vector $B_k(t)$ are piecewise polynomials[6] that are automatically tied together at its breakpoints, where k refers to a number of polynomials used in the expansion which determines a degree of smoothing. Note that we use the same $B_k(t)$ across all observations i. The c_{ik} are the values of a $i \times k$ matrix of unknown parameters that needs to be estimated. The theoretical properties related to spline functions are provided by [2].

The application of basis spline expansion to real data can be done through the familiar technique of fitting statistical models to data by using the method of least squares, which leads to an estimation of unknown parameters c_{ik} as

$$\hat{c}_{ik} = \left(B_k(t)^{\mathrm{T}} B_k(t)\right)^{-1} B_k(t)^{\mathrm{T}} x_{ij} \tag{2}$$

where a subscript T refers to a transpose.

In the next phase of analysis, the features of the functional matrix of observed curves $\chi_i(t)$ are used to fit a regression model by constructing a set of binary independent variables in the $n \times (C+1)$ covariate matrix $\mathbf{Z} = z_{i(c)j}$, where the category $c = 1, \dots, C$ refers to individual regions. The columns $j = 1, \dots, C+1$ of the covariate matrix $\mathbf{Z}$ are constructed as a binary indicator that represents a category of individual observed curves, with ones in the first column that serves for the estimation of the mean function across all curves. The number of curves in each category is denoted by N_c, which is shown in Table 2. It follows that we can formalise the functional ANOVA model [4] as

$$\chi_{i(c)}(t) = \mu(t) + \alpha_c(t) + \varepsilon_{i(c)}(t) \tag{3}$$

where the subscript (c) refers to the category of individual curves, the function $\mu(t)$ is the overall mean across all N sample curves, where $N = \sum_{c=1}^{C} N_C$. The unknown parameters $\alpha_c(t)$ are the specific effect functions which represent departures from $\mu(t)$ specific for each category c. The error function $\varepsilon_{i(c)}(t)$ are the unexplained variations specific to the i^{th} sample curve within category c that are assumed to be independently and identically distributed with 0 mean and a constant variance σ^2. Similarly, as in the classic ANOVA framework, the following constraint has to be satisfied in order to identify individual categories uniquely $\sum_c \alpha_c(t) = 0$ for all t.

The ANOVA model in eq. (3) can be re-expressed by defining a set of regression functions $\beta_j(t)$, where $\beta_1(t) = \mu(t)$, $\beta_2(t) = \alpha_1(t)$ and so on, as

$$\chi_i(t) = \beta_j(t) z_{i(c)j} + \varepsilon_{i(c)}(t) \tag{4}$$

where $\beta_j(t) = \left(\mu(t), \alpha_1(t), \dots, \alpha_C(t)\right)^{\mathrm{T}}$ is the functional vector. Similarly as $\chi_i(t)$, the regression functions $\beta_j(t)$ can also be expressed in terms of the basis spline expansion as $\beta_j(t) = \sum_{k_\beta=1}^{K_\beta} b_{jk_\beta} \theta_{k_\beta}(t)$, where $\theta_{k_\beta}(t)$ is a vector of spline functions of length $k_\beta \times 1$, with a number of spline functions k_β chosen to be sufficiently large not to entail any significant loss of information. The $j \times k_\beta$ matrix of unknown parameters b_{jk_β} remains to be estimated. The functional form of the ordinary least squares method leads to the following estimation of unknown parameters b_{jk_β} [4]

$$vec\left(\hat{b}_{jk_\beta}\right) = \left(\mathbf{J}_{\theta_{k_\beta}\theta_{k_\beta}(t)} \otimes (\mathbf{Z}^{\mathrm{T}}\mathbf{Z})\right)^{-1} vec\left(\mathbf{Z}^{\mathrm{T}}\mathbf{C}\mathbf{J}_{B_k(t)\theta_{k_\beta}(t)}\right) \tag{5}$$

where $vec(.)$ denotes a vectorised matrix, $\otimes$ is the Kronecker product, $\mathbf{J}_{\theta_{k_\beta}(t)\theta_{k_\beta}(t)} = \int \theta_{k_\beta}(t)\theta_{k_\beta}{}^{\mathrm{T}}(t)dt$ and $\mathbf{J}_{B_k(t)\theta_{k_\beta}(t)} = \int B_k(t)\theta_{k_\beta}{}^{\mathrm{T}}(t)dt$ are the derivatives of spline functions over the domain t and $\mathbf{C} = c_{ik}$. The theoretical properties of the estimator in eq. (5) are provided by [4].

[6] Note that we also need to define an order of the polynomial.

2 Results

In this section, we show the application of the theoretical framework outlined in the previous section. Note that for the basis spline expansion of $\chi_i(t)$, we use 30 basis splines[7] of order 3 that are defined on a continuous domain $t \in [1,53]$. Hence, we get 71 curves corresponding to each district, which are displayed in Figure 1 below.

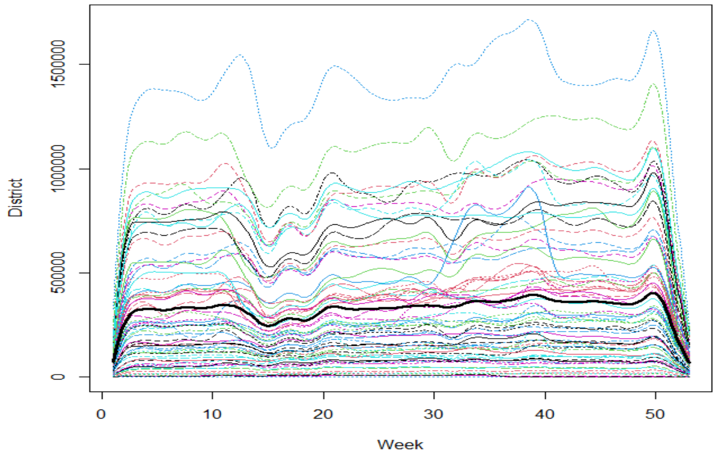


Figure 1: A sample of observed curves for 71 districts (a thick black line represents a mean function across all 71 observed curves).

Figure 1 shows that there is a large difference between individual curves in the sample, i.e. a volume of the freight is significantly different between individual districts. The largest volume of the freight is in the district Svidník (over a million passages in every week, except the first and the last) and the lowest is in the district Medzilaborce (less than a thousand passages in every week).

In the following step, we carry out the application of functional ANOVA, which leads to the estimation of the unknown parameters $\mu(t)$ and $\alpha_c(t)$. Figure 2 displays the estimated parameters for the overall mean across all 71 observed curves $\hat{\mu}(t)$ and for specific effect functions $\hat{\alpha}_c(t)$ for each region $c = 1, \dots, 8$.

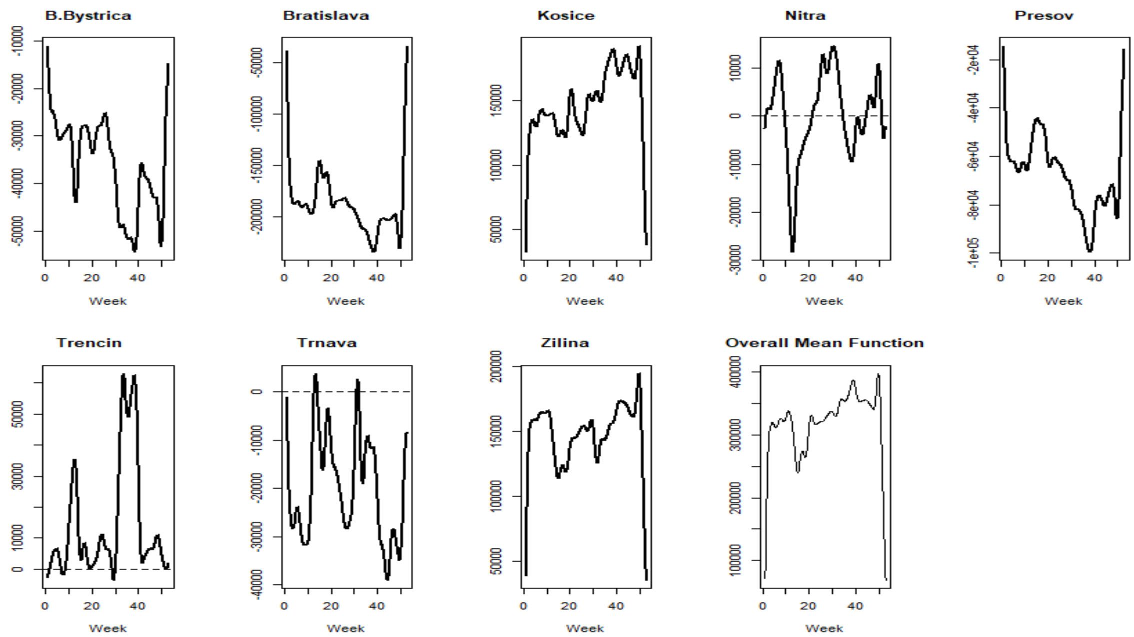


Figure 2: The last panel shows the estimated overall mean function $\hat{\mu}(t)$ across all 71 observed curves. The panels 1-8 show the estimated specific effect functions $\hat{\alpha}_c(t)$ for each region.

[7] A number of basis splines is selected by the author such that we do not lose any significant information from original data.

Figure 2 shows that we are able to study a behaviour of the estimated $\hat{\mu}(t)$ and $\hat{\alpha}_c(t)$ on the entire domain t rather than as point-wise estimates in the classical ANOVA. From the similar pattern of the region Žilina with the overall mean function, we can conclude that the freight in this region is dominant. Moreover, the shape and the magnitude of the deviation from the overall mean function, which is shown as a deviation from zero (dotted line), of the estimated parameters $\hat{\alpha}_c(t)$ are significantly different between individual regions. There are also no clear similarities in the seasonal effect between regions. Some similarities of the freight intensity are shown between regions Bratislava and Prešov between 30th and 40th week, and between the overall trend of regions Banská Bystrica, Bratislava and Prešov. The least deviation from the overall mean function is shown by the region Nitra since its frequent oscillation around zero.

To gain a clearer picture of the deviation of each region from the overall mean function, it is worthwhile examining the estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$, which are displayed in Figure 3 below.

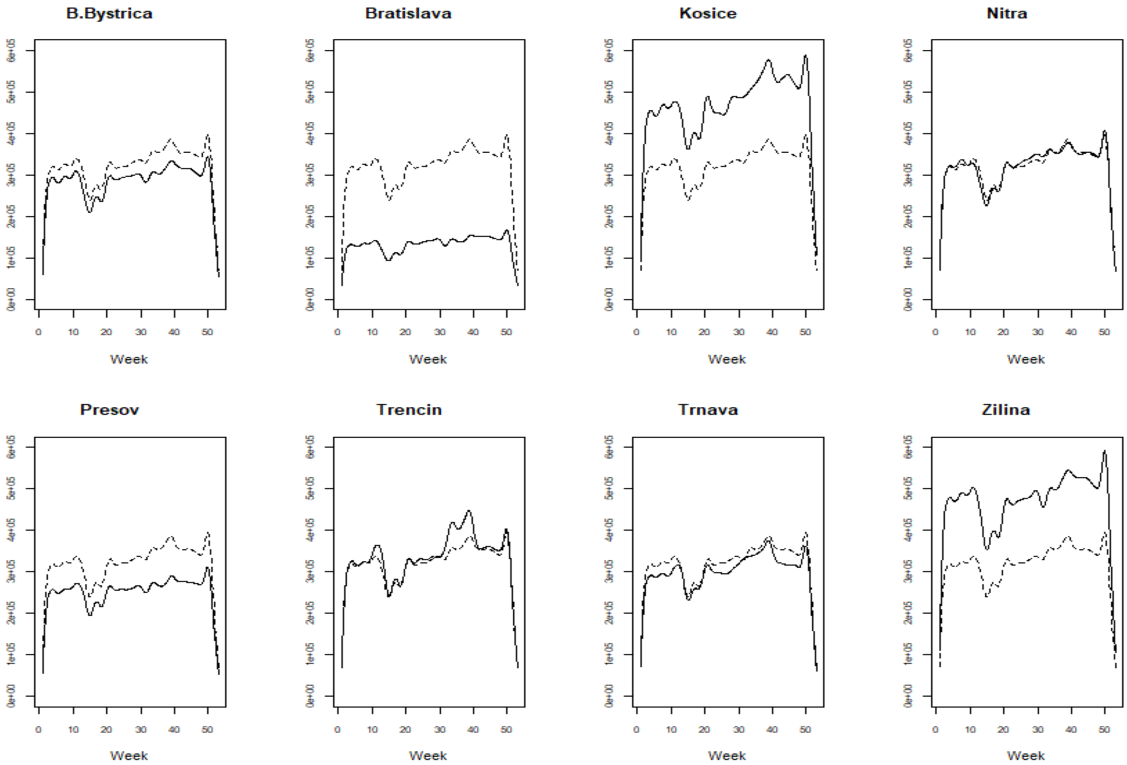


Figure 3: The panels 1-8 show the estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$ for each region, along with the overall mean function $\hat{\mu}(t)$ as a dotted line.

Figure 3 further emphasises that significant differences in the freight intensity among individual regions exist. The regions with the highest freight intensity are Košice and Žilina, and the region with the lowest freight intensity is Bratislava – we note that the geographical size of the region could play a role. Moreover, the regions Nitra, Trenčín and Trnava attain the lowest deviation of the freight from the overall mean function.

Overall, we again emphasise that the functional ANOVA allows for the interpretation of the estimated parameters on its entire domain. This phenomenon can uncover seasonal effects of the freight within individual regions and differences in the shape and the volume of the freight between regions. The empirical results of the analysis in this paper are useful for management control of the freight flow through the road network or can be studied alongside other regional economic variables and environmental indicators.

Additionally, we note that there could be dependencies between observed curves of neighbouring districts that would violate the assumption of independence between the error functions $\varepsilon_{i(c)}(t)$ in the functional ANOVA. The presence of the dependence between observed curves requires the application of more complex statistical models that is part of future research.

3 Conclusion

In this section, we provide a brief summary of the paper and its main objectives that are followed by concluding remarks and proposals for further research.

In this paper, we demonstrate the application of the functional analysis of variance (ANOVA) that explores a variability in the observed sample curves, where the individual curves are classified within categories. The computation strategy for the estimation of the ANOVA model is similar to its classical counterpart, except that the usual summation in the estimator function is replaced by the integration over a continuous domain. The main difference is that the unknown parameters are estimated as functions. Hence, the functional parameters can uncover details of variations on the entire domain of the observed period.

New digital data sources offer an opportunity for scientists to study the application of more complex statistical models. In the empirical application, we use data that are collected by the electronic toll system in Slovakia. An observed dataset contains daily digital records of passages of vehicles through monitored sections of roads during the 2020. The main objective of the paper is to carry out the analysis of the intensity of the regional freight in Slovakia. The results of the research can be useful for the management control of the freight through regions, building new road infrastructure and further regional economic analysis.

The empirical results of ANOVA reveal that there are no common seasonal effects of the freight within individual regions. Moreover, a difference in the shape and the volume of the freight between regions can be very large. A deviation from the overall mean function identifies regions with a high and a low volume of the freight that can serve for a diversification of the road transport and a more efficient investment reallocation for building the new road infrastructure.

A transportation of goods and commodities carried out by vehicles through roads can provide an early indicator of the economic activity for industries that rely on the road transport. A further research could address a development of the indicator that would show a monthly change in the freight. However, a design for this analysis could require a more complex algorithm that would filter out the freight that is transitory through Slovakia and further information about the economic value of the transported goods and commodities.

The analysis in this paper opens numerous further research inquiries. One of the most important is to examine a degree of the dependence between the observed curves that would trigger an application of more complex statistical models, which incorporate a spatial dependence. Moreover, a theoretical framework of the econometric model can be expanded to study the relationship between the freight and various macroeconomic variables or its impact on the environmental indicators.

Acknowledgements

This work was supported by the Grant Agency of the Slovak Republic VEGA 1/0047/23 „The importance of spatial spillover effects in the context of the EU's greener and carbon-free Europe priority “.

References

[1] Aneiros, G., Cao, R., Fraiman, R., Genest, C., Vieu, P. (2019). Recent advances in functional data analysis and high-dimensional statistics, *Journal of Multivariate Analysis*, Volume 170, 3–9, ISSN 0047-259X. https://doi.org/10.1016/j.jmva.2018.11.007.

[2] de Boor, C. (2001). *A Practical Guide to Splines*. Revised Edition. New York: Springer.

[3] Cox, M., Triebel, J., Linz, S., Fries, C., Flores, L.F., Lorenz, A., Ollech, D., Dietrich, A., LeCrone, J., Webel, K. (2020). Daily truck toll mileage index based on digital process data from toll collection system. *Federal Statistical Office (Statistisches Bundesamt)*.

[4] Ramsay, J. O., Silverman, B. W. (2005). *Functional Data Analysis*. Second Edition. Springer Series in Statistics, Springer.

[5] Ramsay, J. O., Hooker G., Graves, S. (2009). *Functional Data Analysis with R and MATLAB*. Use R!, Springer.

[6] Knížat, P., Ceľuchová Bošanská, D., Janík, M., Nguyen, F. (2024). New data sources in statistics: the effect of freight intensity on macroeconomic indicators (in Slovak). *Slovak statistics and demography, 34(1)*, 7–26.

Flexible Job Shop Scheduling with setup, transportation and planned machine idle time.

František Koblasa[1], Miroslav Vavroušek[2]

Abstract. The Flexible Job Shop Scheduling Problem (FJSP) is one of the most popular scheduling models because of its ability to describe various real-life manufacturing systems. Despite being used mainly in its natural form, more practical constraints such as transportation and setup times have attracted attention in the last decade as setups and internal transportation are the most visible non-valued added processes.
This article focuses on Flexible Job Shops with transport and setup times, adding planned idle times between machine operations. Those idle times depend not on job types as sequence-dependent setups but on machine type and represent regular maintenance, administration, scrap management, etc.
This article aims to enhance the known FJSP models with setup and transportation times by idle time constraints and test the real-world approach of dispatching rules against the advanced evolution algorithm technique. The basic scheduling generation technique is compared with the earliest processing job start selection.
Together with the model, known FJSP testing instances are modified to suit the needs of the above-mentioned constraint and experiment. Generalization of testing instances modifications to real-world and combinatorial optimization needs is discussed.

Keywords: Flexible Job Shop Scheduling Problem, transportation, setup, planned idle time, Evolution Algorithm.

JEL Classification: C63, L23
AMS Classification: 90C59

1 Introduction

Thanks to its complexity, the job shop scheduling problem is the most used model among classical scheduling models (Job shop, Flow shop, Open shop). With a hundred WOS papers (WOS search query "Job shop scheduling" (Topic)) in the 2000s, two hundred in the late 2010s, and four hundred per year since 2020 (WOS query), it is significantly more than tens of Flow shops in the 2000s and a hundred and a half in the 2020s. Open shop problem is the least favourite, with up to 15 papers peak in the year, as it can be solved by the Flexible Job Shop Scheduling Problem, which is not only focusing on the scheduling part of the problem as classical Job Shop but also on job assigning to the resource. Thus, it is close to the problem of Open shops.

Flexible job shop scheduling has growing potential (Xie et al. 2019; Dauzère-Pérès et al. 2024) as the number of papers closing to 150 is the same popular as the classical Flow Shop – the most common model describing the Lean-favourable manufacturing system of production lines. The Flexible Job Shop is not only popular for its complexity, thus a challenging problem (Dauzère-Pérès et al. 2024) to be solved by AI-bioinspired algorithms, but due to its potential applicability in practice. There are numerous practical applications in both general and research publications, beginning with steel tubes (Li and Huo 2009), cars (Calleja and Pastor 2014), casting (Tang et al. 2019), injection machines (Tanev et al. 2004) and even weapons manufacturing (Chen et al. 2012).

Since the setup operation is one of the most relevant non-value added times, it is covered in the literature (cca 70 papers in WOS) with a special focus on sequence-dependent setup times (Shen et al. 2018; Gao et al. 2019). Transportation (Homayouni et al. 2023), with an increase in Automated Guided Vehicles topics (Destouet et al. 2023), has become popular as well (over 120 papers). However, combining both (Li and Lei 2021) is seen less often (around 30 papers).

[1,2]Technical university of Liberec, Department of manufacturing systems and automation, Studentská 2, Liberec 1, Czech Republic

[1] frantisek.koblasa@tul.cz

[2] miroslav.vavrousek@tul.cz

This paper focuses on both problems of transportation and setup (FJSP-ST) while adding a break (planned idle time) in which the machine is cleaned, preheated, proactively maintained, administration done, worker breaks, worker-to-worker transfer of work or any other nonproductive activities done. That describes situations and their durations where it is not possible to start with a predefined activity (machining, setup, transport) at the very same moment other activity ended.

This paper is organised as follows. The second chapter focuses on the Flexible JSP model with setup, transportation and idle time. The third chapter describes known modifications to test instances of the defined FJSP-STI problems. The fourth chapter focuses on the definition of heuristic and metaheuristics used to find the best values of makespan objective functions. The fifth defines the experiment and discusses experimental results.

2 FJSSP model with setup, transport and planned machine idle time

FJSSP problem is known NP-hard problem. It consists of two decision problems, the first is assigning operations to machines (assignment), and the second is determining the processing order of jobs on machines (sequencing). Each job $J=\{J_1,\ldots, J_q, \ldots, J_i,\ldots,J_n\}$ consists of a sequence of operations O_{ijp} where i is job, j is a type of job and $k=\{1,...,p, ...,s_i\}$ is a sequence of operations, where s_i is the maximal number of operations in job i. They are allowed to be processed on any of the feasible machines $M=\{M_1,\ldots, M_u, \ldots, M_m,\ldots, M_l\}$. All n jobs and l machines, are in theoretical problems available at time 0, and a machine can only execute one operation at a given time. Preemption is not allowed.

The objective function in our case is chosen to be the minimisation of total competition of jobs C_{max} makespan (1)

$$C_{max} = \max_{1\leq i\leq n}\{C_i\} \tag{1}$$

where C_i is the competition time of job i and n is the maximum number of jobs.

The following constraints (2)(3)(4)(5) (see Demir and Kürşat İşleyen 2013) define FJSP in the matter of the starting t_{st} and ending t_e times of activities such as classical processing time t^{ct} and real-world problem constraints setup time t^{co}, transportation time t^h and newly defined planned idle time t^{id}. No subscript (st or e) indicates duration of such activity. t_{st}^{o} (t_e^o) sugest earliest (lastest) starting (ending) time of operation.

The machine can process only one job, and the job can be processed only on one machine (2)(3)(4) at that moment.

$$t^{ct}_{st(i,j,p,m)} + t^{ct}_{(i,j,p,m)}x_{ijpm} \leq C_{ijp} \quad \forall i,j,p,m \tag{2}$$

$$x_{ijpm} = \begin{cases} 1 \text{ if the } O_{ijp} \text{ is performed on machine } M_m \\ 0 \end{cases} \tag{3}$$

$$\sum_{m=1}^{l} x_{ijpm} = 1 \tag{4}$$

where $t^{o}_{st(i,j,p,m)}$ is starting time of p-th operation O_{ijp} of job i with job type j made on machine m, $t^{ct}_{(i,j,p,m)}$ is duration of cycle time of O_{ijp} operation, C_{ijp} is competition time and x desition variable.

There is a precedence of operations (5) where $t^{J}_{st(i,j,p+1,m)}$ is earliest time job j next operation ($p+1$) can start.

$$t^{o}_{st(i,j,p,m)} + t^{co}_{(i,j,p,m)} + t^{ct}_{(i,j,p,m)} \leq t^{J}_{st(i,j,p+1)} \quad \forall i,j,k,m \tag{5}$$

Setup is made when the machine m has to process a different job type j operation (decision variable y=0) than it was previously processed (6)(7).

$$t^{co}_{st(i,j,p,m)} - t^{co}_{e(i,j,p,m)} = \left(1 - y_{ijpm}\right)t^{co}_{(i,j,p,m)} \qquad \forall i,j,p,m \tag{6}$$

$$y_{ijpm} = \begin{cases} 1 \text{ if } O_{ijp} \text{ has same job type } j \text{ as preceding operation on } M_m \\ 0 \end{cases} \tag{7}$$

The operation O_{ijp+1} cannot start $t^{o}_{st(i,j,p+1,m)}$ (8) before both the (9) job $t^{J}_{st(i,j,p+1)}$and machine (10) $t^{M}_{st(i,j,p+1,m)}$ are available. Job is avaiable after preceding operation is finished $t^{o}_{e(i,j,p,u)}$ and job is transported to next stage by duration $t^{h}_{(u-m)}$. Machine m is ready when any preceding operation O_{qjp} ends(10) and its $t^{id}_{(m)}$ idle time passes.

$$t^{o}_{st(i,j,p+1,m)} \geq max\left\{t^{J}_{st(i,j,p+1)}, t^{M}_{st(i,j,p+1,m)}\right\} \quad \forall i,j,p,m \tag{8}$$

$$t^{J}_{st(i,j,p+1)} = t^{o}_{e(i,j,p,u)} + t^{h}_{(u-m)} \quad \forall i,j,p,m,u \tag{9}$$

$$t^{M}_{st(i,j,p+1,m)} = t^{o}_{e(q,j,p,m)} + t^{id}_{(m)} \quad \forall\, i,q,j,p,m \tag{10}$$

Unlike adding minimal machine idle time $t^{id}_{(m)}$, job will be transported (11) between machines *m* and *u* by duration of $t^{h}_{(u-m)}$ only (desition variable *z*) (12) if the next operation is processe on a different ($m \neq u$) machine. $t^{h}_{st(i,j,p,u)}$ indicates when transport operation starts and $t^{h}_{e(i,j,p,u)}$ ends.

$$t^{h}_{e(i,j,p,m)} - t^{h}_{st(i,j,p,u)} = \left(1 - z_{ijk}\right) t^{h}_{(u-m)} \quad \forall i,j,k,m \tag{11}$$

$$z_{ijk} = \begin{cases} 1 \text{ if the } O_{ijp} \text{ and } O_{ijp-1} \text{ are both processed on machine } M_m \\ 0 \end{cases} \tag{12}$$

Simple Giffler and Thompson constructive algorithm for the active schedule is modified to construct a feasible solution while taking into account equations for transportation, setup and idle time (6)-(12)

The general description of a constructive algorithm for FJSP is scheduling operation *O* in step $t = \{1,r\}$, where *r* is the total number of operations:

1. Creating list V_t of all schedulable *O* operations in step *t*, including all machine variants *v*.
2. Conflict set f_t* creation - find the possible earliest ending time t^{o}_{e}

$$f^{*}_{t} = min\ OinV_t\{t^{o}_{e}\} \tag{13}$$

   and machine M_m on which t^{o}_{e} occurs
3. Choose an optimal operation (by heuristics or metaheuristic – see chapter 4) which requires M_m and its starting time $t^{o}_{st(i,j,p,m)} < t^{o}_{e}$
4. Continue until there is *O* unscheduled.

3 Testing instances

Development of the test instances in this paper, is based on analysis of the general approach to generate FJSP instances and is trying to take the most of the information from the original most used Brandimarte (Brandimarte 1993) approach of uniformly random generation of MK01-MK10.

Instance	n	l	s_i *min*	s_i *max*	v	t^{ct} *min*	t^{ct} *max*	t^{h}	t^{co}	t^{id}
MK01_STI	10	6	5	7	3	1	7	(14)	(15)	(16)
MK02_STI	10	6	5	7	6	1	7			
MK03_STI	15	8	10	10	5	1	20			
MK04_STI	15	8	3	10	3	1	10			
MK05_STI	15	4	5	10	2	5	10			
MK06_STI	10	15	15	15	5	1	10			
MK07_STI	20	5	5	5	5	1	20			
MK08_STI	20	10	5	15	2	5	20			
MK09_STI	20	10	10	15	5	5	20			
MK10_STI	20	15	15	15	5	5	20			

Table 1 MK_STI test instances random ranges

The original range of *n* jobs *J*, *l* machines *M*, the maximal/minimal number s_i of operation *O* per job *J*, the maximal number of parallel machines *v* on which operation can be made and the minimal/maximal duration of processing time t^{ct} is used while calculating the duration of transportation (handling) t^{h} (14), setup t^{co} (changeover) (15) and idle planned time t^{id}. Handling time t^{h} (14) to machine M_m is defined as the difference between the index of machines used for O_{ijp} and O_{ijp-1} , which highlights that transportation duration is 0 if the machine does not change. Real-world scheduling problem of transport and setup requires special case if p=1 that job was on no machine m=0 and no job was on machine q=0 to reflect the situation when the job starts first operation (M_0 represents the warehouse) or there was no job on machine before (preceding job index *q*).

$$t_{u-m}^{h} = |u - m| \quad \forall m, u \tag{14}$$

Setup is sequence-dependent, and its duration $t_{(i,j,p,m)}^{co}$ (15) is calculated, in this paper instances, based on job type indexes *j*,*q* as the difference between the previous *q* and current *j* job type on the scheduled machine *m*. This means if the job type attends the same machine, the setup will be 0. Job type *j* and job *i* index in defined _STI instances *j* = *i* as there are no job types in the original MK instances so job indexes are used. Same as in the case of transport if *p*=1 than *q*=0 so duration setup is equal to current job type.

$$t_{(i,j,p,m)}^{co} = |j - q| \quad \forall i, j, q, p, m \tag{15}$$

Based on the original test instances, the planned idle time is challenging to define, and no current instance information seems relevant to the defined problem. In this case, duration is based on the number of machine variants v_{ijp} per operation O_{ijp}. It represents the difficulty of the operator in preparing the job for a number of machines.

$$t_{(m)}^{id} = v_{ijp} \tag{16}$$

4 Dispatching rules and metaheuristics to optimise FJSP-STI

It is usual to test optimisation algorithms against known test instances against their lower bounds. The difficulty of developed instances is usually shown by comparing the lower and upper bounds. However, that would require extensive use of some branch and bound methods (Kubiak et al. 2020) which is very time-consuming and is out of the scope of this paper. Instead, this paper is using 3 dispatching rules, including the most popular (Tan et al. 2013) Shortest Processing Time (SPT), Critical Ratio (CR) (Červeňanská et al. 2021) and Random generation (Rnd) (Tay and Ho 2008). That can show a possible range of results, thus the difficulty of defined instances.

$$CR = \frac{DD - t_{st,p}^{ct}}{\sum_{k=p}^{s_i} max\left(t_{k,v}^{ct}\right)} \tag{17}$$

Where *DD* is due date given $t_{st,p}^{ct}$ is starting time of current operation *p*, and $t_{p,v}^{ct}$ is cycle time of current operation as classical Critical ratio works with sum of future operations and there are multiple operation variants, we take the maximal out of them.

To optimise the schedule Evolution Algorithm (EA) is used:

- Representation: Random key (Bean 1994).
- Size of population: $\mu = 2jl + 100$ where *j* is total number of jobs and *l* total number of machines.
- A number of generations: Max 1000 or 10 generations without improvement.
- Parent selection: Biased with incest control (Koblasa et al. 2020).
- Crossover: uniform crossover with crossing probability *Pc*=85% that new individual will share gene of 2nd parent. Number of new individuals $\lambda = \mu$.
- Elimination: Elitist strategy surviving of the fittest μ.

5 Tests and result discussion

Test of defined benchmark instances is done by the three dispatching rules (SPT, CR, Rnd) and evolution algorithm. As both Rnd and EA are stochastic-based, test runs are conducted ten times. Rnd, with motivation to find a range of possible lower and upper bound, and EA to see if instances have some hard local optima.

Table 2 consists of 10 benchmarking instances results, 3 priority rules, including minimal, maximal, and average values of both Rnd and EA makespan objective function.

The results show that among dispatching rules CR outperforms both SPT (which was used as the Due Date for CR calculation) and Rnd. CR was better in all cases except one out of 10 solutions generated by Rnd (MK07_STI RnD min < CR). SPT, as one of the popular dispatching rules (except FIFO), gives us the worst results; in most cases, it was worse than the Rnd average. All except MK_40_STI , which, together with CR, hit the “near-optimal” solution of EA.

The evolution algorithm outperformed all dispatching rules, as expected; however, as the range coefficient (EA R.coef) suggests, its results could have been more consistent. It was planned to use the classical Six Sigma *Cp; Cpk* consistency test; however, the results don't fall under the normal distribution.

Instance	*SPT*	*CR*	*Rnd min*	*Rnd max*	*Rnd avg*	*EA min*	*EA max*	*EA avg*	*EA R.coef*
MK01_STI	122	109	112	128	120	89	94	91,6	5,46%
MK02_STI	127	108	110	138	123,4	98	100	98,7	2,03%
MK03_STI	514	346	380	474	415,5	326	336	330,7	3,02%
MK04_STI	164	164	172	223	190,3	147	148	147,2	0,68%
MK05_STI	374	348	353	410	373,3	320	328	325,3	2,46%
MK06_STI	256	210	222	263	237,7	198	201	199,5	1,50%
MK07_STI	436	361	349	431	387	309	324	315,3	4,76%
MK08_STI	1087	799	848	1034	886,2	744	759	752,7	1,99%
MK09_STI	790	616	652	818	717,3	590	609	598	3,18%
MK10_STI	740	503	525	674	598,7	488	494	491,5	1,22%

Table 2 – Makespan test results of MK_STI instances

Table 3 shows basic instances of properties found in preliminary tests. *LB* orig. is showing original lower bound of Brandimarte test instances. Improved Jaya (IJA) by (Yang and Gao 2018) shows the result of one of the best achieved makespan results on original MK instances for comparison (*EA Orig*).

Instance	*LB orig*	*EA Orig*	*LB**	*UB**	*R. coef**	*EA L min*	*EA L n*	*EA G avg*
MK01_STI	36	40	89	128	30,47%	94	3	25,0
MK02_STI	24	27	98	138	28,99%	98	5	19,2
MK03_STI	204	204	326	514	36,58%	331	3	25,1
MK04_STI	48	60	147	223	34,08%	147	6	16,4
MK05_STI	168	172	320	410	21,95%	328	3	19,7
MK06_STI	33	57	198	263	24,71%	199	3	23,2
MK07_STI	133	139	309	436	29,13%	324	3	23,4
MK08_STI	523	523	744	1087	31,55%	757	3	36,3
MK09_STI	299	307	590	818	27,87%	594	5	31,6
MK10_STI	165	197	488	740	34,05%	491	3	31,4

Table 3 – Test instances MK_STI properties

Hypothetical Loverbound (*LB**) and Upper Bound (*UB**) are set based on the worst-case dispatching rule (*UB**) and the best-achieved by EA (*LB**). The range coefficient (*R. Coef**) between *LB** and *UB** shows that EA outperformed the dispatching rule by at least 22%. *EA L n* min shows an interesting local optimum, which was obtained at least 3 times out of 10 tests. The most interesting result is a case of MK04_STI instance where the local optimum was achieved 6 times, and at the same time, it was best found local optimum. *EA G avg* shoves generation number in which optimisation stopped after not getting better results for 10 generations. A value of 16,4 average generations suggests that there were only 6 generations, in this case, which brings improvement.

6 Conclusion

The presented research defined an extended model (8-10) of the classical flexible job shop scheduling problem by joining setup and transportation with planned idle machine time. New testing instances were defined for the newly defined model using an approach based on the original FJSP and using original data as a difference between the machine ID as transportation time (14), job ID difference as setup time (15), and job variants (16) as machine idle time. This allows researchers to use original instances (MK1-10) without searching for the exact seed of randomly generated instances. Newly created instances (_STI) are tested using dispatching rules and the EA algorithm to define their lower and upper bounds. However, this research is limited to a small number of repetitions and the suitability of EA to find *LB*s. Some more suitable algorithms will be necessary to find a better LB.

Further research will focus on finding more suitable ways to find *LB*, such as Branch and bound-based algorithms, of defined test instances and adding more constraints to cover more real-life scheduling problems. Research will also focus on local search-inspired mutation operators to improve the assigning part of scheduling optimization.

Acknowledgements

This publication was written at the Technical University of Liberec, Faculty of Mechanical Engineering with the support of the Institutional Endowment for the Long Term Conceptual Development of Research Institutes, as provided by the Ministry of Education, Youth and Sports of the Czech Republic in the year 2024. The research reported in this paper was supported by institutional support for nonspecific university research.

References

Bean JC (1994) Genetic algorithms and random keys for sequencing and optimization. ORSA J Comput 6:154–160. https://doi.org/10.1287/ijoc.6.2.154

Brandimarte P (1993) Routing and scheduling in a flexible job shop by tabu search. Ann Oper Res 41:157–183. https://doi.org/10.1007/BF02023073

Calleja G, Pastor R (2014) A dispatching algorithm for flexible job-shop scheduling with transfer batches: an industrial application. Prod Plan Control 25:93–109. https://doi.org/10.1080/09537287.2013.782846

Červeňanská Z, Važan P, Juhás M, Juhásová B (2021) Multi-Criteria Optimization in Operations Scheduling Applying Selected Priority Rules. Appl Sci 11:2783. https://doi.org/10.3390/app11062783

Chen JC, Wu C-C, Chen C-W, Chen K-H (2012) Flexible job shop scheduling with parallel machines using Genetic Algorithm and Grouping Genetic Algorithm. Expert Syst Appl 39:10016–10021. https://doi.org/10.1016/j.eswa.2012.01.211

Dauzère-Pérès S, Ding J, Shen L, Tamssaouet K (2024) The flexible job shop scheduling problem: A review. Eur J Oper Res 314:409–432. https://doi.org/10.1016/j.ejor.2023.05.017

Demir Y, Kürşat İşleyen S (2013) Evaluation of mathematical models for flexible job-shop scheduling problems. Appl Math Model 37:977–988. https://doi.org/10.1016/j.apm.2012.03.020

Gao K, Cao Z, Zhang L, et al (2019) A review on swarm intelligence and evolutionary algorithms for solving flexible job shop scheduling problems. IEEECAA J Autom Sin 6:904–916. https://doi.org/10.1109/JAS.2019.1911540

Homayouni SM, Fontes DBMM, Gonçalves JF (2023) A multistart biased random key genetic algorithm for the flexible job shop scheduling problem with transportation. Int Trans Oper Res 30:688–716. https://doi.org/10.1111/itor.12878

Koblasa F, Vavroušek M, Manlig F (2020) Selection Strategies in Evolution Algorithms and Biased Selection with Incest Control. In: 38th International Conference on Mathematical Methods in Economics 2020. Brno

Kubiak W, Feng Y, Li G, et al (2020) Efficient algorithms for flexible job shop scheduling with parallel machines. Nav Res Logist NRL 67:272–288. https://doi.org/10.1002/nav.21901

Li L, Huo J (2009) Multi-objective flexible job-shop scheduling problem in steel tubes production. Syst Eng-Theory Pract 29:117–126. https://doi.org/10.1016/S1874-8651(10)60063-4

Li M, Lei D (2021) An imperialist competitive algorithm with feedback for energy-efficient flexible job shop scheduling with transportation and sequence-dependent setup times. Eng Appl Artif Intell 103:104307. https://doi.org/10.1016/j.engappai.2021.104307

Shen L, Dauzère-Pérès S, Neufeld JS (2018) Solving the flexible job shop scheduling problem with sequence-dependent setup times. Eur J Oper Res 265:503–516. https://doi.org/10.1016/j.ejor.2017.08.021

Tan Z, Chen Y, Zhang A (2013) On the exact bounds of SPT for scheduling on parallel machines with availability constraints. Int J Prod Econ 146:293–299. https://doi.org/10.1016/j.ijpe.2013.07.013

Tanev IT, Uozumi T, Morotome Y (2004) Hybrid evolutionary algorithm-based real-world flexible job shop scheduling problem: application service provider approach. Appl Soft Comput 5:87–100. https://doi.org/10.1016/j.asoc.2004.03.013

Tang H, Chen R, Li Y, et al (2019) Flexible job-shop scheduling with tolerated time interval and limited starting time interval based on hybrid discrete PSO-SA: An application from a casting workshop. Appl Soft Comput 78:176–194. https://doi.org/10.1016/j.asoc.2019.02.011

Tay JC, Ho NB (2008) Evolving dispatching rules using genetic programming for solving multi-objective flexible job-shop problems. Comput Ind Eng 54:453–473. https://doi.org/10.1016/j.cie.2007.08.008

Xie J, Gao L, Peng K, et al (2019) Review on flexible job shop scheduling. IET Collab Intell Manuf 1:67–77. https://doi.org/10.1049/iet-cim.2018.0009

Yang XP, Gao XL (2018) Optimization of dynamic and multi-objective flexible job-shop scheduling based on parallel hybrid algorithm. Int J Simul Model 17:724–733. https://doi.org/doi:10.2507/IJSIMM17(4)CO19

The relationship between investment in machinery, employment, and output of forestry in Switzerland

Richard Kovárník[1]

Abstract. This article deals with time series modelling in the field of forest management. Specifically, it is a time series containing information on investments in machinery and equipment, a time series on the number of work units and a time series on the output of the forestry sector. The data comes from the period between 1992 and 2022 in Switzerland. For this purpose, a vector autoregression model was estimated to model the dynamics between these variables, Granger causality was tested to determine interrelationships and causalities and Impulse-response analysis was performed to understand how quickly and significantly variables respond to shocks. The results obtained indicate that sudden changes in forestry output stimulate additional investment in the following year to meet the increased demand for logging and wood processing. Furthermore, the results indicate that investments in machinery and equipment affect the number of working units. In this direction, there is a decrease in working units if there were additional investments in machinery and equipment in the previous period.

Keywords: vector autoregression, Granger causality, impulse-response, forestry

JEL Classification: C51
AMS Classification: 37M10

1 Introduction

Forest management is not often subjected to quantitative analysis as it is in the case of other sectors. Nevertheless, the tools of quantitative analysis and modern management, which are increasingly in demand these days, have come here as well. In the context of global changes and the growing emphasis on sustainable development, it becomes essential to understand the dynamics of this sector and identify the factors that influence its growth and development.

Several studies have already been published in the field of econometric modelling in forest management. One of them is a study that dealt with the determination of wages in forestry depending on the occurrence of natural disasters (Hampel et al., 2021). Among other things, the paper deals with the lack of information on wages in the forestry sector in the Czech Republic and emphasises the need for systematic scientific research in the field of human resources in the forestry economy. It also analyses the prices of work associated with logging and hauling in forest stands and emphasises that rising wages for forest workers are an impetus for investment in machinery.

This paper focuses on analyzing time series from the forestry sector in Switzerland. Specifically, it examines the overall output of forestry and related secondary activities in relation to investments in machinery and equipment and analyzes the relationships between investments in machinery and employment in forestry. The main goal is to identify and quantify the mutual relationships between these variables, providing a deeper insight into the mechanisms that control the amount of gross investments in machinery and equipment.

2 Data and Methods

To achieve the goals mentioned above, a vector autoregression model (VAR) was used, commonly used in econometrics for time series analysis. VAR models can capture linear interactions between two or more time series and provide a more detailed view of the development of a dynamic system. Another key part of the work was the use of the concept of Granger causality. This approach allows us to see if one time series provides information useful for predicting the other time series, which is necessary to understand the possible causal relationships between investment in machinery and equipment and the total output of the forest economy. Finally, an impulse response analysis was performed, which provides information on how the system reacts to sudden shocks or changes in one variable, regarding the impact on other variables in the model. All calculations were performed using the Python programming language version 3.10.5 and the appropriate libraries (Matplotlib, Numpy, Pandas, Statsmodels).

[1] Mendel University in Brno, Department of Statistics and Operation Analysis, Zemědělská 1, 61300 Brno, xkovarn1@mendelu.cz

2.1 Data

The input dataset contains information on the total economic output of the forestry sector (*forestry_output*), which includes the value of finished production and work intended for sale in millions of euros, information on investments in machinery and equipment (*machinery*) in millions of euros and finally data about the number of annual working units (*employment*) in Switzerland. The data comes from the period between 1992 and 2022 and is captured at an annual frequency. The data comes from the Eurostat database. The time series are shown in Figure 1. For the possibility of plotting the series on a unified graph, the data were standardised, and thus the interrelationships are more obvious.

Figure 1 Graph of standardised time series. Source of data: Eurostat

The graphs show the long-term growth of forestry output and investments. On the other hand, the number of people employed in forestry has been decreasing for a long time. The graphs show a significant increase in all variables around the year 2000. During this period, there was a relatively strong growth in bark beetle populations, which stimulated an increase in timber harvesting and the associated investments in machinery and equipment (Scherrer et al., 2023).

2.2 Vector Autoregression Model

Vector Autoregression (VAR) is a statistical model used mainly in econometrics for the analysis of multivariate time series and processes. VAR models allow each variable in the system to be a function of its own past values and the past values of the other variables in the model. In its basic form, the VAR model is defined as a matrix equation where each variable is a linear function of its own past values and the past values of the other variables in the system (Hamilton, 2020). Mathematically, the VAR model can be expressed as follows:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \varepsilon_t,$$

where y_t is a vector of n variables at time t, $A_1, ..., A_p$ are coefficient matrices, p is the lag order, and ε_t is a vector of residuals governed by multivariate normal distribution.

One of the main assumptions of VAR models is stationarity. The VAR model's variables must be stationary for the statistical tests around the model to be valid. If the variables are not stationary, it is possible to convert them to a stationary form, for example, by differencing them. Tests such as the Dickey-Fuller test or the Kwiatkowski–Phillips–Schmidt–Shin test can be used to verify stationarity (Hamilton, 2020).

Another important aspect for the appropriate estimation and correct specification of the VAR model is the use of the correct lag order, p. Commonly used criteria for selecting an appropriate lag order are the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the Hannan-Quinn Information Criterion (HQC). These criteria usually penalize models with a large number of parameters, aiming to find a compromise between the complexity of the model and its accuracy (Lütkepohl, 2005).

2.3 Granger Causality

Granger causality is a statistical concept used to test whether one time series contains information that can be used to predict the behavior of another time series. This concept was proposed by Clive Granger in 1969 and is based on the idea that if information about the past values of one variable makes it possible to predict the future values of another variable better than it could be predicted without that information, then the first variable is "temporally related" to the second (Granger, 1969). Formally, if we have two stationary time series X and Y, we can say that X is temporally related to Y if the past values of X provide statistically significant information, that improve the predictions of Y. The Granger causality test is usually performed within the VAR model (McCracken, 2016). First, a VAR model containing both variables is estimated, and then tests are performed to determine whether the past values of one variable contribute statistically significantly to the prediction of the other variable. A prerequisite for using Granger causality is that both investigated time series are stationary.

2.4 Impulse-Response Analysis

Impulse-Response analysis evaluates how time series respond to sudden shocks in one of the variables, often using VAR models to understand system dynamics and interactions (Lütkepohl, 2005). It starts with estimating the VAR model to understand dynamic relationships between variables. Shocks, seen as random errors, are identified and isolated using techniques like Cholesky decomposition (Sims, 1980). Impulse-Response Functions (IRFs) show how variables in a VAR model respond to unit shocks over time. They are plotted to observe the system's return to its original state after a shock (Neusser, 2016).

3 Results and Discussion

To ensure correct model estimation, the first step was to verify the stationarity of the time series, which appeared non-stationary at first glance and was confirmed by unit root tests, including the ADF and KPSS tests. Stationarity was achieved by differencing the time series. The differenced time series were then subjected to unit root tests again. For all differenced series, stationarity was confirmed with ADF test p-values < 0.001 and KPSS test p-values > 0.1.

Initially, the relationships between the output of the forest sector and investments in machinery and equipment were tested. The lag order of the model was chosen to be $p = 1$. This choice was made based on the Akaike Information Criterion, the cross-correlation function plot, and the autocorrelation (ACF) and partial autocorrelation function (PACF) plots. The residuals from the model were subsequently tested using the Ljung-Box test, which confirmed the absence of autocorrelation in the residuals, indicating that the assumption about residuals was fulfilled. A summary of the estimated VAR model is captured in Table 1.

Y	X	Coefficient	S.E.	p-value
	const	0.164	0.087	0.059
forestry_output	L1 – *forestry_output*	-0.134	0.197	0.496
	L1 – *machinery*	-0.341	0.191	0.074
	const	0.103	0.075	0.172
machinery	L1 – *forestry_output*	0.501	0.171	0.003
	L1 – *machinery*	-0.344	0.165	0.037

Table 1 VAR model for the relationship between sector output and machinery investment

From the result of the obtained VAR model, statistically significant lagged variable is included in the equation for investments in machinery and equipment with lower p-values than is the critical level of significance $\alpha = 0.05$. This equation can be written in the following way:

$$machinery_t = 0.103 + 0.501 \times forestry\_output_{t-1} - 0.344 \times machinery_{t-1}$$

The estimated model suggest that the increased need for logging and wood processing at time t will trigger additional investments in machinery and equipment at time $t + 1$. After obtaining a suitable VAR model, Granger causality can be tested. Initially, the hypothesis tested whether the time series of investments in machinery and equipment Granger-causes changes in the output of the forestry sector. Subsequently, another hypothesis examined whether the output of the forest sector Granger-causes changes in investments in machinery and equipment. The test results are presented in Table 2.

Null Hypothesis	Parameter	Results
machinery does not Grange-cause *forestry_output*	Test statistic	3.197
	p-value	0.079
	Conclusion	Fail to reject null hypothesis
forestry_output does not Grange-cause *machinery*	Test statistic	8.554
	p-value	0.005
	Conclusion	Reject null hypothesis

Table 2 Granger causality test results for sector output and machinery investment

As part of the test, the null hypothesis could not be rejected, indicating that a change in investments in machinery and equipment does not have a statistically significant impact on the change in economic output from forestry and associated activities. Conversely, statistically significant Granger causality was found in cases where the economic output of forestry affects the level of investment in machinery and equipment. This scenario may occur due to unexpected influences, such as a bark beetle or natural disaster, prompting increased logging and wood processing, necessitating additional investment to boost production. However, it's crucial to consider the nature of Granger causality mentioned earlier. Finally, an Impulse-Response analysis was conducted, with a graphical representation included in Figure 2.

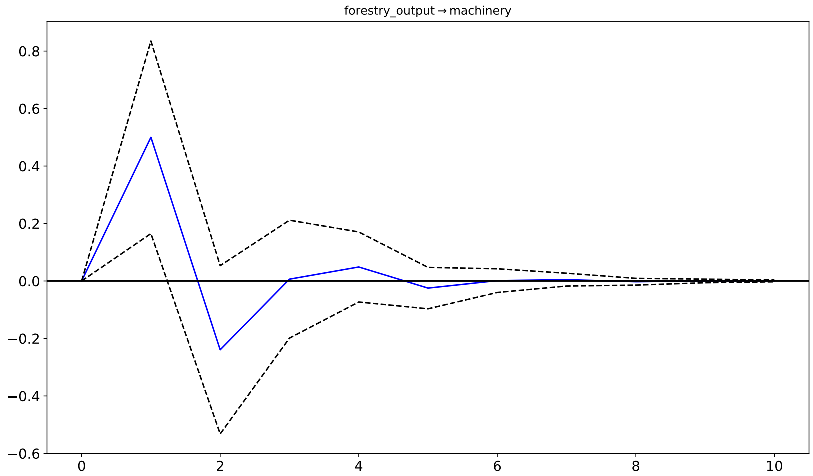


Figure 2 The response of investment to a shock in the output of the forest sector

The graph illustrates how a unit shock in the forestry sector's output influences the level of investment over time. The blue line depicts the response value, while the dashed line represents the 95% confidence interval. Observing the graph, it is evident that a sudden unit shock in the *forestry_output* leads to a response in *machinery* approximately 0.5 standard deviation after 1 year, stabilizing around 3 years thereafter.

Next, a hypothesis was tested to investigate how the output of the forestry sector and investments in machinery and equipment affect employment levels using another VAR model that includes all time series. Regarding the relationship between sector output and employment, a suitable equation could not be estimated because the statistically significant lag at order $p = 0$.

A statistically significant correlation for zero lag suggests that the need for increased logging and subsequent processing in year t will trigger an influx of additional working units already in the same year (t). A summary of the VAR model for the equations describing the dependence of working units on lagged values of other two variables and other dependencies is shown in Table 3.

Y	X	Coefficient	S.E.	p-value
employment	*const*	-0.092	0.171	0.589
	L1 – *machinery*	-0.037	0.017	0.041
	L1 – *employment*	-0.522	0.281	0.064
	L1 – *forestry_output*	0.002	0.003	0.476
forestry_output	*const*	17.189	15.052	0.253
	L1 – *machinery*	-3.437	1.575	0.029
	L1 – *employment*	-48.638	24.851	0.050
	L1 – *forestry_output*	0.342	0.307	0.265
machinery	*const*	2.893	1.562	0.064
	L1 – *machinery*	-0.303	0.163	0.063
	L1 – *employment*	3.893	2.580	0.131
	L1 – *forestry_output*	0.020	0.032	0.529

Table 3 VAR model for the relationship between employment, sector output and investments

The result of the model shows a statistically significant effect of the delayed *machinery* variable. The result suggests that increased investment in machinery and equipment at time $t-1$ results in a decrease in working units at time t. To verify this causality, a Granger causality test was subsequently performed with results shown in Table 4.

Null Hypothesis	Parameter	Results
machinery does not Grange-cause *employment*	Test statistic	4.186
	p-value	0.044
	Conclusion	Reject null hypothesis
employment does not Grange-cause *machinery*	Test statistic	2.279
	p-value	0.135
	Conclusion	Fail to reject null hypothesis

Table 4 Granger causality test results for employment and machinery investment

The results of the Granger causality test suggest a causal relationship between investments in machinery and equipment and the number of working units. No evidence of reverse causality was found. Finally, the reaction of the *employment* variable to a shock in the *machinery* variable was examined. The results are depicted in Figure 3. The graph illustrates that a sudden shock in the *machinery* variable can lead to a negative reaction in the *employment* variable, resulting in a decrease in working units in the following year.

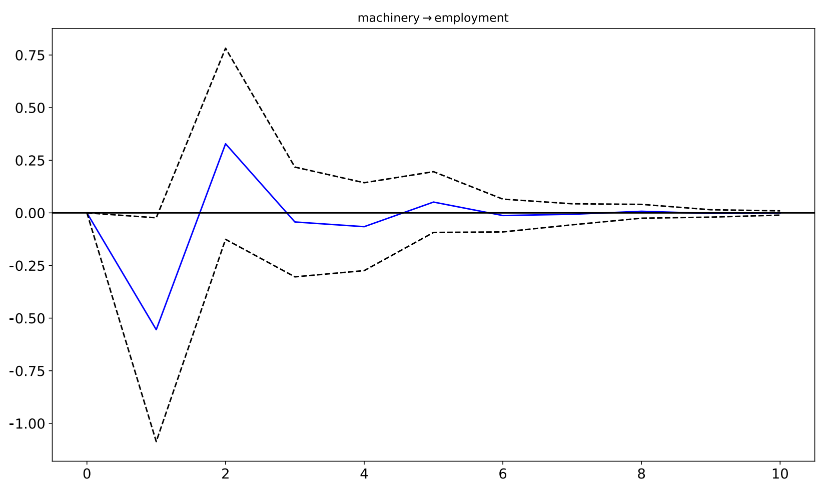


Figure 3 The response of employment to a shock in the machinery investment

4 Conclusion

In this paper, an analysis of three time series from the area of Swiss forestry was carried out with the aim of determining the mutual relations between these series, defining the process that generates the data of this system and describing its dynamics. The aim of the study was to quantify the mutual relations of these series and to test their causal relations.

Based on the estimated models, it was found that the growth in investments in machinery and equipment is due to the increased need for logging and subsequent processing of wood in the previous year. Sudden shocks caused by events such as bark beetle infestations or other calamities force producers to make additional investments to cover these damages. The paper also models the relationship between employment and the level of investment or sector output. The results show that the increased need for production stimulates the arrival of additional working units in the same year. Furthermore, the increase in investment in machinery and equipment causes a decrease in working units in the following period.

In the field of forestry, the methods described above are probably not often used. The proof is the fact that it was not possible to find a large number of studies that would deal at least partially with a similar problem in the area of the forest sector. For that reason, it is not easy to confront the obtained models with the works of other authors and thus verify the validity of the obtained model. However, an article (Hampel et al., 2021) demonstrated the fact that the increasing need for logging increases the wages of forest workers. This means that there is an increasing demand for these workers, which is in line with the findings in this article. Another research direction in this area could be to simulate Swiss forest economics as in (Janová and Hampel, 2016) and determine detailed information at the level of forest managers as in (Janová et al., 2022).

Acknowledgements

This article was supported by grant IGA24-PEF-DP-011 of the IGA PEF MENDELU Grant Agency.

References

Granger CW (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica: journal of the Econometric Society pp 424–438

Hamilton J (2020) Time Series Analysis. Princeton University Press, `https://books.google.cz/books?id=BeryDwAAQBAJ`

Hampel D, Viskotová L, Martiník A (2021) Determination of wages in forestry depending on the occurrence of natural disasters. In: Mathematical Methods in Economics 2021: Conference Proceedings. Česká zemědělská univerzita v Praze, Praha, pp 135–140

Janová J, Hampel D (2016) Optimal managing of forest structure using data simulated optimal control. Central European Journal of Operations Research 24(2):297–307

Janová J, Hampel D, Kadlec J, Vrška T (2022) Motivations behind the forest managers' decision making about mixed forests in the czech republic. Forest Policy and Economics 144:102841

Lütkepohl H (2005) New introduction to multiple time series analysis. Springer Science & Business Media

McCracken JM (2016) Exploratory causal analysis with time series data. Morgan & Claypool Publishers

Neusser K (2016) Time series econometrics. Springer

Scherrer D, Baltensweiler A, Bürgi M, Fischer C, Stadelmann G, Wohlgemuth T (2023) Low naturalness of swiss broadleaf forests increases their susceptibility to disturbances. Forest Ecology and Management 532

Sims CA (1980) Macroeconomics and reality. Econometrica: journal of the Econometric Society pp 1–48

Comparing Measures of Product Relatedness on Data from Czech Drugstore Retail Chain

Petr Krautwurm[1]

Abstract. This article examines Robinson's elasticity of substitution estimator alongside cross-price elasticity, from both theoretical and empirical perspectives, focusing on their application in classifying products as substitutes or complements. Initially, we illustrate the theoretical interconnections between these measures, demonstrating that they should consistently classify products similarly. Furthermore, we introduce a novel approach to adjust for the presence of perfect substitutes in individual transactional datasets, a common challenge in economic analyses. By aggregating perfect substitutes, we maintain the integrity of product relations, thus enhancing the utility of the data without introducing biases. Utilizing this adjusted dataset, we apply both measures to assess product relatedness and find that they effectively and consistently classify products as either substitutes or complements, validating theoretical predictions and our methodological innovations.

Keywords: Elasticity of substitution, Cross-price elasticity, Individual data, Substitutes, Complements

JEL Classification: C13, D01, D11
AMS Classification: 91B16

1 Introduction

Understanding the relationship between products from a consumer perspective—specifically, identifying whether products are complements or substitutes—is a fundamental question in economics. Two primary concepts are prevalent in this area: Robinson's Elasticity of Substitution (Blackorby and Russell, 1989; Robinson, 1933; Stern, 2011) and Cross-Price Elasticity (Auer and Papies, 2019; Deaton, 1987). Although these concepts are primarilly theoretetical, they can be translated into econometric world and estimated by relatively straightforward regression models.

Despite their theoretical utility, these concepts encounter challenges when applied to individual transactional data, primarily due to the lack of sufficient data to meet the hidden but strict requirements of the analytical techniques used. Historically, the emergence of large-scale datasets and advances in computational power shifted focus away from the limitations of small-scale datasets. However, this shift has obscured persistent issues in small datasets, particularly when analyzing individual perceptions of product substitutability and complementarity.

This article focuses on these challenges, offering a methodological adjustment that compensates for data limitations, and compares Robinson's Elasticity of Substitution with Cross-Price Elasticity using a small-scale dataset of individual transactions.

We begin by theoretically contrasting the two methods, followed by a detailed discussion of the dataset and our novel approach. The study concludes with an application of both methods to the dataset, comparing their outcomes to demonstrate that challenges in small-scale data analysis remain pertinent in the era of data abundance.

2 Elasticity of Substitution and Cross-Price Elasticity

In economic theory, two significant concepts used to analyze relations between products are Elasticity of Substitution, especially Robinson's (Robinson 1933), and Cross-Price Elasticity. These concepts can be applied to any utility function $U(\boldsymbol{x})$, obtaining theoretical results, as well as transformed into regression models for empirical data analysis.

For the purposes of our analysis, we will consider that the consumer who purchases n products faces linear constraint $P_1x_1 + \ldots P_nx_n \leq M$ during decision-making, where $\boldsymbol{P}$ represents prices of all relevant products and M is the budget.

[1] Prague University of Economics and Business, Department of Econometrics, W. Churchill sq. 1938/4, Prague, 130 67, The Czech Republic, petr.krautwurm@vse.cz

2.1 Definition with Respect to Utility Function

Robinson's Elasticity of Substitution is mathematically expressed as:

$$\sigma_{i,j} = \frac{\frac{\partial(x_j/x_i)}{x_j/x_i}}{\frac{\partial(MU_i/MU_j)}{MU_i/MU_j}} \tag{1}$$

where $MU_i = \partial U/\partial x_i$ represents the marginal utility of product i.

Conversely, Price Elasticity is defined as:

$$\zeta_{ij} = \frac{\frac{\partial x_i}{x_i}}{\frac{\partial P_j}{P_j}} \tag{2}$$

This metric differentiates between own-price elasticity ($j = i$) and cross-price elasticity ($j \neq i$).

These concepts are interconnected, which can be illustrated using the Constant Elasticity of Substitution (CES) function:

$$U(\boldsymbol{x}) = \left(\sum_i^n a_i x_i^\rho\right)^{\frac{1}{\rho}} \tag{3}$$

Here, x_i denotes the quantity of product i, a_i represents taste parameters, and $\rho \in (-\infty; 1)$ depicts product substitutability.

Applying Robinson's formula (1) to the CES function yields:

$$\sigma_{ij} = \frac{1}{1-\rho} \tag{4}$$

This result categorizes products as complements if $\rho < 0$ and as substitutes if $\rho \in (0,1)$. The scenario where $\rho > 1$ leads to negative elasticity, which actually corresponds to concave indifference curves. Since concave indifference curves represent consumer preferences for extremes over balanced product bundles and the consumers maximize their utility, the negative elasticity of substitution relates to attitudes toward products similar to perfect substitutes.

To apply Cross-Price Elasticity and compare it with Elasticity of Substitution, we need to consider the demand functions x_1^* and x_2^* under a CES utility function (3) with linear budget constraint. Demands are given as follows:

$$x_1^* = \frac{M}{P_1 + P_2\left(\frac{a_1}{a_2}\right)^{\frac{1}{\rho-1}}\left(\frac{P_2}{P_1}\right)^{\frac{1}{\rho-1}}} \qquad x_2^* = \frac{M}{P_1\left(\frac{a_2}{a_1}\right)^{\frac{1}{\rho-1}}\left(\frac{P_1}{P_2}\right)^{\frac{1}{\rho-1}} + P_2} \tag{5}$$

Analyzing how x_2^* varies with P_1 using Cross-Price Elasticity (2) results in:

$$\zeta_{21} = \frac{\partial x_2^*}{\partial P_1}\frac{P_1}{x_2^*} = -\frac{\rho}{\rho-1}\frac{\left(\frac{a_2}{a_1}\right)^{\frac{1}{\rho-1}}\left(\frac{P_1^\rho}{P_2}\right)^{\frac{1}{\rho-1}-1}P_1^\rho}{P_1\left(\frac{a_2}{a_1}\right)^{\frac{1}{\rho-1}}\left(\frac{P_1}{P_2}\right)^{\frac{1}{\rho-1}} + P_2} \tag{6}$$

This analysis reveals that if $\rho \in (0,1)$, both products are substitutes and $\zeta_{ij} > 0$, and if $\rho < 0$, they are complements and $\zeta_{ij} < 0$, aligning with the elasticity of substitution results (4).

A crucial distinction between both measures exists: Elasticity of Substitution is derived solely from consumer preferences, requiring no price information and providing a universal view of consumer attitudes toward products. Since the concepts of complementarity and substitutability are relevant even in scenarios, where prices are either low or not well defined, such as being stranded on an abandoned island, definitions of substitutes and complements based solely on consumer preferences offer a more fundamental understanding of consumer behavior. From a theoretical perspective, this is an essential advantage of Elasticity of Substitution over Cross-Price Elasticity.

2.2 Regression Models

Both Robinson's Elasticity of Substitution (1) and Cross-Price Elasticity (2) can be estimated through regression analysis, requiring several key novel assumptions (de Lima, 2019; Deaton, 1987). Since these measures stem from consumer decision-making during optimization, they presuppose access to comprehensive data at the moment of purchase, including any special discounts and price considerations.

Modeling Cross-Price Elasticity

For Cross-Price Elasticity, observing the quantity of a product purchased and all relevant prices is essential. The associated regression model is formulated as follows:

$$\ln(x_i) = \alpha_0 + \zeta_{ii} \cdot \ln(P_i) + \sum_{j \neq i}^{n} \zeta_{ij} \cdot \ln(P_j) + \beta \ln(M) + u \tag{7}$$

Here, i and j represent different products, x_i is the quantity of product i purchased, P_j is the price of product j, M denotes the consumer's budget at the time of decision, and u is the stochastic error term.

This model faces challenges such as simultaneity bias (Graddy and Kennedy, 2010), where firms adjust prices in response to consumer demand and vice versa, often leading to endogeneity in the estimates (Auer and Papies, 2019; Meyer and Miller, 1954). To address this, instrumental variables like production costs are used (Angrist and Krueger, 2001), though they are not always perfect as they can be in specific markets correlated with demand (Davidoff, 2016).

Additionally, the symmetry problem arises with n products, generating two potentially unequal estimates of cross-price elasticity for each pair (ζ_{ij} and ζ_{ji}). Thus, we estimate them as a system, imposing symmetry constraints on the parameters.

Another complication involves the budget variable M, which is typically unobservable and must be inferred from the total spending ($P_1 x_1 + \cdots + P_n x_n$), assuming full budget expenditure.

Modeling Elasticity of Substitution

Direct empirical evaluation of Elasticity of Substitution is challenging due to the unobservable nature of marginal utility (de Lima, 2019). This can be solved by assuming that consumers allocate their entire budget optimally and it is binding. This leads to a second Gossen's law ($MU_i/MU_j = P_i/P_j$), implying that the ratio of marginal utilities can be substituted by the ratio of prices, resulting in the following regression model:

$$\ln\left(\frac{x_j}{x_i}\right) = \gamma_0 + \sigma_{ij} \cdot \ln\left(\frac{P_i}{P_j}\right) + v \tag{8}$$

In this model, γ_0 represents a constant, capturing the intrinsic preference ratio (a_j/a_i) when prices are equal, and v is the random error term. Unlike cross-price elasticity, this model ensures symmetry ($\sigma_{ij} = \sigma_{ji}$) and is strictly pairwise.

As outlined in a previous subsection, negative σ_{ij} corresponds to perfect substitutes, implying an exclusive preference for one product, which is usually contradictory with observed multi-product purchases. Additionally, from the econometric perspective, a negative σ_{ij} paradoxically aligns with the definition of perfect complements, since an increase in the price of product P_i leads to a decreased demand for product x_j, mirroring the dynamics dictated by cross-price elasticity. This apparent contradiction stems from the application of Gossen's law under the assumption that non-negativity constraints are redundant. Yet, this assumption fails for perfect substitutes, as the marginal utility ratio (MU_i/MU_j) does not equate to the price ratio (P_i/P_j). Consequently, negative estimates of σ_{ij} might suggest noisy data.

Importantly, Elasticity of Substitution has an advantage over Cross-Price Elasticity when applied to individual transaction data as it deals with ratios, which can enhance the model's robustness by allowing for the analysis of product relationships even when there is zero variance in one of the variables.

3 The Dataset

The dataset under examination contains transactional data from a Czech drugstore retail chain, detailing consumer purchases of the 30 most frequently bought products in 2019. It includes records from 62 322 unique customers, accumulating a total of 643 450 instances of products being bought. Each customer in this dataset purchased at least three of these top products, and the purchases were tracked using personal ID cards, enabling data linkage.

Crucially, all prices in the dataset reflect personalized discounts and shopping visit characteristics, such as "buy one, get one free" offers or account-specific discounts. This ensures that the prices recorded are those actually encountered by consumers at the point of sale, which is vital for accurately applying measures of product relatedness like Elasticity of Substitution and Cross-Price Elasticity.

3.1 Characteristics and Challenges of the Dataset

Despite its robustness, the dataset presents several challenges:

1. **Non-simultaneous Purchases:** Often, transactions involve the purchase of a single product rather than multiple products simultaneously, resulting in estimation of both the Elasticity of Substitution and Cross-Price Elasticity not being viable. Transactions where only one product was bought at a time are therefore excluded from the analysis.
2. **Lack of Variance:** Some consumers consistently purchase the same quantity of a particular product. This lack of variance makes estimating Cross-Price Elasticity challenging, as variability in the dependent variable is necessary for robust estimation. Consequently, all data points with zero variance in purchase quantity are excluded from the analysis.
3. **Constant Price and Product Ratios:** The Elasticity of Substitution requires variability in the price and product ratios. Instances where the price or product ratio remains constant are excluded from the analysis.
4. **Dealing with Endogeneity:** A key challenge in this kind of analysis is the potential endogeneity in price, where changes in consumer demand might influence prices, due to the simultaneous nature of the problem. However, it is noted that the pricing policy during 2019 of this particular drugstore retail chain was predominantly driven by suppliers rather than demand dynamics, suggesting minimal in prices. Unfortunately, typical instrumental variables such as production costs are not available, limiting the scope for addressing potential endogeneity issues more robustly.

Despite these challenges, the dataset's detailed transactional records and personalized pricing data provide a unique opportunity to analyze individual consumer behavior.

4 Adjusting the dataset:

Given the complexities associated with the dataset described, it is crucial to explore strategies for increasing the variance in the variables and inferring additional information about prices. We propose an approach suitable when dealing with perfect substitutes.

4.1 Perfect substitutes

Product substitutability can be understood as the consumer's perceived sameness of products. Perfect substitutes are then products perceived by consumers as identical, rendering them interchangeable. This perception is aligned with the Constant Elasticity of Substitution (CES) function at $\rho = 1$, where utility becomes additive: $U(\boldsymbol{x}) = a_1x_1 + \cdots + a_nx_n$. If the taste parameters differ, say $a_1 > a_2$, the lower-taste product behaves as though it were merely a lesser quantity of the one identical product the consumer demands.

Identifying perfect substitutes in datasets is challenging as consumers typically do not purchase such items together unless their price and taste ratios are equivalent, in which situation the consumers are indifferent between buying only one product or a combination of those. Thus, certain products, which have perfect substitutes to them, usually appear missing from datasets, posing a significant obstacle in analyzing product relations.

Importantly, from an analytical standpoint, perfect substitutes can be aggregated into a single composite good. If perfect substitutes are aggregated, the elasticity of substitution (1) between any product within this composite group and any product outside it remains constant, simplifying the analysis. Thus, aggregating solely perfect substitutes does not change the relations between products. This ensues from the fact that under imbalanced taste and price ratio, only one product will be effectively demanded, reducing the complex utility function with perfect substitutes back to CES function (3).

4.2 Equal prices

In scenarios where products are priced equally, consumers will be indifferent to the proportions in which they purchase them, provided the tastes are balanced. Remarkably, if the products are near-perfect substitutes $\rho < 1$, they behave the same as perfect complements under balanced price and taste ratios. This could manifest in datasets as consumers purchasing fixed quantities of each product during a single shopping event. Thus, we are even theoretically able to observe multiple potential perfect substitutes being bought together if price and taste ratio are balanced.

This is a necessary information to understand why we can observe multiple perfect substitutes being purchased together, allowing us to compare their purchasing patterns and prices. Additionally, if multiple products consistently share the same price, this can serve as a basis for inferring price information about other items, thereby expanding the usable dataset.

4.3 Combining "Hello" food pouches

Within our dataset, four products emerge as potential perfect substitutes: "Hello" food pouches in Strawberry, Banana, Apple, and Apricot flavors. These products are generally suited for children over four months, except for the Strawberry flavor, which is intended for those over six months. This minor deviation does not significantly impact their substitutability, except for a narrow age group.

Given that these products are intended to supplement children's diets as they transition to solid foods, parents may view them as interchangeable in achieving this objective, with flavor variations interaction only with personal tastes. Additionally, these products maintained identical pricing throughout 2019 and their purchasing patterns were similar within the decision-making of individual consumers. However, further investigation, involving more rigorous testing of this hypothesis, is necessary.

To capitalize on this observation, we propose a method of combining these "Hello" food pouches into a single category, treating them as a composite good. This approach not only increases the number of observations but also enhances the variance in the dependent variables, which is crucial for robust econometric analysis.

5 Results

Utilizing the adjusted dataset, we examined the purchase history of each consumer to identify patterns in product consumption. Specifically, we included consumers in our analysis who purchased at least two products together on at least six different occasions. We further limited our data to those transactions where the prices, quantities, and price and quantity ratios exhibited non-zero variance throughout the purchasing period.

Out of 62,322 total customers, only 181 combinations of product purchases met our criteria, with an average of 8.6 instances of purchasing products together per customer and a maximum of 46 observations for a single combination.

We applied the regression models (8) and (7) to these data points, estimating both with and without imposing a zero constraint on the constants. Our analysis identified 18 cases of statistically significant negative σ_{ij} when constants were not included, reduced to 3 cases when constants were included. This suggests that excluding the constant from the model, both theoretically and empirically, does not enhance the model's effectiveness and tends to increase discrepancies in the outcomes of both methods.

The contingency table below (Table 1) presents the classification of these interactions. In the table, **N** indicates noisy observations with respect to Elasticity of Substitution, while **S**, **I**, and **C** represent classifications where the appropriate method identified the interactions as statistically significant substitutes, complements, or statistically insignificant from zero, respectively.

		EOS			
		N	**S**	**I**	**C**
XPE	**S**	0	1	3	0
	I	0	12	156	1
	C	3	0	5	0

Table 1 Contingency table: Full results with insignificance measure

A notable finding from our analysis is that many observations were not statistically significant from zero. Given the different scales of measurement, with $\zeta_{ij} \in \mathbb{R}$ and $\sigma_{ij} \in \mathbb{R}_{\geq}$, a non-significant result has differing implications for each measure. For Cross-Price Elasticity, a non-significant result implies a Cobb-Douglas-like relationship (neither substitutes nor complements), whereas for Elasticity of Substitution, it suggests that products are indistinguishable from perfect complements. Yet, even this analysis is valuable to illustrate how many observation of negative σ_{ij} are actually indistinguishable from zero; thus, negative elasticity of substitution may just be negative by a pure chance. In future research, a Non-negative least squares method may be used for estimating elasticity of substitution.

Because of this, we analyzed results further, eliminating all observations for which the $\sigma_{ij} < 0$ and then just compared point estimates of these methods together without paying attention to their significance. The resulting contingency table in table 2 descrives these results:

We refined our analysis by excluding all observations where $\sigma_{ij} < 0$ and compared the point estimates of both methods without considering their statistical significance. The results are presented in the following contingency table (Table 2):

		EOS	
		S	**C**
XPE	**S**	73	16
	C	14	23

Table 2 Contingency table: Identifying substitutes and complements

The chi-square test and analysis of adjusted standard residuals confirm that the classification of products as either substitutes or complements by both methods is statistically significant from a random distribution. Interestingly, the methods appear less consistent in identifying complements, primarily due to the exclusion of negative σ_{ij}. However, 46 out of 55 cases of negative σ_{ij} were identified as complements by Cross-Price Elasticity, underscoring a divergence in econometric interpretation from the theoretical interpretation. This suggests that while negative values of Elasticity of Substitution can hint at complementary relationships, the theoretical underpinnings require further investigation to align empirical findings with theoretical constructs. Further research should explore the relationship between negative empirical Elasticity of Substitution and its theoretical implications, particularly in the context of perfect substitutes where Gossen's law may not apply.

Acknowledgements

This work was supported by The Internal Grant Agency of Prague University of Economics and Business [VŠE IGS F4/52/2023]

References

Angrist J D, Krueger A B (2001) Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. Journal of Economic Perspectives 15(4): 69–85.

Auer J, Papies D (2019) Cross-price elasticities and their determinants: a meta-analysis and new empirical generalizations. Journal of the Academy of Marketing Science 48(3): 584–605.

Blackorby C, Russell R R (1989) Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities). The American Economic Review 79(4): 882–888.

Davidoff T (2016) Supply Constraints Are Not Valid Instrumental Variables for Home Prices Because They Are Correlated With Many Demand Factors. Critical Finance Review 5(2): 177–206.

de Lima J E (2019) Alternative definitions of elasticity of substitution: review and application. Revista de Economia e Sociologia Rural 38(1): 119-153.

Deaton A (1987) Estimation of own- and cross-price elasticities from household survey data. Journal of Econometrics 36(1–2): 7–30.

Graddy K, Kennedy P (2010) When Are Supply And Demand Determined Recursively Rather Than Simultaneously? Eastern Economic Journal 36(2): 188–197.

Meyer J R, Miller H L (1954) Some Comments on the "Simultaneous-Equations Approach." The Review of Economics and Statistics 36(1): 88-92.

Robinson J (1933) The Economics of Imperfect Competition (1st edn). London, Macmillan.

Stern D I (2011) Elasticities of substitution and complementarity. Journal of Productivity Analysis 36(1): 79–89.

Economic Growth and Agricultural Sector Dynamics in the Visegrad Group: A Panel Analysis

Radmila Krkošková[1], Zuzana Neničková[2], Lucie Waleczek Zotyková[3]

Abstract. This empirical investigation analyzes the interplay between employment in the agricultural sector, food production index, and economic growth within the Visegrad Group. Utilizing panel analysis, the study investigates annual time series data spanning from 2005 to 2023. Employing panel data analysis and the Autoregressive Distributed Lag (ARDL) model, the study aims to clarify the research objectives. Findings reveal that, in the short and long terms, the food production index significantly impacts economic growth. Specifically, an increase in the food production index correlates with boosted economic growth. Conversely, a decline in the agricultural sector labor force tends to spur economic growth.

The agricultural sector continues to be a vital pillar of economic growth in the Visegrad Group. Understanding the dynamics, challenges, and opportunities within this sector is essential for policymakers, stakeholders, and investors to formulate strategies that promote sustainable development across the countries.

Keywords: agricultural, ARDL, economic growth, food production index, panel analysis

JEL Classification: C01, Q10, Q19
AMS Classification: 91C99

1 Introduction

The relationship between economic growth and employment in the agricultural sector within EU countries is a key issue in studies focusing on economic development and labour market structure within the European Union. This relationship is complex and can be influenced by various factors. The agricultural sector can be a key component of the economy of certain countries, and economic growth in these countries can be closely linked to agricultural development. Increased agricultural production can lead to GDP growth and increased employment in the sector. Another influencing factor is that high economic growth can lead to structural changes in the economy, often resulting in a shift towards the industrial and service sectors, which may reduce employment in agriculture. Technological progress is also important, as the modernisation of agriculture and the introduction of new technologies can increase productivity, allowing economic growth without necessarily increasing employment in the agricultural sector. Rising urban populations and rural-urban migration also play a crucial role, often leading to a reduction in agricultural employment as people seek employment opportunities in other sectors in urban areas. Finally, the impact of policies and support programmes cannot be underestimated. Subsidies, investment in infrastructure and education can support growth and employment in the agricultural sector. In summary, the relationship between economic growth and employment in the agricultural sector is dynamic and influenced by a variety of factors. It depends on the specific conditions and policies of each country within the EU.

The following hypotheses were formulated to fulfil the main objective. The core of this paper is to analyse two hypotheses concerning the relationship between GDP and the employment in agriculture (EMAG), and the food production index (FPI) in the V4 group.

During economic development, countries often transition their workforce from agriculture to industry and services. This shift can result in a decline in agricultural employment, but simultaneously lead to growth in the industrial and service sectors, which contributes to GDP growth. This statement presents the initial hypothesis.

[1] School of Business Administration in Karviná, Silesian University in Opava, Department of Informatics and Mathematics, Univerzitní náměstí 1934/3, 733 40 Karviná, Czech Republic, e-mail: krkoskova@opf.slu.cz
[2] School of Business Administration in Karviná, Silesian University in Opava, Department of Informatics and Mathematics, Univerzitní náměstí 1934/3, 733 40 Karviná, Czech Republic, e-mail: nenickova@opf.slu.cz
[3] School of Business Administration in Karviná, Silesian University in Opava, Department of Informatics and Mathematics, Univerzitní náměstí 1934/3, 733 40 Karviná, Czech Republic, e-mail: waleczekzotykova@opf.slu.cz

H1: There is a negative relationship between GDP and EMAG in the V4 countries.

The hypothesis assumes a positive correlation between GDP and the food production index. This is because economic growth leads to an increase in the standard of living, resulting in higher demand for food and increased production. This is the second hypothesis.

H2: There is a positive relationship between GDP and FPI in the V4 countries.

2 Literature Review and Data

2.1 Literature Review

The ongoing debate on the role of agriculture in the economy and society has been a topic of interest for over 250 years. Lains and Pinilla (2009) posit that this debate primarily takes place within the fields of economic theory and economic history. Economic theory examines agriculture from various perspectives, and there is no single unified approach to studying.

Research suggests that there is a negative relationship between the relative GDP share of agriculture and the unemployment rate in Central and Eastern European countries, indicating that higher agricultural GDP shares may lower unemployment rates (Bartóková, 2019). The paper analyzes the agriculture sector in V4 countries, expecting a decline in production and employment from 2000-2014, except for Poland showing relatively stronger domestic position.

Additionally, studies on India show that employment may not be a significant factor affecting the GDP contribution of sectors, with GDP growth potentially influenced by technological advancements and quality parameters (Bein and Ciftcioglu, 2017). Authors claim that the relative GDP share of agriculture is negatively related to unemployment rates in Central and Eastern European countries, suggesting a potential role for agriculture in lowering unemployment.

Furthermore, in China, the employment in the primary sector has been found to have a dynamic relationship with the total output value index of farming, with changes from negative to positive correlations over time (Agarwal et al., 2019).

Zhu (2023) investigated the relationship between the employment in the primary sector and the total output value index of farming, verifying the effect of labor factor on the relationship and showed that the relationship changes from negative to positive.

In Europe, the relationship between Gross Domestic Product and employment in the agricultural sector is influenced by various factors. Economies with larger agricultural sectors tend to have smoother fluctuations in aggregate employment and agricultural employment is less correlated with overall GDP (Sancar and Sancar, 2017). Over time, the share of the agricultural sector in GDP and employment has been declining (Ali et al., 2021). This decline aligns with the structural changes observed in high-income regions like Europe, characterized by a shift away from agriculture towards urbanization and jobs in the modern industrial sector (Ramachandran, 2021).

Research indicates a bilateral causality relationship between GDP and employment rates in different sectors, including agriculture (Alonso-Carrera and Raurich, 2018). Factors such as labour mobility, structural changes, and wage disparities between the agriculture and non-agriculture sectors affect the employment share in agriculture (Pfister, 2022). Moreover, the agricultural sector's contribution to GDP and employment can vary based on factors like sectoral prices, output, and market demand for agricultural labour (Farag and Ab-Rahim, 2022).

Research has indicated that factors such as the share of agriculture in national exports and structural changes away from agriculture influence the shares of agricultural GDP and employment (Anderson and Ponnusamy, 2022). Furthermore, a study on the innovative development of the agricultural sector in the Visegrad Group countries offers insights into the trends and dynamics shaping agriculture in these regions (Burdiuzha, 2021). These studies emphasize the importance of comprehending the evolving relationship between GDP and employment in the agricultural sector within the V4 countries.

2.2 Data and Methods

Data

In this research, computations were based on yearly data ranging from 2005 to 2023 due to the unavailability of quarterly data, prompting the utilization of panel data analysis. The information was sourced from The World Bank, and models pertaining to economic expansion will be developed. According to economic theory and existing literature, factors such as employment within the agricultural industry or the food production index may impact economic growth. A detailed explanation of the variables utilized in this analysis is presented in Table 1, with each variable being adjusted for seasonal variations.

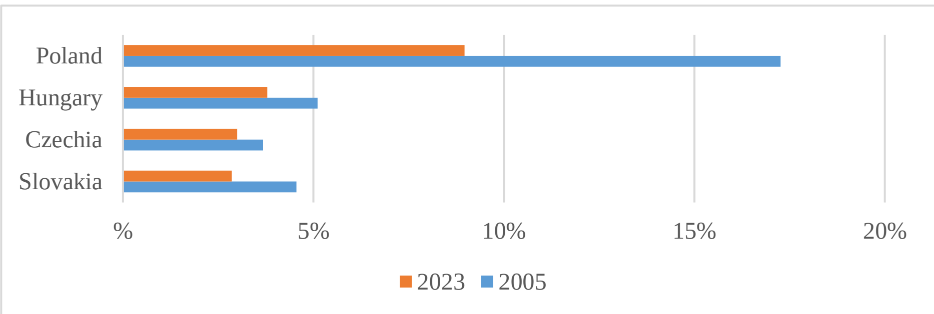
Variables designation		Description of variables
GDP (current US$)	GDP	The gross value added by all resident producers in the economy, plus any product taxes and less any subsidies not included in the value of the items, is the GDP at purchaser's prices. Current U.S. dollars are used for data. GDP numbers expressed in dollars are derived from local currencies using official exchange rates valid for a single year.
Employment in agriculture (% of total employment)	EMAG	People of working age who have engaged in any activity to create things or perform services for compensation or profit are considered to be employed. Activities related to agriculture, hunting, forestry, and fishing are included in the agriculture sector.
Food production index (2014-2016 = 100)	FPI	Nutrient-containing, edible food crops are included in the food production index. Tea and coffee are not included because they are edible but lack nutritional value.

Source: The World Bank (2024)

Table 1 Description of variables

The evolution of employment within the agricultural domain in the European Union during the period 2005-2023 has been subject to transformations influenced by a myriad of factors. The workforce engaged in agriculture has experienced a gradual decrease over time, a trend attributed to the progression of economies towards industrialization and the expansion of service sectors. Historically a prominent source of employment, agriculture has undergone a shift in relevance as most economies have transitioned into industrial and urbanized entities. Across numerous nations, the agricultural sector now represents a minor segment of the overall labor force. Moreover, the integration of advanced technology and modernization initiatives within the agricultural sphere have led to enhanced productivity levels with a reduced workforce requirement. The implementation of automation and novel technological solutions has significantly diminished the necessity for manual labor in agricultural activities. It is noteworthy that agriculture remains a sector heavily influenced by seasonal variations.

All things considered, it can be said that employment in agriculture in the V4 group is decreasing (Figure 1), and the sector is becoming less labour-intensive.



Source: Eurostat, https://ec.europa.eu/eurostat/databrowser/view/NAMA_10_A64_E__custom_7249372/default/table?lang=en

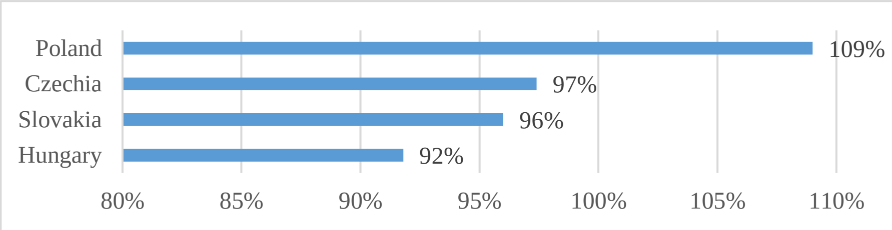
Figure 1 Employment in agricultural (% of total employment, 2005 and 2023)

The Food Production Index (FPI) serves as a metric employed to monitor fluctuations in the quantity of food production across time. The utility of the Food Production Index lies in its role as a gauge of food security and a nation's capacity to fulfill the nutritional requirements of its populace. Enhanced food security is often associated with a heightened index value, indicating increased production and accessibility of food resources. Conversely, a diminished index value may highlight obstacles in agricultural output and potential deficiencies in food supply.

The formula to calculate the Food Production Index involves comparing the current year's agricultural output to a base period:

$$Food\ Production\ Index = \frac{Agricultural\ Output\ in\ Current\ Year}{Agricultural\ Output\ in\ Base\ Period\ 2014{-}2016} x100 \quad (1)$$

This index can be calculated at the national, regional, or global level and is used by policymakers, researchers, and international organizations to monitor and analyse trends in food production. Figure 2 shows that the highest FPI in 2023 had Poland.



Source: Eurostat, https://ec.europa.eu/eurostat/databrowser/view/NAMA_10_A64_E__custom_7249372/default/table?lang=en

Figure 2 Food production index (%) in 2023

Methods – ARDL model

In the present article, the Autoregressive Distributed Lag (ARDL) model, as formulated by Pesaran et al. (2011), is utilized. A cointegration test based on the ARDL framework is employed to ascertain the presence of a sustained relationship among the designated variables. The ARDL model constructs a singular reduced form equation. In the context of ARDL methodology, the conditional error correction model (ECM) is represented as delineated by Turan et al. (2020).

$$\Delta GDP_t = \beta_0 + \sum_{i=1}^{I} \beta_i \Delta GDP_{t-i} + \sum_{i=1}^{I} \beta_i \Delta EMAG_{t-i} + \sum_{i=1}^{I} \beta_i \Delta FPI_{t-i} + \\ + \theta_1 GDP_{t-1} + \theta_2 EMAG_{t-1} + \theta_3 FPI_{t-1} + \varepsilon_t \quad (2)$$

The ECM comprises of short-term dynamics and the cointegration equation. Over the long term, the short-term dynamics are eliminated, leaving only the cointegrating, or equilibrating equation. As indicated by Pesaran et al. (2011), the presence of a cointegrating relationship is assessed using the F-test. Upon discovering cointegration, the short and long-term coefficients are examined. ARDL theory is elucidated, as seen in the work of Greena (2002).

3 Data Analysis

Stationarity tests can be categorized into two groups based on their underlying assumptions. The first group assumes cross-sectional independence of data, which includes the majority of commonly used tests. The second group rejects the assumption of cross-sectional independence of data. In this study, tests that assume cross-sectional independence are used, specifically the tests by Levin, Lin, and Chu (2002) – LLC test, Im, Pesaran, and Shin (2003) – IPS test, Maddala and Wu (2002), Choi (2001) – Fisher-ADF test, and Fisher-PP test. Although these tests are referred to as panel unit root tests, they essentially involve parallel application of individual time series unit root tests within the panel data structure.

All variables GDP, EMAG and FPI are non-stationary and have an integration order of one, i.e., I (1).

The equation for the long-run relationship is presented in Table 2. Employment in agriculture affects GDP in a negative way, while the food production index has a positive relationship with GDP. The coefficients of the long-term relationship are statistically significant for both the EMAG variable and the FPI variable.

In the equation for the short-run relationship, the values of the COINTEQ correction terms is negative, which expresses the return of the variables to the long-run equilibrium. The system adjusts to long-term equilibrium at a rate of 71.2 % per year. In other words, it would take 1.4 years (1/0.712) to completely eliminate instability. From a short-term perspective, there is a significant negative relationship between FPI and GDP, while a significant positive relationship between EMAG and GDP was not confirmed.

Variable	Coefficient	Std. Error	t-Statistic	Significance
		Long Run Equation		
EMAG	–22.5	1.72	–13.08	0.000***
FPI	2.12	0.41	5.17	0.000***
		Short Run Equation		
COINTEQ	–0.712	0.21	–3.39	0.004**
D(GDP(-1))	0.287	0.34	0.84	0.451
D(EMAG)	–24.71	25.21	–0.98	0.467
D(EMAG(-1))	20.22	18.18	1.11	0.262
D(EMAG(-2))	17.58	10.73	1.64	0.219
D(FPI)	–1.75	0.51	–3.43	0.003**
D(FPI(-1))	–1.47	0.63	–2.33	0.029*
D(FPI(-2))	–1.25	0.61	–2.05	0.038*
C	63.23	21.33	2.96	0.006**

Statistical significance at the level 0.001 (***), 0.01 (**), 0.05 (*)
Source: The World Bank, EViews, own calculation

Table 2 Model ARDL (2, 3, 3) – V4 group – Dependent Variable D (GDP)

4 Conclusion

Smil (2021) points out that agriculture is considered a peripheral and uninteresting sector for economists because it constitutes only a negligible portion of the GDP in developed countries. However, he suggests that these economists should try consuming products and services from the finance, insurance, real estate, rental, and leasing sectors. Without agriculture, life is not possible.

The correlation between GDP and employment in the agricultural sector is intricate and subject to various factors and circumstances in diverse nations. Within developing nations, there is a tendency for a higher rate of employment in agriculture, often representing a substantial segment of the labor force. In these instances, the expansion of agriculture can yield a favorable effect on GDP as a greater number of individuals are engaged in the production of food and other agricultural goods. Conversely, in developed nations, agriculture typically assumes a lesser role in the economic landscape, with lower levels of employment in this particular sector. Consequently, GDP growth in these countries is more significantly influenced by the industrial and service sectors.

Model shows a negative relationship between GDP and employment in the agricultural sector (Hypothesis H1 was accepted) and also there is a positive relationship between GDP and FPI in the V4 countries (Hypothesis H2 was accepted).

The modernization of agriculture and the consequent enhancement of its productivity lead to a reduction in the workforce required to maintain the same level of food output. As a result, there is a decrease in agricultural employment, yet there is a potential for an augmentation in GDP due to the amplified production of agricultural goods. Ultimately, the correlation between GDP and employment in the agricultural sector is not fixed and is contingent upon numerous variables, such as the level of economic advancement, agricultural efficiency, and governmental regulations.

The correlation between GDP and the food production index serves as a pivotal gauge for evaluating the economic and agricultural well-being of a specific nation. Considering all factors, it can be posited that the connection between GDP and the food production index is intricate and reliant on various elements, encompassing economic progression, agricultural efficiency, governmental directives, and global market circumstances. This correlation may exhibit positivity when food production escalates, thereby bolstering GDP expansion, but it is also susceptible to external influences like drought, climatic occurrences, and alterations in market dynamics.

Acknowledgements

This research was supported by the Ministry of Education, Youth and Sports Czech Republic within the Institutional Support for Long-term Development of a Research Organization in 2024.

References

Agarwal S, Gupta D, Verma P (2019) Impact of Employment on GDP Contribution of Various Sectors in India. Global Journal of Enterprise Information System 11(1):47–53.

Ali N, Hayat U, Khan M, Khattak S (2021) Sources of agricultural productivity growth in SAARC countries: the role of financial development, trade openness and human capital. Sarhad Journal of Agriculture 37(2). https://doi.org/10.17582/journal.sja/2021/37.2.586.593

Alonso-Carrera J, Raurich X (2018) Labor mobility, structural change and economic growth. Journal of Macroeconomics 56:292–310. https://doi.org/10.1016/j.jmacro.2018.03.002

Anderson K, Ponnusamy S (2022) Structural transformation away from agriculture in growing open economies. Agricultural Economics 54(1):62–76. https://doi.org/10.1111/agec.12745

Bartóková L (2019) The Position of Agriculture and Food Sector in V4 Countries. AGRIS on-line Papers in Economics and Informatics 11(3):13–22. https://doi.org/10.22004/AG.ECON.294161

Bein M. A, Çiftçioğlu S (2017) The relationship between the relative GDP share of agriculture and the unemployment rate in selected Central and Eastern European countries. Agricultural Economics-zemedelska Ekonomika 63(7):308–317. https://doi.org/10.17221/372/2015-AGRICECON

Burdiuzha A (2021) Innovative development of the agricultural sector in the visegrad group countries. Economic and Regional Studies / Studia Ekonomiczne I Regionalne 14(1):1–13. https://doi.org/10.2478/ers-2021-0001

Eurostat (2024). Database. Available at: https://ec.europa.eu/eurostat/web/main/data/database

Farag F, Ab-Rahim R (2022) An empirical analysis of causal relationships between agricultural labour and agricultural GDP in Libya. International Journal of Academic Research in Business and Social Sciences 12(9):1565–1571. https://doi.org/10.6007/ijarbss/v12-i9/14747

Greene W.H (2002) Econometric Analysis. (5th ed.). Prentice Hall, Upper Saddle River

Im K.S, Pesaran M.H, Shin Y (2003) Testing for unit roots in heterogenous panels. Journal of Econometrics 115(1):53–74. https://doi.org/10.1016/S0304-4076(03)00092-7.

Lains P, Pinilla V (2008) Agriculture and Economic Development in Europe Since 1870. (1st ed.). London, Routledge. https://doi.org/10.4324/9780203928776

Levin A, Lin C.F, Chu C (2002) Unit root tests in panel data: asymptotic and finite sample properties. Journal of Econometrics 108:1–24. https://doi.org/10.1016/S0304-4076(01)00098-7

Maddala G.S, Wu S (2002) A comparative study of unit root tests with panel data and new simple test. Oxford Bulletin of Economics and Statistics 61:631–652. https://doi.org/10.1111/1468-0084.0610s1631

Pesaran M. H, Shin Y, Smith R (2011) Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics 16(3):289–326. https://doi.org/10.1002/jae.616

Pfister U (2022) Economic growth in Germany, 1500–1850. The Journal of Economic History 82(4):1071–1107. https://doi.org/10.1017/s002205072200033x

Ramachandran V (2021) Convergence, development, and energy-intensive infrastructure in Africa: a review of the evidence. Sustainability 13(19),10572. https://doi.org/10.3390/su131910572

Sancar C, Sancar C (2017) Is urban employment a potential for economic growth? (the case of EU countries and Turkey). Uluslararası İktisadi Ve İdari İncelemeler Dergisi. https://doi.org/10.18092/ulikidince.287053

Smil V (2021) Velké Proměny: Jak se utvářel moderní svět. Kniha Zlín, Praha, pp 106–107

The World Bank. (2024). Data bank. World development indicators. Available at: https://databank.worldbank.org/source/world-development-indicators/preview/on

Turan T, Karakas M, Ozer H. A (2020) How do oil price changes affect the current account balance? Evidence from co-integration and causality tests. Journal of Economics 68(1):55–68.

Zhu Z (2023) The relationship between rural employment rate and rural development: An analysis based on a bivariate VAR model. BCP business & management 38:104–109. https://doi.org/10.54691/bcpbm.v38i.3675

From Petrochemicals to Produce: Unveiling the Fertilizer-Crude Oil Nexus

Oleh Kurinnyi[1], Lukáš Frýd[2]

Abstract. The study explores the intricate relationship between crude oil and global fertilizer prices, highlighting fertilizer's significant role in food commodity costs. While prior research has emphasized crude oil's impact, this study unveils the extensive economic and social repercussions of fertilizer market shocks. Investigating potential asymmetric price transmission, the study employs a Nonlinear Autoregressive Distributed Lag (NARDL) model to analyze the nonlinear dynamics between spot prices of various fertilizers and crude oil. By utilizing the Bai-Perron method to identify structural changes, the study unveils both symmetric and asymmetric responses in fertilizer prices over different time periods, underscoring the dynamic nature of this crucial relationship.

Keywords: NARDL, rockets-and-feathers effect, breakpoints, asymmetric

JEL Classification: C22, O13, Q18
AMS Classification: 91B84

1 Introduction

The "rockets and feathers" phenomenon, coined by economist Robert W. Bacon in Bacon 1991, used to originally describe the asymmetric adjustment of gasoline prices to crude oil prices. Gasoline prices rise rapidly like rockets when crude oil prices increase but fall slowly like feathers when crude oil prices decrease. Bacon found that retailers quickly pass on cost increases to maintain profit margins, while price reductions are delayed due to market power, inventory management, and adjustment costs. According to Bremmer and Kesselring 2016, the opposite of rockets-and-feathers effect – balloons-and-rocks effect – can be observed during periods of generally falling prices, as in case with gasoline, when prices decrease more rapidly when crude oil prices drop.

In fertilizer production, costs are heavily influenced by oil and natural gas prices, which are critical inputs, especially for nitrogen-based fertilizers. The energy-intensive nature of fertilizer production means that natural gas accounts for 70-90% of production costs. Additionally, oil prices impact transportation costs due to the heavy and costly nature of transporting fertilizers (Bekkerman et al. 2020).

The objective of this study is to investigate the nonlinear dynamics between the spot prices of various fertilizers and crude oil using a Nonlinear Autoregressive Distributed Lag (NARDL) model. By employing the Bai-Perron method to identify structural changes, this research aims to uncover both symmetric and asymmetric responses in fertilizer prices over different time periods, thereby providing a comprehensive understanding of the intricate relationship between crude oil and global fertilizer prices.

2 Methodology

2.1 NARDL modeling approach

The NARDL (Nonlinear Autoregressive Distributed Lag) model was developed by Shin et al. 2014 to extend the well-known ARDL (Autoregressive Distributed Lag) cointegrating procedure by Pesaran et al. 2001 to consider cointegration between decomposed parts of all or some explanatory variables and the dependent variable. This extension addresses the limitations of traditional linear models by incorporating the possibility of asymmetric relationships between variables. As described here, the general NARDL modelling approach can be easily generalized to multiple explanatory variables.

First, the independent variable x_t is decomposed into positive (x_t^+) and negative (x_t^-) partial sums of changes in x_t as $x_t^+ = \sum_{j=1}^{t} \max(\Delta x_j, 0)$ and $x_t^- = \sum_{j=1}^{t} \min(\Delta x_j, 0)$.

[1] University of Economics in Prague, Department of Econometrics, Winston Churchill Square 4, 13067 Prague, Czech Republic, kuro01@vse.cz
[2] University of Economics in Prague, Department of Econometrics, Winston Churchill Square 4, 13067 Prague, Czech Republic, lukas.fryd@vse.cz

Then the following ARDL(p,q), $p,q \in \mathbb{Z}_{\geq 0}$ model is considered: $y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=0}^{q} (\psi_j^+ x_{t-j}^+ + \psi_j^- x_{t-j}^-) + e_t$, where $e_t \sim$ i.i.d.$(0, \sigma^2)$, which can always be transformed into an error correction representation to capture both short-run dynamics and long-run equilibrium relationships: $\Delta y_t = \rho \xi_{t-1} + \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + \sum_{j=0}^{q-1} \left(\pi_j^+ \Delta x_{t-j}^+ + \pi_j^- \Delta x_{t-j}^- \right) + \varepsilon_t$, where $\xi_t = y_t - \beta^+ x_t^+ - \beta^- x_t^-$ is the error correction term, which is usually expanded further and does not enter the final model explicitly:

$$\Delta y_t = \rho y_{t-1} + \theta^+ x_{t-1}^+ + \theta^- x_{t-1}^- + \sum_{i=1}^{p-1} \delta_i \Delta y_{t-i} + \sum_{j=0}^{q-1} (\pi_j^+ \Delta x_{t-j}^+ + \pi_j^- \Delta x_{t-j}^-) + \varepsilon_t \quad (1)$$

where the long-run coefficients β^+ and β^- can be identified as $-\theta^+/\rho$ and $-\theta^-/\rho$, respectively. The new coefficients in 1 are simply a particular linear combination of the original coefficients of the model described by the ARDL representation. Model of type 1 is then easily estimated with OLS by applying appropriate variable decomposition.

The bounds testing approach, developed by Pesaran et al. 2001, is then used to test for a cointegrating relationship between a dependent variable and regressors. It involves fitting an ARDL model and using a Wald or F-statistic to see if the statistic falls outside critical value bounds, indicating the integration/cointegration status of the regressors. The test uses two sets of asymptotic critical values for regressors that are purely $I(1)$ or purely $I(0)$, supplemented by a t-test from Banerjee et al. 1998. To test for cointegration, an F-test of the joint null $\rho = \theta^+ = \theta^- = 0$ (Bounds test, denoted F_{PSS}) or $\rho = 0$ (BDM test, denoted t_{BDM}) are conducted. These test statistics have non-standard asymptotic distributions, and a cointegrating relationship is found only if both nulls are rejected (see Pesaran et al. 2001.

According to Shin et al. 2014, the short-run dynamic parameters (=differenced regressors) follow an asymptotically normal distribution, allowing inference using the Wald test, which is also used to test for short-run and long-run asymmetry. Short-run asymmetry is tested as $\sum_{i=0}^{q-1} \pi_i^+ = \sum_{i=0}^{q-1} \pi_i^-$, while long-run asymmetry is tested as $\theta^+ = \theta^-$. Rejection of these nulls implies asymmetry, meaning that data does not suggest the presence of rockets-and-feathers or balloons-and-rocks effects either in the short or long run. A typical regression model with long-run asymmetry only is written as $\Delta y_t = \rho y_{t-1} + \theta^+ x_{t-1}^+ + \theta^- x_{t-1}^- + \sum_{i=1}^{p-1} \delta_i \Delta y_{t-i} + \sum_{j=0}^{q-1} \gamma_j \Delta x_{t-j} + \varepsilon_t$, while one with long-run symmetry and short-run asymmetry as $\Delta y_t = \rho y_{t-1} + \theta x_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta y_{t-i} + \sum_{j=0}^{q-1} (\gamma_j^+ \Delta x_{t-j}^+ + \gamma_j^- \Delta x_{t-j}^-) + \varepsilon_t$.

Shin et al. 2014 also provide a way to further assess and visualize the adjustment path to the new equilibrium, given a permanent shock to the regressors by + or - 1 unit (percentage) change — the asymmetric dynamic multipliers, which are derived from the ARDL in-levels representation of the NARDL model. As time passes, $h \to \infty$, these multipliers converge to the asymmetric long-run coefficients β^+ and β^-. Usually, these two multipliers and their difference, for which the confidence intervals are bootstrapped, are plotted together.

2.2 Bai-Perron breakpoints estimation approach

Bai and Perron 2003 provide a framework for identifying and estimating multiple structural breaks in a time series regression model without prior knowledge of the breaks. This methodology assumes a linear model with stationary or trend-stationary variables.

Bai and Perron 2003 consider the following model for multiple structural changes:

$$y_t = \boldsymbol{x}_t' \boldsymbol{\beta} + z_t' \boldsymbol{\delta}_j + u_t, \quad t = T_{j-1} + 1, \ldots, T_j \quad (2)$$

for $j = 1, \ldots, m+1$, where $T_0 = 0$, $T_{m+1} = T$, and T is the sample size. Here, y_t is the dependent variable, $\boldsymbol{x}_t$ is a vector of covariates with coefficients $\boldsymbol{\beta}$, z_t is a vector of covariates with coefficients $\boldsymbol{\delta}_j$ which may change at the breakpoints T_j, and u_t is the error term.

1. No break vs $1 \leq s \leq s^*$ breaks (UDmax test):

$$H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_2 = \ldots = \boldsymbol{\delta}_{s+1} \quad \text{vs} \quad H_1 : \boldsymbol{\delta}_k \neq \boldsymbol{\delta}_j \quad \text{for some} \quad j \neq k \quad \text{and} \quad s = 1, \ldots, s^* \quad (3)$$

   The UDmax test is a double maximum test designed to test the null hypothesis of no structural breaks against the alternative of an unknown number of breaks.

2. s breaks vs $s+1$ breaks (sequential supF test):

$$H_0 : \boldsymbol{\delta}_j = \boldsymbol{\delta}_{j+1} \quad \text{for one} \quad j = 1, \ldots, s \quad \text{vs} \quad H_1 : \boldsymbol{\delta}_j \neq \boldsymbol{\delta}_{j+1} \quad \text{for all} \quad j = 1, \ldots, s \quad (4)$$

   If the null of the UDmax test is rejected, break dates need to be estimated by testing each s segment for an additional break. Instead of comparing the SSR for models with one and two breaks, each segment is tested for another break. The test continues until no significant additional break is found.

For both tests, the test statistic is defined in Bai and Perron 2003, and asymptotic critical values depend on the number of breaks s and regressors q.

3 Results

For our analysis, we use the Weekly Green Markets North American Fertilizer Price Index from Green Markets, a subsidiary of Bloomberg. This index is constructed from the weighted average prices of US Gulf Coast Urea, US Cornbelt Potash, and NOLA Barge DAP, with weights based on the annual global demand for each nutrient. It serves as an approximation of overall fertilizer price changes in the North American fertilizer market. We include weekly oil and gas prices as explanatory variables, specifically the Cushing, OK WTI Spot Price and the Henry Hub natural gas spot price, following Crespi et al. 2022 and Bekkerman et al. 2020. The negative WTI spot price on April 20, 2020, along with several other missing entries, were treated as NAs and imputed using the Kalman filter. For the analysis, these variables are log-transformed to $lwti$, $lgas$, and $lindex$, while differenced variables are prefixed by Δ.

We then apply the Augmented Dickey-Fuller (ADF) unit root tests with drift and lags ranging from 0 to 5 to determine the order of integration for each variable, using both the full dataset and two subsets, namely before and after the breakpoint of June 25, 2021. In all cases, the null hypothesis of a unit root was not rejected for the logged variables, but it was rejected after taking the first difference of these variables, indicating that $lwti$, $lgas$, and $lindex$ are all I(1) for all time periods.

We estimate the following NARDL model (Equation 5) according to Shin et al. 2014 using data from August 11, 2008, to April 5, 2024, with $p, q = 12$. Then, we apply a general-to-specific approach, starting with a complete model and progressively removing insignificant short-run parameters at the $\alpha = 0.05$ level of significance. We then perform F_{pss} and t_{bdm} tests to check for the presence of a cointegrating relationship.

$$\begin{aligned} lindex_t &= \alpha + \theta_1^+ lwti_{t-1}^+ + \theta_1^- lwti_{t-1}^- + \theta_2^+ lgas_{t-1}^+ + \theta_2^- lgas_{t-1}^- \\ &+ \sum_{i=1}^{p} \delta_i lindex_{t-i} + \sum_{i=0}^{q} \left(\phi_i^+ \Delta lwti_{t-i}^+ + \phi_i^- \Delta lwti_{t-i}^-\right) + \sum_{i=0}^{q} \left(\gamma_i^+ \Delta lgas_{t-i}^+ + \gamma_i^- \Delta lgas_{t-i}^-\right) + \varepsilon_t \end{aligned} \tag{5}$$

Our analysis spans periods of instability, including the 2008 financial crisis, the 2011 food crisis, significant fertiliser industry mergers, and exogenous shocks like the 2021 Texas storm or The Great Texas Freeze (see Crespi et al. 2022). These factors could lead to structural breaks that may impact the consistency of OLS estimates. To tackle this problem, we estimate a nested model of model II from 1, excluding decomposed and endogenous variables such as $lindex_{t-1}$, $lwti_{t-1}^+$, etc. in order to identify structural breaks through the Bai-Perron procedure according to Bai and Perron 2003.

The UDMax test statistic is 28.207, exceeding the 5% critical value of 27.230 (Table 2), indicating that at least one structural breakpoint is present. Sequential supF tests are as follows: supF(1|0) = 23.576, which is less than the 10% critical value of 24.650, but supF(2|1) test statistic equals 30.025, exceeding the 5% critical value of 29.240, which, in turn, suggests that an additional breakpoint might be needed. Further supF tests do not reject the null hypothesis. To resolve these somewhat contradictory findings, we estimate a single breakpoint identified as June 25, 2021, with a 95%

Analysing the entire dataset from 8/11/2008 to 4/5/2024 (Table 1, model I), we observe a cointegrating relationship, as the T_{BDM} and F_{PSS} statistics for the Bounds and BDM tests are rejected. Model II, which considers long-run asymmetry only for natural gas, also indicates a cointegrating relationship (Table 1). Both models reject the null hypotheses of no rockets and feathers effect for crude oil in the short and long run. Precisely, a 1% increase in oil prices is estimated to cause approximately a 0.849 and 0.814 increase in the value of the index for models I and II, respectively. Conversely, a 1% decrease in oil prices causes fertiliser prices to fall on average by -0.516 and -0.493 for models I and II, respectively. Figures 1a and 1c show that both positive and negative effects of oil price changes settle around the 40$^{\text{th}}$ week, and the difference between positive and negative dynamic multipliers becomes statistically significant by the 20$^{\text{th}}$ week. In the short run, however, the effect of balloons and rocks is apparent for the first few weeks. Furthermore, for natural gas, the opposite rocks and balloons effect is present as well, regardless of the imposition of short-run symmetry for natural gas or not (Figures 1b, 1d).

Next, we estimate the NARDL model for two separate periods: 8/11/2008 to 6/18/2021 and 6/5/2021 to 4/5/2024, using a general-to-specific approach by Shin et al. 2014, since re-estimating models I and II did not yield a cointegrating relationship for any of these two new time periods. For both model III and IV from Table 1, the nulls of the bounds tests and its supplementary BDM tests were rejected. Model III suggests that the presence of a rockets and feathers effect in both the short and long run, more pronounced as $|L^{+,III}_{lwti}| > |L^{+,I}_{lwti}|, \quad |L^{+,III}_{lwti}| > |L^{+,II}_{lwti}|$. Dynamic multipliers for oil are similar across models I and III (see Figure 1e). For natural gas, the balloons and rocks effect is less pronounced and not statistically significant (long-run asymmetry p-val. = 0.104; Figure 1f) compared to models I and II.

Model IV again finds a cointegrating relationship while revealing the balloons and rocks effect. Here, both long-term coefficients for oil are negative: -0.256 for a 1% increase and -0.843 for a 1% decrease. Moreover, the dynamic multipliers for oil and gas, both positive and negative, converge more quickly to their respective long-run parameters (Figures 1g, 1h). The adjustment takes place much faster and finishes at approximately 20th or 25th week compared to previous models, which were estimated either partially or fully using data for the time period from 8/11/2008 to 6/18/2021.

The reasons for the rockets and feathers effect in both the entire dataset and the pre-2021 data could include significant consolidation and market power issues in the North American fertiliser market, aligning with Bacon 1991. Bekkerman et al. 2020 highlights this consolidation in the U.S. nitrogen fertiliser production sector, noting a reduction in production facilities from 59 to 22 and a decline in firms from 46 to 13, with the four largest firms accounting for approximately 75% of total domestic production by 2019, indicating high market concentration. Bekkerman et al. 2020 expected this consolidation to lead to higher fertiliser prices, especially nitrogen-based ones, due to fewer firms controlling the market. This market power concentration is further corroborated by Crespi et al. 2022, who states that by 2018, four large firms—CF Industries, Nutrien, DynoNobel, and Koch Industries—controlled 75.3% of domestic production capacity. Additionally, mergers and acquisitions, such as CF Industries acquiring Terra and the merger of Potash Corp. and Agrium to form Nutrien, have further intensified market concentration (see Crespi et al. 2022).

The reasons for the balloons and rocks effect might be explained by the surge in synthetic fertiliser prices since 2021 due to energy costs and higher demand, which made reducing fertiliser use more economically attractive for farmers (see Schaub and Benni 2024). Schaub et al. provide an example of strategies such as substituting synthetic fertilisers with legumes in grassland-based milk production and employing precision farming technologies in wheat production to help maintain yields and cut costs.

4 Conclusion

Our NARDL model analysis shows significant dynamics between fertilizer prices and crude oil and natural gas prices. Models I and II reveal a rockets and feathers effect for crude oil from 2008 to 2024, where a 1% increase in oil prices leads to a larger rise in fertilizer prices than the fall caused by a 1% decrease. While Model III confirms this effect, Model IV, estimated after the June 25, 2021, break date, reveals the opposite balloons-and-rocks effect. The rockets and feathers effects might be explained by market consolidation, where a few firms control 75% of production, which might, in turn, lead to higher fertilizer prices. Conversely, the balloons and rocks effect since 2021 could be due to surging fertilizer prices from energy costs, making reduced fertilizer use more attractive for farmers.

Acknowledgements

The work was supported by the Internal Grant Agency of Prague University of Economics and Business under Grant F4/24/2023.

References

Bacon, R. W. (1991). Rockets and feathers: the asymmetric speed of adjustment of uk retail gasoline prices to cost changes. *Energy Economics*, 13(3):211–218.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. 18:1–22.

Banerjee, A., Dolado, J., and Mestre, R. (1998). Error-correction mechanism tests for cointegration in a single-equation framework. *Journal of Time Series Analysis*, 19(3):267–283.

Bekkerman, A., Brester, G. W., and Ripplinger, D. (2020). The history, consolidation, and future of the u.s. nitrogen fertilizer production industry. *Choices*.

Bremmer, D. S. and Kesselring, R. G. (2016). The relationship between us retail gasoline and crude oil prices during the great recession:"rockets and feathers" or "balloons and rocks" behavior? *Energy Economics*, 55:200–210.

Crespi, J., Hart, C., Wongpiyabovorn, O., Schulz, L., and Zhang, W. (2022). An examination of recent fertilizer price changes.

Pesaran, M. H., Shin, Y., and Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16(3):289–326.

Schaub, S. and Benni, N. E. (2024). How do price (risk) changes influence farmers' preferences to reduce fertilizer application? *Agricultural Economics*, 55(2):365–383.

Shin, Y., Yu, B., and Greenwood-Nimmo, M. (2014). Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ardl framework. *Festschrift in Honor of Peter Schmidt*, pages 281–314.

Table 1: NARDL models estimation and diagnostics results

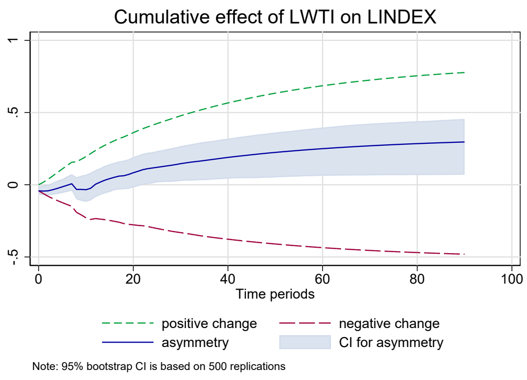
I			II			III			IV		
8/11/2008–4/5/2024			8/11/2008–4/5/2024			8/11/2008–6/18/2021			6/25/2021–4/5/2024		
Var.	Coeff.	S.E.	Var.	Coeff.	S.E.	Var.	Coeff.	S.E.	Var.	Coeff.	S.E.
$lindex_{t-1}$	−0.02 ***	0.005	$lindex_{t-1}$	−0.022***	0.005	$lindex_{t-1}$	−0.133**	0.055	$lindex_{t-1}$	−0.127***	0.033
$lwti^+_{t-1}$	0.173***	0.005	$lwti^+_{t-1}$	0.018***	0.005	$lwti^+_{t-1}$	0.014***	0.005	$lwti^+_{t-1}$	−0.033	0.041
$lwti^-_{t-1}$	0.01 **	0.004	$lwti^-_{t-1}$	0.011**	0.004	$lwti^-_{t-1}$	0.008*	0.004	$lwti^-_{t-1}$	0.107**	0.532
$lgas^+_{t-1}$	−0.001	0.004	$lgas^+_{t-1}$	0.0005	0.04	$lgas^+_{t-1}$	0.004	0.005	$lgas^+_{t-1}$	0.070***	0.026
$lgas^-_{t-1}$	0.002	0.004	$lgas^-_{t-1}$	0.005	0.004	$lgas^-_{t-1}$	0.006	0.005	$lgas^-_{t-1}$	0.240*	0.013
$\Delta lindex_{t-1}$	0.22 ***	0.03	$\Delta lindex_{t-1}$	0.226***	0.035	$\Delta lindex_{t-1}$	0.249***	0.038	$\Delta lindex_{t-1}$	0.179**	0.80
$\Delta lindex_{t-2}$	0.13 ***	0.035	$\Delta lindex_{t-2}$	0.135***	0.035	$\Delta lindex_{t-2}$	0.121***	0.038	$\Delta lindex_{t-1}$	−0.160**	0.78
$\Delta lindex_{t-10}$	0.11 ***	0.034	$\Delta lindex_{t-10}$	0.105***	0.034	$\Delta lindex_{t-10}$	0.100***	0.038	$\Delta lgas$	0.027	0.286
$\Delta lindex_{t-11}$	−0.08 **	0.035	$\Delta lindex_{t-11}$	−0.076**	0.035	$\Delta lindex_{t-11}$	−0.063	0.038	$\Delta lwti_{t-1}$	0.185**	0.085
$\Delta lindex_{t-12}$	−0.11 ***	0.034	$\Delta lindex_{t-12}$	−0.115***	0.035	$\Delta lindex_{t-12}$	−0.147***	0.037	$\Delta lindex_{t-6}$	−0.246***	0.081
$\Delta lwti^+_{t-8}$	−0.019	0.017	$\Delta lwti^+_{t-8}$	0.022	0.017	$\Delta lwti^+_{t-4}$	0.008	0.015	$\Delta lindex_{t-12}$	−0.178**	0.793
$\Delta lwti^-_{t}$	−0.042***	0.15	$\Delta lwti^-_{t}$	0.044***	0.15	$\Delta lwti^+_{t-7}$	0.053	0.029	Cons.	0.875***	0.230
$\Delta lwti^-_{t-8}$	0.30 *	0.15	$\Delta lwti^-_{t-8}$	0.028*	0.015	$\Delta lwti^+_{t-8}$	−0.030	0.017			
$\Delta lwti^-_{t-12}$	0.013	0.015	$\Delta lwti^-_{t-12}$	−0.013	0.015	$\Delta lwti^-_{t}$	−0.042***	0.013			
$\Delta lgas^+_{t}$	−0.022	0.16	$\Delta lgas_{t-3}$	0.021**	0.009	$\Delta lwti_{t-8}$	0.062	0.024			
$\Delta lgas^+_{t-4}$	0.043**	0.018	Cons.	0.136***	0.034	$\Delta lgas^+_{t-3}$	0.025***	0.001			
$\Delta lgas^+_{t-6}$	0.027	0.018				Cons.	0.084	0.034			
$\Delta lgas^+_{t-10}$	0.028*	0.015									
$\Delta lgas^-_{t-3}$	0.044***	0.016									
$\Delta lgas^-_{t-5}$	0.036**	0.016									
$\Delta lgas^-_{t-8}$	0.032**	0.014									
Cons.	0.127***	0.034									
Stat.	Value	p-val.	Stat.	Value	p-val.	Stat.	Value	p-val.	Stat.	Value	p-val.
t_{BDM}	−3.875	>0.10	t_{BDM}	−4.117	>0.05	t_{BDM}	−4.223	<0.05	t_{BDM}	−3.810	<0.10
F_{PSS}	3.997	>0.10	F_{PSS}	4.35	>0.05	F_{PSS}	4.147	<0.05	F_{PSS}	3.521	<0.10
L_{lwti^+}	0.849	>0.01	L_{lwti^+}	0.814	>0.01	L_{lwti^+}	1.013	0.006	L_{lwti^+}	−0.256	0.453
L_{lwti^-}	−0.516	0.003	L_{lwti^-}	−0.493	>0.01	L_{lwti^-}	−0.623	0.024	L_{lwti^-}	−0.843	0.018
L_{lgas^+}	0.070	0.729	L_{lgas^+}	0.021	0.908	L_{lgas^+}	0.267	0.473	L_{lgas^+}	0.545	0.008
L_{lgas^-}	−0.122	0.492	L_{lgas^-}	−0.208	0.204	L_{lgas^-}	−0.461	0.244	L_{lgas^-}	−0.187	0.068
$W_{LR,lwti}$	7.822	0.005	$W_{LR,lwti}$	8.550	0.004	$W_{LR,lwti}$	3.735	0.054	$W_{LR,lwti}$	8.447	0.004
$W_{SR,lwti}$	6.433	0.011	$W_{SR,lwti}$	6.867	0.009	$W_{SR,lwti}$	4.870	0.028	$W_{SR,lwti}$		
$W_{LR,lgas}$	5.971	0.015	$W_{LR,lgas}$	6.51	0.011	$W_{LR,lgas}$	2.647	0.104	$W_{LR,lgas}$	4.704	0.032
$W_{SR,lgas}$	1.184	0.277	$W_{SR,lgas}$			$W_{SR,lgas}$			$W_{SR,lgas}$		
$\bar{R}^2$	0.163		$\bar{R}^2$	0.151		$\bar{R}^2$	0.195		$\bar{R}^2$	0.241	
χ^2_{SC}	9.179	0.164	χ^2_{SC}	7.9795	0.240	χ^2_{SC}	5.952	0.429	χ^2_{SC}	6.731	0.346
χ^2_{FF}	0.115	0.734	χ^2_{FF}	1.10	0.295	χ^2_{FF}	0.436	0.510	χ^2_{FF}	0.494	0.483
χ^2_{HET}	69.206	<0.01	χ^2_{HET}	71.641	<0.01	χ^2_{HET}	59.097	<0.01	χ^2_{HET}	9.02	0.620
χ^2_{NOR}	593.96	<0.01	χ^2_{NOR}	600.43	<0.01	χ^2_{NOR}	871.37	<0.01	χ^2_{NOR}	0.335	0.846

Notes: The sample period is indicated in the second row. t_{BDM} is the t-statistic for the Banerjee-Dolado-Mestre test, and F_{PSS} is the F-statistic for the Pesaran-Shin-Smith bounds test. L_{x^+} and L_{x^-} denote the long-run coefficients associated with positive and negative changes of output, respectively The Wald test for long-run symmetry is denoted as W_{LR} (H$_0$: $L_{x^+} = L_{x^-}$); Wald test for the additive short-run symmetry condition is denoted W_{SR} (H$_0$: $\sum_{i=0}^{q-1} \pi_i^+ = \sum_{i=0}^{q-1} \pi_i^-$). $\bar{R}^2$ is the adjusted R^2. χ^2_{SC} is the Breusch-Godfrey Lagrange Multiplier test for serial correlation with maximum order of 6, χ^2_{FF} is the Ramsey RESET test for functional form, χ^2_{HET} is the Breusch–Pagan test for heteroscedasticity, χ^2_{NOR} is the Jarque-Bera test for normality. In case of models I, II, and III Newey-West standard errors were used.

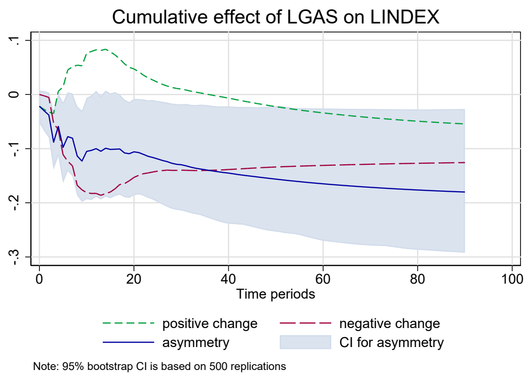
Table 2: supF and UDmax tests results

	supF(1|0)	supF(2|1)	supF(3|2)	supF(4|3)	supF(5|4)	UDmax
Test stat.	23.576	30.025	11.332	16.200	0.000	28.207
10% CV	24.650	26.920	28.260	29.180	29.880	24.900
5% CV	27.030	29.240	30.450	31.450	32.120	27.230

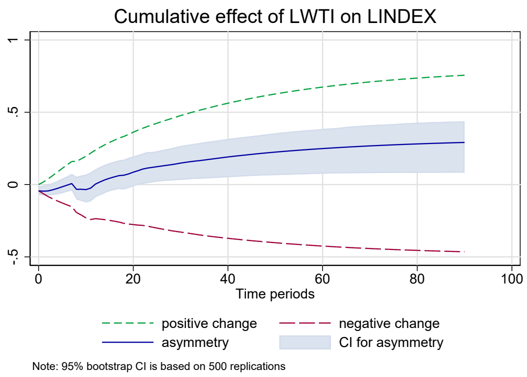
Note: CV stands for critical value.



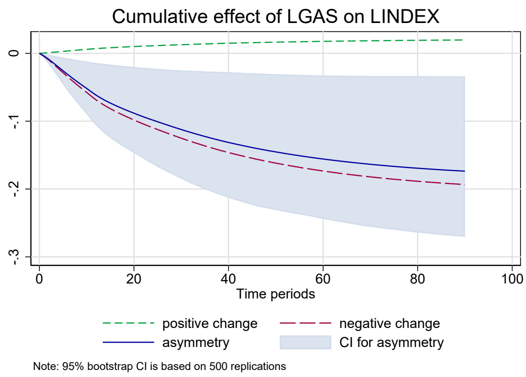
(a) model I



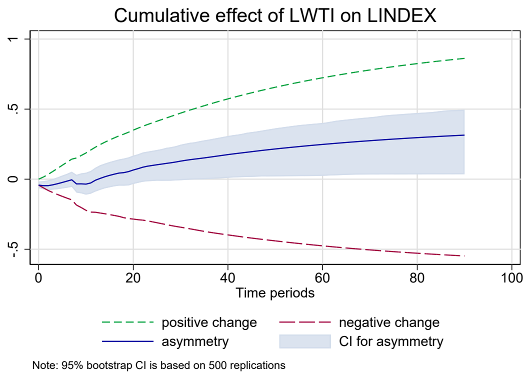
(b) model I



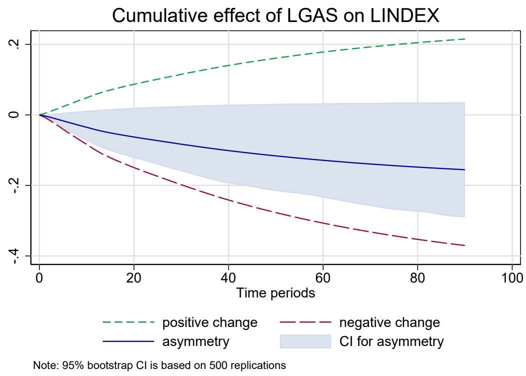
(c) model II



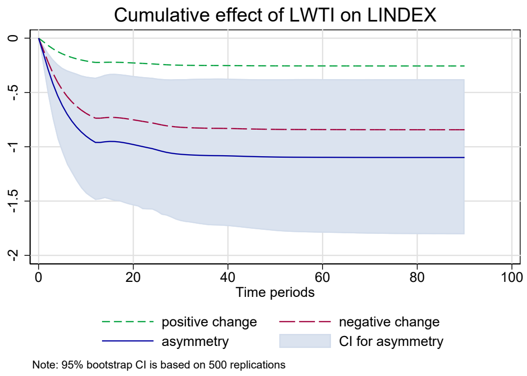
(d) model II



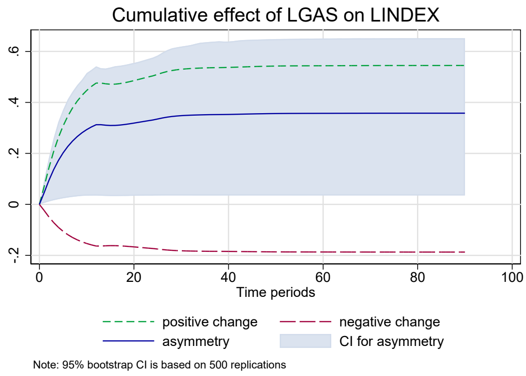
(e) model III



(f) model III



(g) model IV



(h) model IV

Figure 1: Multiplier for models from 1

Swap Heuristic Parameter Sensitivity

Marek Kvet[1], Jaroslav Janáček[2]

Abstract. Swap operation used in neighborhood search heuristics can be modified in many ways introducing various parameters, which play an important role in the search termination rule, move accessibility definition and heuristic strategy determination. These parameters may considerably influence efficiency of the optimization process taking into account the objective function improvement and associated computational time. In this paper, we deal with the general problem of finding a good approximation of Pareto front, which consists of non-dominated *p*-location problem solutions. The applied gradual refinement process consists of enormous number of runs of the neighborhood search heuristic influenced by the parameter settings. In connection with the search for the correct setting of parameter values, we conducted research on swap heuristic behavior with the aim of revealing which parameters belong to the class of sensitive parameters.

Keywords: Bi-criteria decision-making problems, parameter settings, swap heuristics

JEL Classification: C44, C61
AMS Classification: 90C05, 90C06, 90C10, 90C27

1 Introduction

When a big series of *p*-location problems is solved, a simple neighborhood search algorithm based on swap operation proved to be an efficient tool of optimization. The specific task, which asks for repeating application of the neighborhood search, is the construction of a good approximation of the Pareto front of public service system designs by gradual refinement approach.

A Pareto front is a common tool for decision-makers, which enables qualified negotiation with public representatives to balance different points of view of various groups of people involved in the system designing process. It is a typical output used in such optimization problems, in which at least two conflicting objective functions need to be considered. It is a small collection of problem solutions that meet a specific property: none of its elements can be improved in any criterion without worsening the other one [6, 7]. Obtaining the Pareto front by exact methods is a complex and time-consuming challenge. Therefore, several heuristic approaches have been recently developed to achieve good approximation of the original Pareto front [8, 9, 10, 13].

One of possible strategies applicable in creation of Pareto front approximation is the gradual refinement, which is based on processing some sequence of solutions with the goal to find a new solution, which could extend current set of elements. When an individual solution is being processed, its neighborhood is explored. The neighborhood of any solution can be defined in many ways. Nevertheless, a common way is to exchange one location, in which a service center is located for a place, that does not contain a center. Simply said, the neighborhood search uses a swap operation that is performed several times during the whole algorithm run.

As we all know, many heuristics are sensitive to various parameters, which could affect not only their performance, but also the quality of obtained results. Within this paper, we concentrate our attention on study of relation between setting of swap heuristic parameters and the quality of Pareto front approximation.

The studied problem of service system designing with two objectives belongs to the family of location problems, which is widely covered by many authors who have developed exact and approximate modelling and solving approaches [1, 2, 3, 4, 5].

The remaining part of the paper is organized as follows. After the Introduction, the second section provides the readers with explanation of parametrized swap algorithm for *p*-location problem. The third section brings concrete forms of objective functions and discusses the Pareto front approximation quality measurement. The fourth section is devoted to performed computational study and finally, the concluding part of the paper summarizes the obtained results and suggests possible goals for future research activities.

[1] University of Žilina, Faculty of Management Science and Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovakia, marek.kvet@fri.uniza.sk
[2] University of Žilina, Faculty of Management Science and Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovakia, jaroslav.janacek@fri.uniza.sk

2 Parametrized Swap Algorithm for p-Location Problem

The p-location problem is formulated as a task of selection of p locations from m potential ones to minimize an objective function. If a solution of the problem is described by an m-dimensional vector $\boldsymbol{y}$ of zeros and units, the problem can be formally expressed by (1).

$$\min\left\{ f\left(\mathbf{y}\right) : \mathbf{y} \in \{0, 1\}^{m}, \sum_{i=1}^{m} y_i = p \right\} \tag{1}$$

Swap operation performed with a solution $\boldsymbol{y}$ consists in selection of two components y_i and y_j, where $y_i = 1$ and $y_j=0$, and adjusting the solution $\boldsymbol{y}$ so that $y_i = 0$ and $y_j =1$. The resulting vector $\boldsymbol{y}$' is further denoted by *swap*($\boldsymbol{y}$, i, j). The set of all solutions, which can be obtained from $\boldsymbol{y}$ by exactly one swap operation is called the neighborhood of $\boldsymbol{y}$ and is denoted by $N(\boldsymbol{y})$.

The standard swap algorithm starts with an input solution $\boldsymbol{y}^c$ and inspects the solutions from the neighborhood $N(\boldsymbol{y})$. If a better solution $\boldsymbol{y}^b$ is found, then substitution $\boldsymbol{y}^c = \boldsymbol{y}^b$ is performed and the neighborhood search is repeated with new $\boldsymbol{y}^c$. If no better solution is found, the swap algorithm terminates.

The determination of $\boldsymbol{y}^b$ can be performed according several strategies. The most known are the first admissible and best admissible strategies. In this paper, the generalized strategy depending on parameters *maxNos* and *thr* is studied. In addition, a parameter *aTime* is introduced as a limit of time of swap algorithm performance. The studied algorithm is described by the following steps.

SwapAlgorithm($\boldsymbol{y}^c$, *aTime*, *maxNos, thr*)

0. Set *OK* = *true.*
1. If $CPU \leq aTime$ and *OK*, then set $Nos = 0, f^* = f(\boldsymbol{y}^c)$, $i^* = 0$, $P = \{i \in \{1, \ldots, m\}, y^c_i = 1\}$, $Q = \{j \in \{1, \ldots, m\}, y^c_j = 0\}$ and continue with step 2. Otherwise return $\boldsymbol{y}^c$ and terminate.
2. While $Nos < maxNos$ and *OK* perform the following commands for each pair (i, j), $i \in P$, $j \in Q$.
   If $f(\boldsymbol{y}^c) - f(\text{swap}(\boldsymbol{y}^c, i, j)) \geq thr$, then do $Nos = Nos+1$ and if $f(\text{swap}(\boldsymbol{y}^c, i, j)) < f^*$, then set $f^* = f(swap(\boldsymbol{y}^c, i, j))$, $i^* = i, j^* = j$.
3. If $i^* > 0$, then substitute $\boldsymbol{y}^c = swap(\boldsymbol{y}^c, i^*, j^*)$, else set *OK* = *false.*
   Continue with step 1.

3 Application of Swap Algorithm to Pareto Front Approximation

As it was mentioned in the Introduction, a Pareto front is formed by a list of solutions, which are evaluated by two different objective functions and cannot be improved in both of them at the same time. Let us concentrate on the objectives themselves, now.

Let I represent a finite set of candidates from which exactly p items are to be selected. For every subset P that is a subset of I and has a cardinality equal to p, it indicates a possible solution to the problem. This study focuses on the analysis of system and fair criteria. The system objective function can be defined based on equation (2), where J represents a finite collection of users who have made b_j service requests. Moreover, let t_{ij} correspond to the traversing time from a service provider located at $i \in$I to the client's location j and finally, let q_k denote the probability value of the case that the k-th nearest service center will be the first available one. The operator $\min_k\{V\}$ gives the k-th smallest value in V. More details can be found in [12].

$$f_1\left(P\right) = \sum_{j \in J} b_j \sum_{k=1}^{r} q_k \min_k \left\{ t_{ij} : i \in P \right\} \tag{2}$$

The second criterion f_2 considers the aspect of fairness. It minimizes the total number of those users' calls, for which the response time exceeds given limit T. This objective $f_2(P)$ can be formally written by the formula (3).

$$f_2\left(P\right) = \sum_{j \in J} b_j \max\left\{0, sign\left(\min\left\{t_{ij} : i \in P\right\} - T\right)\right\} \tag{3}$$

As the objective (2) follows the concept of generalized disutility, the parameter r was set to 3. The associated coefficients q_k were set in percentage in the following way: q_1 = 77.063, q_2 = 16.476 and q_3 = 100 - q_1 - q_2. Parameter T used in the fair criterion (3) was set to the value of 10 minutes [6, 7, 8, 9, 10, 11, 12].

Since Pareto front completion is a too complex challenge for practical usage, several heuristics have been developed. The gradual refinement used in our research is based on the following principle: The scheme of the used incremental process is based on the idea that it step-by-step improves the current Pareto front approximation by including a newly obtained non-dominated problem solution. A simple step of the process consists of processing one member of the approximation, what means that the member is used as a starting solution of a routine, which performs series of swap operations. If we want to evaluate the quality of the Pareto front approximation, we may apply the concept of the region created by certain *NDSS* set (set of non-dominated solutions) pieces [8, 9, 10, 13]. Undoubtedly, in order to ensure the evaluation is precise and significant, it is important to correctly ascertain the bordering solutions of the collection. The approximating collection of non-dominated solutions (*NDSS*) will be stored by a sequence of *noNDSS* solutions $\mathbf{y}^1, \ldots, \mathbf{y}^{noNDSS}$ ordered according to increasing values of f_2. Here, the symbol *noNDSS* means the number of *NDSS* elements. Under these assumptions, the quality of the approximation *NDSS* can be evaluated using $A(NDSS)$ computed according to the formula (4).

$$A(NDSS) = \sum_{k=1}^{noNDSS-1} \left(f_1\left(\mathbf{y}^k\right) - f_1\left(\mathbf{y}^{noNDSS}\right)\right)\left(f_2\left(\mathbf{y}^{k+1}\right) - f_2\left(\mathbf{y}^k\right)\right) \tag{4}$$

4 Computational Experiments

The goal of the computational study is to investigate relation between parameters *aTime*, *maxNos* and *thr* and quality of the Pareto front approximation evaluated by the resulting gap. The study is carried out on eight benchmarks derived from existing public service systems ensuring emergency medical aid in the individual self-government regions (higher territorial units - HTU) of the Slovak Republic. The exact Pareto fronts were presented in [6, 7] and the associated characteristic are reported in Table 1, where each studied instance corresponds to one row of the table. The first two columns denoted by m and p contain the problem size. The *NoS* stands for number of solutions forming the Pareto front (*PF*) and the *A*(*PF*) denotes associated *Area*. The list of problem instances contains the HTU of Bratislava (BA), Banská Bystrica (BB), Košice (KE), Nitra (NR), Prešov (PO), Trenčín (TN), Trnava (TT) and Žilina (ZA). In the used input data, all inhabited network nodes represent the set of possible candidate locations of service centers and the possible demand locations as well.

Table 1 Benchmarks and the exact Pareto fronts characteristics

HTU	*m*	*p*	*NoS*	*A*(*PF*)
BA	87	14	34	569039
BB	515	36	229	1002681
KE	460	32	262	1295594
NR	350	27	106	736846
PO	664	32	271	956103
TN	276	21	98	829155
TT	249	18	64	814351
ZA	315	29	97	407293

The first phase of the computational study is devoted to the search for such setting of *aTime*, *maxNos* and *thr* , which minimizes average gap of the all eight instances.

We started with determination of the parameter ranges taking into account that their setting near to upper limit can hardly bring a good result. Based on previous experience, the following ranges were specified. Interval [0.01 s, 1.0 s] stands for *aTime* value range, interval [1, 100] stands for *maxNos* range and interval [0, V^{max}] contains inspected values of *thr*. The value of V^{max} was determined proportionally to the area of rectangle determined by bordering solutions $\boldsymbol{x}^0$ and $\boldsymbol{x}^{no}$ of Pareto front associated with the individual instance. The solution $\boldsymbol{x}^0$ is the member of Pareto front with minimal value of the objective function f_2 and $\boldsymbol{x}^{no}$ is the member with maximal value of the objective function f_2. Then, the value V^{max} is determined according to (5).

$$V^{\max} = (f_2(\mathbf{x}^{no}) - f_2(\mathbf{x}^0))(f_1(\mathbf{x}^0) - f_1(\mathbf{x}^{no}))/100 \quad (5)$$

Each range [a_0, a_{max}] was unevenly divided by five values a_0, a_1, …, a_4, according to the formula $a_k = a_{k-1} + A.2^k$ for k= 1, …, 4, where $A \approx (a_{max}- a_0)/15$. The concrete values of the individual parameters are plotted in Table 2.

Table 2 The value V is determined as $V^{max}/15$

	a_0	a_1	a_2	a_3	a_4
aTime	0.01	0.07	0.19	0.43	0.91
maxNos	1	7	19	43	91
thr	0	V	$3V$	$7V$	$15V$

To find a good setting of the parameters, the following heuristic procedure was used. Based on preliminary experiments the starting setting [*aTime*=a_0, *maxNos* = a_1, *thr* = a_2] was chosen. Sum of gaps associated with individual instances for the setting was 13.93.

The search for good parameter setting continues so that the first two parameters stayed fixed and the third one was changed over its range. The best result was obtained for *thr* = a_0, where the associated sum of gaps equaled to 7.02. Then the *thr* was fixed at a_0 and the parameter *aTime* stayed fixed at a_0 and parameter *maxNos* was changed over its range. The best solution was obtained for *maxNos* = a_2 with the value of the sum of gaps equaled to 6.77. In the third run, the parameters *thr* and *maxNos* stayed fixed at their optimized values and *aTime* was tested for each value from its range. The best value of the sum was 6.77 and it was obtained for a_4. Result of this test was verified by fourth run, when range of *aTime* was extended. The setting *aTime* = a_4. *maxNos* = a_2 *thr* = a_0 was submitted further verification by repeating the tests.

Table 3 Results of numerical experiments – table of gaps for fixed *aTime* = a_0, *maxNos* = a_1

thr	a_0	a_1	a_2	a_3	a_4
ZA	1.25	2.02	1.79	1.79	1.79
TT	0	0.65	0.65	0.65	0.65
TN	0.72	4.38	4.38	4.38	4.38
PO	0.13	1.31	1.31	1.31	1.31
NR	0.99	0.99	0.99	0.99	0.99
KE	1.52	1.66	1.66	1.73	1.59
BB	0.23	1.06	1.06	1.06	1.06
BA	2.18	5.55	2.08	2.08	2.08
sum	7.02	17.62	13.92	13.99	13.85

Table 4 Results of numerical experiments – table of gaps for fixed *aTime* = a_0, *thr* = a_0

maxNos	a_0	a_1	a_2	a_3	a_4
ZA	1.81	1.25	1.03	1.03	1.03
TT	0.09	0	0	0	0
TN	0.89	0.72	0.72	0.72	0.72
PO	0.17	0.13	0.13	0.13	0.13
NR	0.57	0.99	0.99	0.99	0.99
KE	2.98	1.52	1.48	1.48	1.48
BB	0.6	0.23	0.23	0.23	0.23
BA	5.88	2.18	2.18	2.18	2.18
sum	12.99	7.02	6.76	6.76	6.76

Table 5 Results of numerical experiments – table of gaps for fixed *maxNos* = a_2, *thr* = a_0

aTime	a_0	a_1	a_2	a_3	a_4
ZA	2.02	2.02	1.03	1.03	1.03
TT	0.65	0	0	0	0
TN	4.38	0.72	0.72	0.72	0.72
PO	4.32	4.32	1.31	0.13	0.13
NR	0.99	0.99	0.99	0.99	0.99
KE	3.56	3.56	1.87	1.48	1.48
BB	1.08	1.08	1.11	0.35	0.23
BA	2.18	2.18	2.18	2.18	2.18
sum	19.18	14.87	9.21	6.88	6.76

The second phase of the computational study deals with parameter sensitivity based on results plotted in the tables above.

As concerns the parameter *thr*, it can be stated that in all cases with tiny deviation of results obtained for benchmark BA, the best gaps were obtained for the same setting of the parameter. Then, we can conclude that this parameter is not sensitive.

The similar characteristic has the parameter *maxNos*, for which such setting exists that it is optimal for each benchmark. Nevertheless, the experiment showed that for bigger values of this parameter, the results are also optimal.

Also, parameter *aTime* proved not to be sensitive to the problem being solved.

5 Conclusions

The swap operation utilized in neighborhood search heuristics can be altered in numerous ways by adding various factors and parameters. These parameters have a significant impact on the search termination rule, move accessibility definition, and heuristic approach determination. These characteristics can have a significant impact on the effectiveness of the optimization process, considering the improvement of the objective function and the computing time connected with it. This study addressed the overarching issue of identifying an effective approximation of the Pareto front, which encompasses solutions to the non-dominated *p*-location problem. The applied iterative refining process involves a large number of iterations of the neighborhood search heuristic, which is controlled by the parameter choices. We performed a study on the behavior of swap heuristics to identify the sensitive parameters that need to be correctly configured. Based on the achieved results we can conclude that we have fulfilled the scientific goals of this paper at a satisfactory level.

Obviously, development of heuristic methods is a never-ending story. Therefore, we suggest several directions for future. Future research in the area of bi-criteria location science could be aimed at development of other approaches for good Pareto front approximation. Another research topic could be aimed at generalization of current approaches and their adjustment for multi-objective system designing.

Acknowledgements

This work was supported by the following grants: VEGA 1/0077/22 “Innovative prediction methods for optimization of public service systems”, and VEGA 1/0654/22 “Cost-effective design of combined charging infrastructure and efficient operation of electric vehicles in public transport in sustainable cities and regions”. This paper was supported by the Slovak Research and Development Agency under the Contract no. APVV-19-0441.

References

[1] Arroyo, J. E. C., dos Santos, P. M., Soares, M. S. and Santos, A. G. (2010). A Multi-Objective Genetic Algorithm with Path Relinking for the p-Median Problem. In: Proceedings of the 12th Ibero-American Conference on Advances in Artificial Intelligence, 2010, pp. 70–79.

[2] Bertsimas, D., Farias, V. F., Trichakis, N. (2011). The Price of Fairness. In Operations Research, 59, 2011, pp. 17-31.

[3] Brotcorne, L, Laporte, G, Semet, F. (2003). Ambulance location and relocation models. Eur. Journal of Oper.Research, 147, pp. 451-463.

[4] Buzna, Ľ., Kohání, M., Janáček, J. (2013). Proportionally Fairer Public Service Systems Design. In: Communications - Scientific Letters of the University of Žilina 15(1), pp. 14-18.
[5] Current, J., Daskin, M. and Schilling, D. (2002). Discrete network location models, Drezner Z. et al. (ed) Facility location: Applications and theory, Springer, pp. 81-118.
[6] Grygar, D., Fabricius, R. (2019). An efficient adjustment of genetic algorithm for Pareto front determination. In: TRANSCOM 2019: conference proceedings, Amsterdam: Elsevier Science, pp. 1335-1342.
[7] Janáček, J., Fabricius, R. (2021). Public service system design with conflicting criteria. In: IEEE Access: practical innovations, open solutions, ISSN 2169-3536, Vol. 9, pp. 130665-130679.
[8] Janáček, J., Kvet, M. (2023). Adaptive Parameter Setting for Public Service System Design. In: Strategic management and its support by information systems 2023, in print
[9] Janáček, J., Kvet, M. (2021). Quality measure of Pareto front approximation. In: Quantitative methods in economics: Multiple criteria decision making 21: Proceedings of the international scientific conference, Bratislava: Letra Edu, pp. 95-103.
[10] Janáček, J., Kvet, M. (2022). Repeated refinement approach to Bi-objective p-location problems. In_: INES 2022: 26th IEEE International Conference on Intelligent Engineering Systems, New York: Institute of Electrical and Electronics Engineers, pp. 41-45.
[11] Jankovič, P. (2016). Calculating Reduction Coefficients for Optimization of Emergency Service System Using Microscopic Simulation Model. In: *17th International Symposium on Computational Intelligence and Informatics*, pp. 163-167.
[12] Kvet, M. (2014). Computational Study of Radial Approach to Public Service System Design with Generalized Utility. In: *Proceedings of International Conference DT 2014*, Žilina, Slovakia, pp. 198-208.
[13] Kvet, M., Janáček, J. (2023). Hyperheuristic as Tuning Tool of Generalized Swap Strategy. In: META 2023: The 9th International Conference on Metaheuristics and Nature Inspired Computing, 1.-4.11. 2023, in print

Comparative Analysis of Bankruptcy Prediction Models in Metallurgical Industry: Logistic Regression versus Artificial Intelligence Techniques

Stanislav Letkovský[1], Sylvia Jenčová[2], Marta Miškufová[3], Petra Vašaničová[4]

Abstract. Bankruptcy prediction becomes part of the financial managers' toolkit, enabling them to address the potential threat of bankruptcy with the aid of a suitable tool. AI is increasingly becoming a popular method in this area as well. Will this technology supplant classic logistic regression in terms of performance? This study aims to compare the prediction accuracy of logistic regression and selected AI methods. The research is conducted on a sample of over 4,600 enterprises from the metallurgical industry in the conditions of the Slovak Republic from 2019 to 2021. This period allows for a comparison between the pre-crisis period and the period of crisis during the COVID-19 pandemic. Currently, no model focuses on this specific industry in the conditions of this country. This study offers a unique tool for identifying bankruptcy in the metallurgical industry of the Slovak Republic, which can be easily adapted to other countries with a similar underdeveloped capital market. A critical aspect of bankruptcy prediction is the selection of reliable predictors. Based on the analyzed literature, 40 financial indicators are empirically investigated. The proposed prediction models contain optimally selected indicators potentially significant in predicting bankruptcy under these conditions. All proposed models achieve high accuracy.

Keywords: bankruptcy prediction, artificial intelligence, logistic regression, Slovak metallurgical industry, financial indicators

JEL Classification: C45, C53, G33
AMS Classification: 62J12, 62M45

1 Introduction

An economic entity relies on various financial resources to carry out business activities. Financial-economic analysis is used for a comprehensive assessment of financial health. According to Vochozka (2020), it is a formalized method that enables an understanding of the company's financial health. An analytical approach to individual activities is advantageous. Evaluative financial indicators belong to the following groups: profitability, liquidity, activity, indebtedness, and market value. Banks and financial institutions require predictive models to evaluate a client's creditworthiness and estimate the risk of insolvency. Thanks to these estimates, they can more accurately and quickly assess creditworthiness and classify businesses according to potential credit risk. Even company managers can apply these forecasts to prevent problems in time and thus increase their rating and competitiveness. Businesses go bankrupt mainly due to the inability to repay their obligations on time. The reason is often high indebtedness and a lack of liquid assets. The accuracy of predictions depends significantly on both the method employed and the conditions (environment) and data quality. Models designed in foreign conditions are not very suitable for use in Slovakia. It is necessary to adapt them and propose new ones for the specific conditions of the given industry in the Slovak Republic. In this context, our research uses a sample of enterprises from the Slovak metallurgical industry from 2019 to 2021. This study aims to compare the prediction accuracy of logistic regression and selected AI methods.

[1] University of Prešov, Faculty of Management and Business, Department of Finance, Accounting and Mathematical Methods, Konštantínova 16, 080 01 Prešov, Slovakia, stanislav.letkovsky@smail.unipo.sk
[2] University of Prešov, Faculty of Management and Business, Department of Finance, Accounting and Mathematical Methods, Konštantínova 16, 080 01 Prešov, Slovakia, sylvia.jencova@unipo.sk
[3] University of Prešov, Faculty of Management and Business, Department of Finance, Accounting and Mathematical Methods, Konštantínova 16, 080 01 Prešov, Slovakia, marta.miskufova@unipo.sk
[4] University of Prešov, Faculty of Management and Business, Department of Finance, Accounting and Mathematical Methods, Konštantínova 16, 080 01 Prešov, Slovakia, petra.vasanicova@unipo.sk

2 Literature Review

Historically, prediction models began with the work of Fitzpatrick (1932), who compared indicators of healthy and bankrupt companies, and linear discriminant analysis using five financial indicators, which is among the most famous (z-score) by Altman (1968). The next stage of development involved probit analysis (Zmijewski 1984) and logistic regression (Ohlson 1980), which are still often used today. With the advancement of technology comes artificial intelligence (AI). Rosenblatt (1962) described the initial use of an artificial neural network, followed by its application for prediction by Odom and Sharda (1990).

Garcia (2022) supports the theory that non-linear machine learning models achieve better results and determine bankruptcy more accurately than traditional statistical methods. Also, Klieštik et al. (2023) used artificial intelligence to make predictions. The existing literature does not provide a conclusive determination regarding whether artificial intelligence-based methods can statistically significantly enhance accuracy.

The selection of appropriate predictors is a crucial aspect of modeling. The literature does not provide an exact or guaranteed way to select the best. Predictors are often chosen based on popularity, statistical selection, or empirical investigation. The ambiguity of the selection was described by Gajdosikova et al. (2022), Vochozka (2020), Du Jardin (2018). Valašková et al. (2018) use multiple regression to select suitable predictors, while Mihalovič (2018) applies evolutionary algorithms. An overview of the most frequently used predictors in existing studies is provided in Table 1**.**

Predictor		**Authors**
AV/S	Added Value to Sales Ratio	Ben Jabeur (2017); Ptak-Chmielewska (2019); Becerra-Vicario et al. (2020); Jenčová et al. (2021)
CL/TA	Current Liabilities to Total Assets	du Jardin (2018); Korol (2020)
CR	Current Ratio	Becerra-Vicario et al. (2020); Tumpach et al. (2020); Korol (2020); Jenčová et al. (2021); Ngo Thanh-Long (2022)
ROA	Return on Assets	Altman (1968); du Jardin (2018); Ptak-Chmielewska (2019); Becerra-Vicario et al. (2020); Tumpach et al. (2020); Korol (2020)
AT	Assets Turnover	Altman (1968); Ptak-Chmielewska (2019); Jenčová et al. (2021); Ngo Thanh-Long (2022); Sigrist and Leuenberger (2023)
NWC/TA	Net Working Capital to Total Assets	Altman (1968); Ben Jabeur (2017); Valašková et al. (2018); Becerra-Vicario et al. (2020); Jenčová et al. (2021); Ngo Thanh-Long (2022); Sigrist and Leuenberger (2023)

Table 1 Frequently used predictors in existing research

3 Data and Methodology

Slovakia ranks among the most industrialized countries in the EU. Our research focused on a sample of enterprises from the metallurgical industry operating in the Slovak Republic between 2019 and 2021. The sample was divided into two periods: the first set comprised 4,601 samples from 2019 (pre-crisis period), and the second set included 4,667 samples from 2021 (crisis period). This study aims to compare the prediction accuracy of logistic regression and selected AI methods. Financial data were sourced from The Register of Financial Statements of the Slovak Republic. Based on the literature, a compilation of 40 financial indicators was created, covering a range of areas. Subsequently, we identified six potentially significant indicators that encompass metrics of liquidity, indebtedness, profitability, and activity. The selection of predictors followed the methodology outlined by Gavurova et al. (2022), informed by previous research and empirical testing on this dataset. Moreover, the chosen predictors were ensured to be uncorrelated, fulfilling a condition for logistic regression. In this study, we use the predictors according to the **Table 1**.

The data underwent rigorous preprocessing steps. After removing inconsistent samples and outliers, the dataset was reduced to 3,205 samples in the first set (with 446 bankruptcies) and 3,246 samples in the second set (with 472 bankruptcies). Each set was randomly split into two parts for training and testing, maintaining a ratio of 70:30.

We conducted outlier detection using the interquartile range method, where values outside the interval calculated as three times the interquartile range were flagged as outliers. Bankruptcy, as defined by Act no. 7/2005 Coll. on Bankruptcy and Restructuring, is determined by the equity-to-debt ratio. Act no. 513/1991 Coll. The Commercial Code sets the minimum ratio for each year. For the analyzed period, it is 0.08. Sample scaling was performed using the standardization method.

In addition to the detailed results presented in the confusion matrix, an AUC (Area Under the Curve) was also calculated to facilitate comparison and interpretation.

3.1 Logistic Regression (LR)

Logistic regression is one of the basic statistical methods frequently employed in predictive analysis, where it demonstrates good performance and relatively high reliability. The resulting variable falls within the interval of 0 to 1.

This method assumes a logistic probability distribution and is utilized to model one-way dependence between variables, particularly when the dependent variable is discrete. Unlike some other methods, logistic regression does not necessitate a normal distribution of the independent variables. The relationship between the probability of occurrence and the explanatory variable is represented by a logistic curve (Jenčová et al., 2021). The primary criterion for its application is the low collinearity of independent variables (Ptak-Chmielewska, 2019).

In the logistic regression notation:

$$logit(\pi) = a\sum_{i=1}^{n} b_i X_i \tag{1}$$

after logit transformations, the non-linear dependence changes to a linear one, where π represents the probability of the event, a is a constant, b_i are regression coefficients, and X_i are predictors. These coefficients are typically estimated using the "maximum likelihood" method.

3.2 Support Vector Machine (SVM)

Support vector machine maps input vectors into a multidimensional feature space, utilizing a linear model to implement nonlinear relationships (López Iturriaga and Sanz, 2015). The task is to find subsets in the feasible space to maximize the set distance, resembling a quadratic optimization solution and aiming to find the optimal solution with a low number of samples. Unlike neural networks, SVMs require only two parameters: an upper bound and a kernel parameter (Shin et al., 2005). SVMs can handle both categorical and continuous variables for classification and regression tasks (Ptak-Chmielewska, 2019). These models classify the data by employing quadratic optimization, searching for sub-spaces ("hyper-planes") in space to maximize the distance from the data points. The capability to optimize even with a small sample size appears advantageous in the context of the Slovak Republic, where the number of bankrupt samples is minimal (less than 2%). This method is thoroughly described by Cortes and Vapnik (1995).

3.3 Decision Tree (DT)

The basis of this method lies in decision rules that are hierarchically structured according to their importance. The more crucial the factor, the higher it is positioned in the decision tree. The procedure segments the samples based on these rules, gradually dividing them into smaller groups. Ultimately, each sample ends up at one of the branches of this tree, referred to as leaves, which cannot be further divided.

As a supervised learning algorithm, the DT model hierarchically traverses sample by sample from the root condition to the target value (bankruptcy/healthy enterprise). DT is a supervised learning algorithm that uses decision making in the form of a tree structure. This predictive model draws a conclusion from the samples at the beginning (the root of the tree), gradually to the result (the leaves of the tree).

As a result, the output is straightforward to interpret, sensitive to missing data, and does not necessitate a normal distribution. Additionally, this method does not require coefficients (Ptak-Chmielewska, 2019).

3.4 Artificial Neural Network (ANN)

An artificial neural network operates on a simplified principle that mimics the nervous structure of the brain. Learning occurs through the accumulation of experiences from examples, allowing it to recognize specific patterns and retain them within its structure. The fundamental unit is the artificial neuron, which receives input signals, combines them, applies an activation function, and produces an output signal. Mathematically, this is often represented as:

$$y = f\left(\sum_{i} x_i w_i\right) \tag{2}$$

where y is the output, x_i are the inputs, w_i are the synaptic weights, and f is the activation function.

In predictive analysis, financial indicators serve as inputs to the network, with the output being a binary value indicating bankruptcy or non-bankruptcy. The number of hidden layers and nodes varies, with no strict rules. Too few nodes may lack the necessary dynamics for diverse inputs, while too many nodes can unnecessarily complicate and slow down the system without improving performance. Generally, a single hidden layer is sufficient for such tasks, with the number of hidden nodes being a characteristic variable of the model. Initially, synaptic weights are chosen randomly and adjusted through the learning process to optimize results. Choosing an appropriate learning method, such as Back-propagation, where errors at the output are gradually propagated back to the input, can improve accuracy. The dataset is typically divided into training and validation sets, with training used exclusively for learning and validation employed to assess the network's ability to generalize.

4 Results and Discussion

When comparing the crisis period to the pre-crisis period, there is no significant increase in bankrupt companies within this industry. The rise in bankruptcy samples between the periods is less than 0.6%. All models (see **Table 2**) achieved nearly 100% accuracy for non-bankrupt samples, primarily due to the imbalance of bankrupt samples. Conversely, the classification accuracy for bankruptcy samples was less than 43%. Non-bankrupt samples constitute a larger portion, skewing the model towards this group. This phenomenon is consistent with the broader literature on the subject, as noted by Tumpach et al. (2020), and Garcia (2022). Authors often attempt to address this issue by balancing samples, such as generating synthetic samples or resampling techniques.

Sample		Predicted											
		LR			SVM			DT			ANN		
2019	observed	0	1	*Correct (%)*	0	1	*Correct (%)*	0	1	*Correct (%)*	0	1	*Correct (%)*
Training	0	1,909	22	*98.86*	1,928	3	*99.84*	1,906	25	*98.71*	1,912	19	*99.02*
	1	226	86	*27.56*	280	32	*10.26*	188	124	*39.74*	225	87	*27.88*
	Overall (%)	*95.19*	*4.81*	*88.94*	*98.44*	*1.56*	*87.38*	*93.36*	*6.64*	*90.50*	*95.27*	*4.73*	*89.12*
Testing	0	819	9	*98.91*	826	2	*99.76*	814	14	*98.31*	818	10	*98.79*
	1	100	34	*25.37*	118	16	*11.94*	89	45	*33.58*	97	37	*27.61*
	Overall (%)	*95.53*	*4.47*	*88.67*	*98.13*	*1.87*	*87.53*	*93.87*	*6.13*	*89.29*	*95.11*	*4.89*	*88.88*
2021	observed	0	1	*Correct (%)*	0	1	*Correct (%)*	0	1	*Correct (%)*	0	1	*Correct (%)*
Training	0	1,916	26	*98.66*	1,937	5	*99.74*	1,922	20	*98.97*	1,929	13	*99.33*
	1	234	96	*29.09*	280	50	*15.15*	214	116	*35.15*	251	79	*23.94*
	Overall (%)	*94.63*	*5.37*	*88.56*	*97.58*	*2.42*	*87.46*	*94.01*	*5.99*	*89.70*	*95.95*	*4.05*	*88.38*
Testing	0	821	11	*98.68*	831	1	*99.88*	820	12	*98.56*	825	7	*99.16*
	1	92	50	*35.21*	116	26	*18.31*	81	61	*42.96*	91	51	*35.92*
	Overall (%)	*93.74*	*6.26*	*89.43*	*97.23*	*2.77*	*87.99*	*92.51*	*7.49*	*90.45*	*94.05*	*5.95*	*89.94*

Table 2 Bankruptcy prediction results

For model comparison, the focus is on the AUC metric for the test dataset (**Table 3**). Upon comparing the methods, it is evident that SVM achieves the weakest fit (0.56 for 2019 and 0.59 for 2021). Conversely, DT emerges as the most effective (0.66 for 2019 and 0.71 for 2021), followed by ANN (0.63 and 0.68), and relatively LR (0.62 and 0.67). The selection of predictors was contingent upon meeting the condition of non-correlation, which is also the minimum requirement for LR. Multicollinearity diagnostics showed a low correlation of predictors, with all VIF values being less than 4.

AUC		LR	*	**	SVM	*	**	DT	*	**	ANN	*	**
2019	test	0.62	0.57	0.67	0.56	0.51	0.61	0.66	0.61	0.71	0.63	0.58	0.68
	overall	0.63	0.60	0.66	0.55	0.53	0.58	0.68	0.66	0.71	0.63	0.61	0.66
2021	test	0.67	0.62	0.72	0.59	0.54	0.64	0.71	0.66	0.75	0.68	0.63	0.72
	overall	0.65	0.62	0.67	0.58	0.55	0.60	0.68	0.66	0.71	0.63	0.61	0.66

Table 3 Result comparison by AUC metric

*Note: * and ** denote lower and upper limit of 95% confidence interval*

According to all models, the CL/TA predictor exhibits the highest significance. These findings suggest that these datasets contain the most critical information regarding bankruptcy prediction within this predictor. The relative importance of CL/TA in the DT models is 71% (2019) and 79% (2021), showing consistent prominence across other models as well. The second most significant information carrier is NWC/TA, contributing 23% (2019) and 21% (2021) to the predictive power of the models. Other predictors demonstrate only negligible added value in the models.

5 Conclusion

From the achieved results, it is inconclusive to determine the superiority of AI over logistic regression. The outcome depends on the specific method employed. In the case of SVM, logistic regression exhibits higher accuracy, while with ANN, the accuracy is comparable, differing only at the level of statistical error. Only DT demonstrates significantly better results. Thus, it can be concluded that AI cannot definitively be considered significantly more accurate in predicting bankruptcy overall. Individual methods must be carefully considered, and logistic regression still achieves comparable performance in this context.

The contribution of this study lies in the development of prediction models tailored to the specific conditions of the metallurgical industry in the Slovak Republic. These models can assist managers in early problem detection, thereby facilitating the industry's further development. Additionally, this study provides a comparison of various methods, highlighting the ongoing discrepancy in opinions regarding their quality.

However, this work has certain limitations. The non-uniformity of the samples poses a challenge, suggesting the need to balance individual sample groups using resampling methods or generating synthetic samples to enhance model validity. Moreover, the selection of predictors was constrained by the requirement for uncorrelated variables, a condition necessary for logistic regression but not for AI. Utilizing more robust techniques such as LASSO or Evolutionary Algorithms for predictor selection could yield improved results and is recommended for future studies.

Acknowledgements

This paper was supported by the research grant KEGA no. 001PU-4/2022 “Application of Modern Trends in Quantitative Methods in the Teaching of Financial and Managerial Subjects”.

References

Altman E I (1968) Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. The Journal of Finance 23(4):589–609. https://doi.org/10.2307/2978933

Becerra-Vicario R, Alaminos D, Aranda E, Fernández-Gámez M A (2020) Deep Recurrent Convolutional Neural Network for Bankruptcy Prediction: A Case of the Restaurant Industry. Sustainability 12(12):5180. https://doi.org/10.3390/su12125180

Ben Jabeur S (2017) Bankruptcy prediction using Partial Least Squares Logistic Regression. Journal of Retailing and Consumer Services 36:197–202. https://doi.org/10.1016/j.jretconser.2017.02.005

Cortes C, Vapnik V (1995) Support-vector networks. Machine learning 20(3):273-297. https://doi.org/10.1007/bf00994018

du Jardin P (2018) Failure pattern-based ensembles applied to bankruptcy forecasting. Decision Support Systems 107:64–77. https://doi.org/10.1016/j.dss.2018.01.003

Fitzpatrick P (1932) A comparison of ratios of successful industrial enterprises with those of failed firms. Certified Public Accountant 2:598–605.

Gajdosikova D, Valaskova K, Lazaroiu G (2024) The Relevance of Sectoral Clustering in Corporate Debt Policy: The Case Study of Slovak Enterprises. Administrative Sciences 14(2):26. https://doi.org/10.3390/admsci14020026

Garcia J. 2022 Bankruptcy prediction using synthetic sampling. Machine Learning with Applications. 9:100343. https://doi.org/10.1016/j.mlwa.2022.100343

Gavurova B, Jencova S, Bacik R, Miskufova M, Letkovsky S (2022) Artificial intelligence in predicting the bankruptcy of non-financial corporations. Oeconomia Copernicana 13(4):1215–1251. https://doi.org/10.24136/oc.2022.035

Jenčová S, Petruska I, Lukacova M, Abu-Zaid J (2021) Prediction of Bankruptcy in Non-financial Corporations Using Neural Network. Montenegrin Journal of Economics 17(4):123-134.

Klieštik T, Nica E, Durana P, Popescu GH (2023) Artificial intelligence-based predictive maintenance, time-sensitive networking, and big data-driven algorithmic decision-making in the economics of Industrial Internet of Things. Oeconomia Copernicana 14(4):1097-1138. https://doi.org/10.24136/oc.2023.033

Korol T (2020) Long-term risk class migrations of non-bankrupt and bankrupt enterprises. Journal of Business Economics and Management 21(3):783–804. https://doi.org/10.3846/jbem.2020.12224

López Iturriaga F J, Sanz I P (2015) Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks. Expert Systems with Applications 42(6):2857–2869. https://doi.org/10.1016/j.eswa.2014.11.025

Mihalovič M (2018) Applicability of scoring models in firms' default prediction. the case of Slovakia. Politická Ekonomie 66(6):689–708. https://doi.org/10.18267/j.polek.1226

Odom M D, Sharda R (1990) A Neural Network Model for Bankruptcy Prediction. In IJCNN International Joint Conference on Neural Networks, pp. 163–168.

Ohlson J A (1980) Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research 18(1):109–131. https://doi.org/10.2307/2490395

Ptak-Chmielewska A (2019) Predicting Micro-Enterprise Failures Using Data Mining Techniques. Journal of Risk and Financial Management 12(1):30. https://doi.org/10.3390/jrfm12010030

Rosenblatt F (1962) Principles of Neurodynamics. Spartan Books, New York.

Sigrist F, Leuenberger N (2023) Machine learning for corporate default risk: Multi-period prediction, frailty correlation, loan portfolios, and tail probabilities. European Journal of Operational Research 305(3):1390–1406. https://doi.org/10.1016/j.ejor.2022.06.035

Shin K-S, Lee T S, Kim H (2005) An application of support vector machines in bankruptcy prediction model. Expert Systems with Applications 28(1):127–135. https://doi.org/10.1016/j.eswa.2004.08.009

Thanh-Long N, Minh T, Hong-Chuong L (2022) A Back Propagation Neural Network Model with the Synthetic Minority Over-Sampling Technique for Construction Company Bankruptcy Prediction. International Journal of Sustainable Construction Engineering and Technology 13(3):68-79. https://doi.org/10.30880/ijscet.2022.13.03.007

Tumpach M, Surovičová A, Juhászová Z, Marci A, Kubaščíková Z (2020) Prediction of the Bankruptcy of Slovak Companies Using Neural Networks with SMOTE. Ekonomický časopis/Journal of Economics 68(10):1021-1039. https://doi.org/10.31577/ekoncas.2020.10.03

Valašková K, Klieštik T, Kováčová M (2018) Management of financial risks in Slovak enterprises using regression analysis. Oeconomia Copernicana 9(1):105-121. https://doi.org/10.24136/oc.2018.006

Vochozka M (2020) Metody komplexního hodnocení podniku. Praha: Grada.

Zmijewski M (1984) Methodological issues related to the estimation of financial distress prediction models. Journal of Accounting Research 22:59–82. https://doi.org/10.2307/2490859

Applying the FIML dynamic structural model in tourism industry

Lukáš Malec[1]

Abstract. Even though the standard approaches of structural techniques can offer pseudo-maximum likelihood estimates in time series data, better criteria are needed under current technological development. The dynamic structural equation model (DSEM) used in this study originates from Ciraki, D. (2007): Dynamic structural equation model, Estimation and inference. This procedure enables the lagged latent endogenous as well as exogenous variables to arrive at a solution in one process, together with variances of model errors. We constructed one variance-covariance matrix for the entire vectorized dataset, and the likelihood is then evaluated for a single observation. Because such matrices generally suffer from numerical problems, regularization has been introduced. The initial evaluation using the 3SLS approach based on stationary methodology and identification are both performed in the observed form. Despite the indisputable advantages of the method, computational difficulty is probably the reason for the full dynamic system not yet being incorporated into econometric packages.

Keywords: dynamic latent systems, structural equations, tourism

JEL Classification: C3, C5, C6, Z3
AMS Classification: 91G70

1 Introduction

With its connection to society, the tourism industry accelerates national economies, provides work opportunities, and positively influences culture in a predominant number of examples. This complex phenomenon, covering many direct, indirect, and induced effects, forms an ideal environment for studying structural dynamic relationships. Tourism can be analyzed based on various parameters, e.g., gross domestic product, exchange rate, consumer price index, tourism price given relative to the country of origin, living cost, and the price of competing destinations as the explanatory variables (Song, Witt and Li, 2008). Also, climate, country of origin, and visitor allowance are applied. Major studies focus on arrivals for forecasting and interpretational purposes, but only a limited number of overnight stays. In this sense, we moreover present an approach to applying latent construct to achieve robustness and better involvement in reality. Regarding the current digital technology era, data can be gathered more intensively and with high density, where novel approaches are requested, capturing a higher computational capacity. They cover, e.g., Monte Carlo simulations, neural networks, and structural equations. This study applies the model first mentioned by Ciraki (2007), p. 87. The area of dynamic factor models is in some sense similar (Stock and Watson, 2011), where mention should be made of several methods of parameter estimation, especially maximum likelihood and Kalman filters, as well as techniques based on principal components, and Bayesian methods, specifying a prior for integrating instead of searching extrema. Originally, Geweke (1977), Sargent and Sims (1977), and Forni et al. (2000) dealt with the estimation of dynamic factor models. Although some of them are nonparametric, they can achieve consistent estimates. Note, the multi-subject Toeplitz method can also be applied for pure time series in structural equations, resulting in moment estimates.

DSEM studies the processes of latent variables over time, covering incorporated the lagged relationships. While time series modeling is generally based on correlations caused by the proximity of observations, structural methodologies are concentrated on relationships between variables. Even though this study can be viewed as a generalization of the stationary structural approach, the theory and practice of our pure time series examined by latent variable methods should be presented more intensively. However, sampling independence cannot be guaranteed because the sample covariance matrix, known from stationary examples, is unavailable. For this reason, full information likelihood is incorporated in the sense that the solution vector must contain all observed data. The method is proposed here due to an empirical covariance matrix based solely on parameters, including covariance matrices

[1]Prague University of Economics and Business, Department of Statistics and Probability, W. Churchill Sq. 1938/4, 130 67 Prague, Czech Republic, lukas.malec@vse.cz

of errors for observed and latent variables, by shifting matrix and vec operator. Note that such system methods are more sensitive to misspecification of the model, e.g., omitting a relevant explanatory variable from the model. Despite the maximum likelihood estimate having a relatively good small sample properties, departures from normality are especially significant in such small samples. This is contrary to cross-sectional techniques (Asparouhov and Muthén, 2023). Although a series of equality constraints can solve the time series relationships among variables in the other environments to get a stationary dynamic structure, discussed in other studies, we apply the R programme with a full information estimation procedure on the score vector. This process still needs to be implemented into the computational package.

FIML and 3SLS are known to be asymptotically equivalent in the sense of the same asymptotic distribution under the given assumptions (Mardia, Kent and Bibby, 1979). Due to some statistical difficulties, the 3SLS and 2SLS estimators are usually applied using the instrumental variables methodology. Identification is also important for consistent estimates covering nonstationary simultaneous equation systems and stationary cases as well, although we concentrate on stationary simultaneous equations in this topic. Simultaneous equation systems for analyses and forecasting are applied in many branches, such as diverse academic works, private banks, and government economics departments. An alternative is various vector autoregression systems, but incorporated in a reduced form not directly connected to economic theory. We apply differenced data, thus expecting autoregressive processes whose roots lie on the boundary or within the unit circle. Such data are expected to be mean and covariance stationary, as the first two theoretical moments are defined and constant. We omit weak exogeneity and distributional discussion, as such assumptions are perceived as fulfilled.

2 Data used and methods

The data are gathered from Eurostat (European Statistical Office, 2024) and cover the Czech Republic for annual time series from 1995 to 2022. The first two observed endogenous variables consisted of nights spent and arrivals to hotels, holiday and other short-stay accommodations, camping grounds, recreational vehicle parks, and trailer parks. Then, length of stay is the next variable, together with the non-residents and residents ratio, where both originated from arrivals. The only exogenous variable is the artificial measure of crises. Such parameter, connected to the main consequences in the tourism industry, covers the 2000–2001 Early 2000s Recession, the 2007–2009 Global Financial Crisis, and the pandemic till 2022. Although strongly skewed and often modeled as a probit in the practice of structural equations, this variable serves properly as direct input to the structural dynamic model. To approximately eliminate the unit root, we take the data in the first differences, which are then standardized. We omit any intercept for descriptive purposes. It is assumed that many variables are perfectly measured by a single indicator, omitting observed errors in the model. But by measuring the target latent by one variable only, we ignore the mutual action of individual components and their contributions, also suppressing the specific variance generally. Moreover, latent constructs given by two observed variables (where one is the indicator) are sometimes problematic for interpretational reasons, although here they behave properly.

In optimization, we use the *optimx* library with some functionality for solving the algorithm and a maximum lag 3. We measure the adjusted residual sum as a model fit that is penalized for complexity despite the fact that the tests for normal distribution of residues can also be performed. Generally, two equivalent ways are known to characterize a rough property for square matrix inversion, as the matrix is of full rank and has a nonzero determinant. In our case, the empirical covariance matrix due to the FIML approach is sometimes ill-conditioned for the consequent application of the regularization process. The initial evaluation in this study is performed via stationary methodology in such a system by 3SLS in the *systemfit*. A known advantage in the simultaneous equation system is that the exogenous variables have unit variances in the case of standardization. As instrumental variables, the exogenous one and then all lagged endogenous ones are selected. Although not applied here, some of the parameters corresponding to errors can also be evaluated from the simultaneous model through essential transformations. According to Judge et al. (1988), the overidentified system is revealed for all cases, evaluated from the simultaneous system in observed form.

2.1 The model

For $t = 1, 2, \dots, T$, we apply the DSEM (Ciraki, 2007) introduced, e.g., in the recursive form as follows

$$\eta_t = \sum_{i=0}^{p} B_i \eta_{t-i} + \sum_{i=0}^{q} \Gamma_i \xi_{t-i} + \zeta_t$$

$$\xi_t = \sum_{i=1}^{s} R_i \xi_{t-i} + \nu_t$$

$$y_t = \Lambda_y \eta_t + \varepsilon_t$$

$$x_t = \Lambda_x \xi_t + \delta_t$$

with parameter matrices $B_i\ (f \times f)$, $\Gamma_i\ (f \times g), \Lambda_y$ and Λ_x. The covariances of model errors covering latent and observed variables are incorporated into the matrices $\Psi\ (m \times m)$ and $\Theta\ (r \times r)$, respectively. Here, η_t and ξ_t, resp. y_t and x_t, are vectors of latent and observed variables. The second equation is the VAR(s) process for $s = max(p, q)$. The equations for y_t and x_t can then be perceived as individual factor analysis models.

Due to several reasons, the relationship should be rewritten in the form

$$\underbrace{\begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix}}_{h_t\ (m\times 1)} = \sum_{i=0}^{s} \underbrace{\begin{pmatrix} B_i & \Gamma_i \\ 0 & R_i \end{pmatrix}}_{C_i\ (m\times m)} \underbrace{\begin{pmatrix} \eta_{t-i} \\ \xi_{t-i} \end{pmatrix}}_{h_{t-i}\ (m\times 1)} + \underbrace{\begin{pmatrix} \zeta_t \\ \nu_t \end{pmatrix}}_{z_t\ (m\times 1)}$$

$$\underbrace{\begin{pmatrix} y_t \\ x_t \end{pmatrix}}_{w_t\ (r\times 1)} = \underbrace{\begin{pmatrix} \Lambda_y & 0 \\ 0 & \Lambda_x \end{pmatrix}}_{\Lambda\ (r\times m)} \underbrace{\begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix}}_{h_t\ (m\times 1)} + \underbrace{\begin{pmatrix} \varepsilon_t \\ \delta_t \end{pmatrix}}_{e_t\ (r\times 1)}$$

The $R_0 = 0$, $m = f + g$ and $r = n + k$. Then, the vectorization proceeds by shifting matrix S_T to get

$$\underbrace{vec\{h_t\}_1^T}_{H_T\ (mT\times 1)} = \left(\sum_{i=0}^{s} S_T^i \otimes C_i\right) \underbrace{vec\{h_t\}_1^T}_{H_T\ (mT\times 1)} + \underbrace{vec\{z_t\}_1^T}_{Z_T\ (mT\times 1)}$$

$$\underbrace{vec\{w_t\}_1^T}_{W_T\ (rT\times 1)} = (I_T \otimes \Lambda) \underbrace{vec\{h_t\}_1^T}_{H_T\ (mT\times 1)} + \underbrace{vec\{e_t\}_1^T}_{E_T\ (rT\times 1)}$$

We can present this expression with observed variables only on the left side of the equation

$$W_T = (I_T \otimes \Lambda)\left(I_{mT} - \sum_{i=0}^{s} S_T^i \otimes C_i\right)^{-1} Z_T + E_T$$

The vector of parameters is given by $\theta' = [(vecC_i)', (vec\Lambda)', (vec\Psi)', (vec\Theta)']$. Then, the empirical covariance matrix in the closed form is

$$\Sigma(\theta) = E(W_T W_T') = AX^{-1}BX'^{-1}A' + I_T \otimes \Theta,$$

where $A = I_T \otimes \Lambda$, $X = I_{mT} - \sum_{i=0}^{s} S_T^i \otimes C_i$ and $B = I_T \otimes \Psi$.

To solve the problem via likelihood, our objective function to be minimized is

$$F_{ML} = ln|\Sigma(\theta)| + W_T' \Sigma^{-1}(\theta) W_T$$

Covering minor printing errors in the original text, we introduce the final analytic expressions for the score vector solved by matrix derivative calculus given by Turkington (2002). However, the set of formulas needed for final expressions is mentioned first. The maximum likelihood function is scalar, and its derivation has the shape of a vector. The derivative proceeds according to the columns, while other ways also exist. The derivative of the function according to any matrix M is expressed by a partial derivative of $vec\ M$ or $vech\ M$. At the same time, the matrix function F has the shape of vec operator, trace, or logarithm. The fundamentals are rules for a derivative of product and function composition. The zero-one matrices are used, e.g. duplication one (D_a) given by (Turkington, 2002, p. 43) $vec\ M = D_a vech\ M$ for symmetric M of order a. There also holds (Turkington, 2002, p. 70)

$$\frac{\partial vec\ M}{\partial vech\ M} = D_a'$$

The commutation matrix (K_{ab}) to $M\ (a \times b)$ is an orthogonal matrix of dimension $(ab \times ab)$ for which $K_{ab} vec\ M = vec\ M'$ (Turkington, 2002, p. 30). Then, K_{ab}^* is the $devec$ operator, covering the assigned commutation matrix.

For substitution, the composite function $o = o \circ p(x)$ of the vectors gives (Turkington, 2002, p. 71 under the chain rule)

$$\frac{\partial o}{\partial x} = \frac{\partial p}{\partial x}\frac{\partial o}{\partial p}$$

and for $o = o \circ [p(x), q(x)\ ]$ is (Turkington, 2002, p. 71 under generalization of the chain rule)

$$\frac{\partial o}{\partial x} = \frac{\partial p}{\partial x}\frac{\partial o}{\partial p} + \frac{\partial q}{\partial x}\frac{\partial o}{\partial q}$$

In case of derivation of the product and elements of matrices $M\ (a \times\ b)$ and $N\ (b \times\ c)$ are scalar functions of a vector δ, then (Turkington, 2002, p. 72)

$$\frac{\partial vec\ MN}{\partial \delta} = \frac{\partial vec\ M}{\partial \delta}(N \otimes I_a) + \frac{\partial vec\ N}{\partial \delta}(I_c \otimes M')$$

During the solution, the next formulas are used as well (Turkington, 2002, pp. 70–74, 80 and 86):

$\frac{\partial vec\ MON}{\partial vec\ O} = N \otimes M'$

$\frac{\partial vec\ \ M^{-1}}{\partial vec\ M} = -(M^{-1} \otimes M'^{-1})$ for nonsingular square matrix M

$\frac{\partial vec\ MO^{-1}N}{\partial vec\ O} = -O^{-1}N \otimes O'^{-1}M'$ here for nonsingular square matrix O

$\frac{\partial vec\ M'OM}{\partial vec\ M} = (I_b \otimes OM)K_{bb} + I_b \otimes O'M$ for square matrix O and $M\ (a \times\ b)$

$\frac{\partial tr\ M'OM}{\partial vec\ M} = vec(O'M + OM)$ for square matrix O

$\frac{\partial log\ |M|}{\partial vec\ M} = vec\ M'^{-1}$ if $|M|$ is positive and $log\ |M|$ exists

The resulting analytical formulas to construct a score vector are as follows

$$\frac{\partial F_{ML}}{\partial vec\ C_i} = \left[K_{Tm}^{*}\left(I_{mT} \otimes S_T'^{i}\right) \otimes I_m\right](X^{-1} \otimes X'^{-1})\left[BX'^{-1} \otimes I_{mT} + (BX'^{-1} \otimes I_{mT})K_{mT,mT}\right]$$
$$\times (A' \otimes A')[vec\ \Sigma^{-1}(\theta) - (\Sigma^{-1}(\theta) \otimes \Sigma^{-1}(\theta))(\ W_T \otimes W_T)]$$
$$\frac{\partial F_{ML}}{\partial vec\ \Lambda} = (K_{Tm}^{*} \otimes I_r)\left[(X^{-1}BX'^{-1}A' \otimes I_{rT}) + (X^{-1}BX'^{-1}A' \otimes I_{rT})K_{rT,rT}\right]$$
$$\times [vec\ \Sigma^{-1}(\theta) - (\Sigma^{-1}(\theta) \otimes \Sigma^{-1}(\theta))(\ W_T \otimes W_T)]$$
$$\frac{\partial F_{ML}}{\partial vech\ \Psi} = D_m'(K_{Tm}^{*} \otimes I_m)(X'^{-1}A' \otimes X'^{-1}A')[vec\ \Sigma^{-1}(\theta) - (\Sigma^{-1}(\theta) \otimes \Sigma^{-1}(\theta))(\ W_T \otimes W_T)]$$
$$\frac{\partial F_{ML}}{\partial vech\ \Theta} = D_r'(K_{Tr}^{*} \otimes I_r)[vec\ \Sigma^{-1}(\theta) - (\Sigma^{-1}(\theta) \otimes \Sigma^{-1}(\theta))(\ W_T \otimes W_T)]$$

Covering the last relation, the diagonal structure can be imposed on Θ matrix to get

$$\frac{\partial F_{ML}}{\partial vech\ \Theta} = D_{diag}(K_{Tr}^{*} \otimes I_r)[vec\ \Sigma^{-1}(\theta) - (\Sigma^{-1}(\theta) \otimes \Sigma^{-1}(\theta))(\ W_T \otimes W_T)]$$

3 Results and discussion

Two fundamentally different models are discussed in this study, first with latent variable η_1 incorporated for endogenous observed variables in the form of arrivals (y_1, as an indicator) and length of stay (y_2). Below, this model will be titled as structural. The second is accounting for η_1 only in nights spent (y_1 remarked) to form a pure simultaneous formulation. Then, the non-residents and residents ratio (y_3) gives another endogenous latent construct η_2, where the exogenous latent variable ξ_1 is determined by various crisis periods (x_1). Note that our simulation study does not apply the matrix R_i. But despite some interpretational drawbacks, its numerical inclusion to solve the whole problem works well. Table 1 introduces an arrangement of the model. We discuss the relationships for coefficients > 0.3 in magnitude. Covering the following tables, if introduced, the first line always corresponds to the final and the second line to initial estimates. In the first model (Table 2 introduces the results), the objective function ends in 48.056 with 48 iterations, covering 64 function and 49 gradient evaluations. The adjusted residual absolute sum is 3532.069.

C_0	$\eta_{1,0}$	$\eta_{2,0}$	$\xi_{1,0}$
$\eta_{1,0}$	0	$\beta_{12,0}$	0
$\eta_{2,0}$	0	0	$\gamma_{21,0}$
$\xi_{1,0}$	0	0	0

C_1	$\eta_{1,1}$	$\eta_{2,1}$	$\xi_{1,1}$
$\eta_{1,1}$	$\beta_{11,1}$	$\beta_{12,1}$	$\gamma_{11,1}$
$\eta_{2,1}$	0	$\beta_{22,1}$	0
$\xi_{1,1}$	0	0	0

C_2	$\eta_{1,2}$	$\eta_{2,2}$	$\xi_{1,2}$
$\eta_{1,2}$	0	0	$\gamma_{11,1}$
$\eta_{2,2}$	0	0	0
$\xi_{1,2}$	0	0	0

Λ	$\eta_{1,0}$	$\eta_{2,0}$	$\xi_{1,0}$
y_1	1	0	0
y_2	λ_{21}	0	0
y_3	0	1	0
x_1	0	0	1

Ψ	$\eta_{1,0}$	$\eta_{2,0}$	$\xi_{1,0}$
$\eta_{1,0}$	ψ_{11}	0	0
$\eta_{2,0}$	0	ψ_{22}	0
$\xi_{1,0}$	0	0	ψ_{33}

Θ	y_1	y_2	y_3	x_1
y_1	θ_{11}	0	0	0
y_2	0	θ_{22}	0	0
y_3	0	0	θ_{33}	0
x_1	0	0	0	θ_{44}

Table 1 Structure of the models

Covering the structural model, the first latent endogenous variable is given by the arrivals indicator with negative influence for the length of stay, which can be perceived as peaks in visiting the country supplied by short-stay accommodation. Such a latent variable is influenced negatively by its autoregressive coefficient given by word-of-mouth effects and positively by the current as well as one-lagged non-residents share. The differences in final estimates, opposite to the initial ones, are probably caused by the composition of the latent variable, contrary to only one parameter in the simultaneous equation system given by arrivals. In this way, the short-stay peaks in arrivals are supplied by non-residents. While insignificant in the first latent variable, the crises are related negatively to the non-residents share, which is an expected conclusion. The negative autocorrelation coefficient exists for non-residents, balancing the equilibrium state. The error estimate for the first latent endogenous variable corresponds to a lower border in the constrained optimization task, but two others are relatively high. However, there are significantly nonzero errors in the observed variables for length of stay and arrivals. The errors of non-residents share and crises are zero, fitting the theory of latent systems per only one variable in the construct.

C_0	$\eta_{1,0}$	$\eta_{2,0}$	$\xi_{1,0}$
$\eta_{1,0}$	0	0.81 0.90	0
$\eta_{2,0}$	0	0	−0.51 −0.45
$\xi_{1,0}$	0	0	0

C_1	$\eta_{1,1}$	$\eta_{2,1}$	$\xi_{1,1}$
$\eta_{1,1}$	−0.83 0.17	0.38 −0.10	−0.17 0.13
$\eta_{2,1}$	0	−0.39 0.08	0
$\xi_{1,1}$	0	0	0

C_2	$\eta_{1,2}$	$\eta_{2,2}$	$\xi_{1,2}$
$\eta_{1,2}$	0	0	−0.20 −0.05
$\eta_{2,2}$	0	0	0
$\xi_{1,2}$	0	0	0

Λ	$\eta_{1,0}$	$\eta_{2,0}$	$\xi_{1,0}$
y_1	1	0	0
y_2	−0.82 −0.39	0	0
y_3	0	1	0
x_1	0	0	1

Ψ	$\eta_{1,0}$	$\eta_{2,0}$	$\xi_{1,0}$
	0.00 0.3	0.74 0.3	0.96 0.3

Θ	y_1	y_2	y_3	x_1
	0.33 0.1	0.53 0.1	0.00 0.1	0.00 0.1

Table 2 Results of structural model

In the second model (Table 3 introduces the results), the objective function ends in 74.366 with 38 iterations, covering 45 function and 38 gradient evaluations. The adjusted residual absolute sum is 4703.048. According to our assumptions, the objective function gains a higher value, but in the same way, the residual absolute sum handling the model complexity increases. Instead of a structural system, we have only one nights spent variable to form a latent construct in the simultaneous system. Herein, the differences from the preceding case are especially discussed. Again, non-residents share is decisive for the first latent endogenous variable, but only in the current state. The autoregression of this variable seems insignificant. The relationships for the second latent variable are similar despite the more pronounced autoregressive part. All latent errors seem important, from which the crises variable stands out. The error of y_2 variable is inactive, covering the whole system, due to omitting it from any considerations. It is interesting that the observed error for non-residents share is nonzero. However, we expect the simultaneous system is limited in interpretation.

C_0	$\eta_{1,0}$	$\eta_{2,0}$	$\xi_{1,0}$
$\eta_{1,0}$	0	0.82 0.88	0
$\eta_{2,0}$	0	0	-0.57 -0.45
$\xi_{1,0}$	0	0	0

C_1	$\eta_{1,1}$	$\eta_{2,1}$	$\xi_{1,1}$
$\eta_{1,1}$	0.09 0.11	-0.15 0.01	-0.25 0.12
$\eta_{2,1}$	0	-0.88 0.07	0
$\xi_{1,1}$	0	0	0

C_2	$\eta_{1,2}$	$\eta_{2,2}$	$\xi_{1,2}$
$\eta_{1,2}$	0	0	-0.03 -0.11
$\eta_{2,2}$	0	0	0
$\xi_{1,2}$	0	0	0

Λ	$\eta_{1,0}$	$\eta_{2,0}$	$\xi_{1,0}$
y_1	1	0	0
y_2	0	0	0
y_3	0	1	0
x_1	0	0	1

Ψ	$\eta_{1,0}$	$\eta_{2,0}$	$\xi_{1,0}$
	0.43 0.3	0.39 0.3	0.96 0.3

Θ	y_1	y_2	y_3	x_1
	0.00 0.1	0.96 0.1	0.23 0.1	0.00 0.1

Table 3 Results of simultaneous model

4 Conclusion

DSEM applied in this study is a fully competitive technique compared to others, which are usually utilized in real examples. After regularization, the numerical evaluation of the score vector in the specific example from tourism works well. The latent system better describes reality than the simultaneous one. The reason is probably the proper interpretation of short-stay peaks, where more coefficients are significant for a subsequent discussion of the results. Considering the final values of the objective function and residual sums are also utilized. Further use of the statistical induction and extension of mathematical part should be contributory.

References

Asparouhov T, Muthén B (2023) Residual structural equation models. Structural Equation Modeling: A Multidisciplinary Journal 30(1): 1-31

Ciraki D (2007) Dynamic structural equation models: Estimation and interference. United Kingdom, London School of Economics and Political Science

Forni M, Hallin M, Lippi M, Reichlin L (2000) The generalized dynamic-factor model: Identification and estimation. Review of Economics and Statistics 82(4): 540-554

Geweke J (1977) The dynamic factor analysis of economic time series. In DJ Aigner and AS Golberger (eds.) Latent Variables in Socio-Economic Models, North-Holland, Amsterdam

Judge GG, Hill RC, Griffiths WE, Lütkepohl H, Lee T (1988) Introduction to the theory and practice of econometrics. John Wiley & Sons, New York

Mardia KV, Kent JT, Bibby JM (1979) Multivariate analysis. Academic Press, Orlando

Sargent TJ, Sims CA (1977) Business cycle modeling without pretending to have too much a priori economic theory. New Methods in Business Cycle Research 1: 145-168.

Song H, Witt SF, Li G (2008) The advanced econometrics of tourism demand. Routledge, New York

Stock JH, Watson MW (2011). Dynamic factor models. University Press, Oxford

Turkington DA (2002) Matrix calculus and zero-one matrices. University Press, Cambridge

Construction of a new DEA-based Composite Index for Circular Economy Assessment in the EU

Markéta Matulová [1]

Abstract. In recent years, there has been growing interest in exploring the concept of the circular economy as a potential solution for enhancing the sustainability of our economic system. The development of circular economy indicators provides valuable insights allowing the evaluation of the progress on the path to circularity and sustainability. On the other hand, composite indicators often stir controversies about the unavoidable subjectivity that is connected with their construction. Usually, the normalized sub-indicators are just added, sometimes with certain weights associated with the sub-indicators. This study addresses existing research gap by focusing on the need for more robust and fair composite indices for circular economy assessment. We depart from usual approach and compute alternate composite index for EU countries using flexible weights obtained by Data Envelopment Analysis. Using flexible weighting can promote buy-in from relevant stakeholders, making the final results more widely accepted. Additionally, DEA-based indicator provides more information on the relative performance of evaluated units and offers implications such as identifying target values of sub-indicators or selecting peer units for benchmarking purposes.

Keywords: Circularity, Composite index, Data Envelopment Analysis

JEL Classification: Q51, Q56, Q60
AMS Classification: 90C90

1 Introduction

The increasing emphasis on sustainability in Europe has underscored the necessity for measuring and comparing Circular Economy (CE) performance across countries and tracking its changes in time. In January 2018, the European Commission introduced the EU monitoring framework for the circular economy (EUmonitor, 2018) composed of a set of (ten) key indicators grouped to four categories according to the stages of circular economy:

- Production and consumption
- Waste management
- Secondary raw materials
- Competitiveness and innovation

In 2023, the European Commission communication presented a revised monitoring framework (Eurostat, 2023) that captures interlinkages between circularity, climate neutrality and the zero pollution ambition. It also puts more focus on the production side rather than on waste by some methodological changes in several subindicators and using more footprint indicators. Moreover, it adds a new dimension: Global sustainability and resilience including several sub-indicators.

Despite the diverse methodologies for the construction of CE indices, with prevalent use of single indices, there is a growing demand from policymakers for aggregate indicators in the form of Composite Indices (CI). These are considered valuable tools that can be easily interpreted and communicated to the general public. The goal of this paper is to develop a new DEA-based composite index for assessing circular economy performance in EU countries, aiming to improve the objectivity and applicability of such indices. The rest of the paper is organized as follows: In the second chapter, we describe the data sample and two models used to construct the CE index. The third chapter presents the results and compares them across European countries. We compare our findings with other studies in the fourth chapter.

[1] Masaryk University, Faculty of Economics and Administration, Lipová 41a, Brno, Czechia, 8987@mail.muni.cz.

2 Methodology and Data

The approach used in our study is based on employing Data Envelopment Analysis (DEA is a tool originally designed for the efficiency evaluation of homogeneous decision making units) as an aggregation tool. In the context of composite indices, it was first used by Mahlberg and Obersteiner (2001) to reassess Human Development Index, and since then, DEA-based CIs has been used in many application areas, such as assessing European social inclusion policy (Cherchye, et al., 2004), technology achievement (Cherchye et al., 2008) or road safety (Shen et al., 2013). Many other applications are mentioned in the survey of Greco et al. (2019).

2.1 DEA models for the construction of the composite index

Basic properties of the DEA-based CIs are described in the paper of Cherchye et al. (2007), where it refers to the method as Benefit of the Doubt (BoD) approach. We will use two models of this kind. Next to the basic BoD model, we also employ extended model capturing the multilayer structure of the data for the composite index.

Model 1: One-layer Benefit of the Doubt DEA

To construct BoD composite index using *m* indicators (as we assume all of them to be of the benefit-type, we often refer to them as outputs) for *n* units (countries) denoting y_{ij} the value of indicator i in unit j, we define value of CI for unit j as the solution of the following linear program with v_i representing the weight of the i-th indicator:

$$CI_j = \max_{v_i} \sum_{i=1}^{n} v_i \, y_{ij}, \tag{1}$$

subject to $\sum_{i=1}^{n} v_i \, y_{ik} \leq 1, \ \forall k = 1, \dots, n, \ \ v_i \geq 0, \ \forall i = 1, \dots, m.$

We obtain $CI_j \leq 1$ for each unit j, with higher values indicating better relative performance. The indices of constraints binding in optimal solution identify peer units for "inefficient" units.

Model 2: Two-layer Benefit of the Doubt DEA

To comprise a multilayer index structure, where the indicators are composed of several sub-indicators and to provide better discrimination among decision making units (DMUs), lets consider an improved BoD DEA composite index. The structure of the model is depicted in Figure 1.

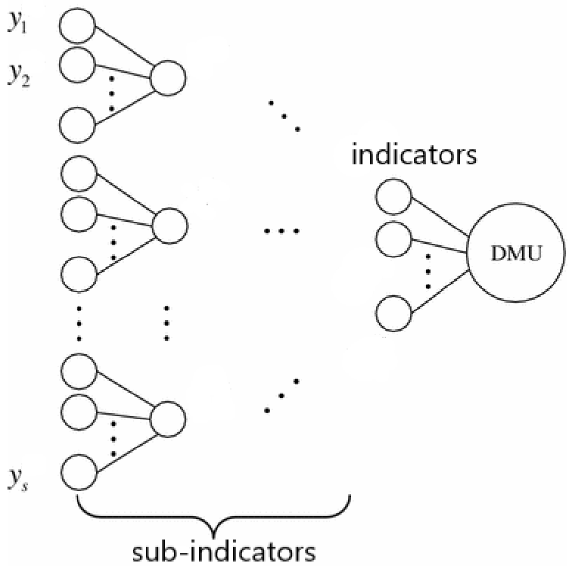


Figure 1 Multilayer index structure. Source: own elaboration according to Shen et al. (2013)

We use the following notation:

- L – a set of indicators;
- I_l – a set of sub-indicators creating an indicator $l \in L$;

- $0 \leq a_l \leq 1$ – upper bound of the performance level for an indicator $l \in L$, under the condition that $\sum_{\forall l \in L} a_l \geq 1$ to ensure feasibility.
- B_r, U_r – lower and upper bounds of the performance level for a sub-indicator r, imposed by decision-maker under the conditions that $\sum_{r \in \cup I_l} B_r \leq 1, \sum_{r \in \cup I_l} U_r \geq 1$

We define value of *CI* for unit *j* using y_{ik}, the values of sub-indicator i and v_i representing its weight:

$$CI_j = \max_{v_i} \sum_{i \in \cup I_l} v_i \, y_{ij} \tag{2}$$

subject to $\sum_{i \in \cup I_l} v_i \, y_{ik} \leq 1, \ \forall k = 1, \dots, n, \ \ v_i \geq 0, \ \forall i \in \bigcup_{l \in L} I_l,$

and $\sum_{i \in I_l} v_i \, y_{ij} \leq a_l, \forall l \in L, B_i \leq v_i \, y_{ij} \leq U_i, \forall i \in \bigcup_{l \in L} I_l$

The virtual indicator values in the inequalities show the relative contribution of each indicator or sub-indicator to the composite index. Their levels can be restricted subjectively by decision-makers or by setting data-driven bounds.

2.2 Data for the analysis

We use available data on the variables covered in the CE monitoring framework provided by Eurostat for the year 2019, comprising 28 EU countries. The individual dimensions are represented by 1-2 (sub)indicators. The overview and the description of the indicators is given in Table 1.

Dimension	(sub)indicators	Description
Production and consumption	Resource productivity	GDP divided by domestic material consumption
Waste management	Recycling rate of municipal waste	Share of recycled municipal waste in the total municipal waste generation
Secondary raw materials	Circular material use rate	Ratio of the circular use of materials to the overall material use
Competitiveness and Innovation	Persons employed in CE sectors	Share of persons employed in the recycling, repair, reuse, and rental sectors
	Private investment and gross added value related to CE sectors	Gross investment in tangible goods and value added at factor costs in CE sectors divided by GDP
Sustainability and resilience	Material import independency	Ratio of domestic material use over direct material inputs

Table 1 The indicators used for the construction of the CE composite index

Descriptive characteristics of the data are displayed in Table 2, the values are from the year 2019.

Indicator	Units	Mean	STD	Min	Max
Resource productivity	EUR/kg	1.94	1.23	0.36	4.96
Recycling rate of municipal waste	%	39.77	14.33	9.10	66.70
Circular material use rate	%	9.72	7.09	1.30	30.00
Employment in CE sectors	%	0.86	0.27	0.32	1.46
Private investment in CE sectors	%	0.78	0.36	0.20	1.63
Material import independency	%	60.53	19.07	8.90	90.50

Table 2 Descriptive statistics of the data

3 Results

Firstly, we applied **Model 1** to obtain plain Benefit of the Doubt index (BoD CEI). The interpretative value of the results was not satisfactory, as 16 out of 28 countries achieved the maximum value. To improve the discrimination power, we applied **Model 2** with the values $a_4 = 0.25$ and $B_i = 0.05, U_i = 0.25, \ i = 1, \ldots, 6$ to obtain two-layer DEA index (2L BoD CEI). In this case, the maximum value of 1 was achieved by three countries only. The results of both models are listed in Table 3 and the visualization of the results showing geographical distribution of the circularity performance is shown in Figure 2.

	BoD CEI	**2L BoD CEI**	**peers**		**BoD CEI**	**2L BoD CEI**	**peers**
Begium	1.00	1.00		Lithuania	1.00	0.82	IT, UK
Bugaria	1.00	0.76	DE	Luxembourg	0.95	0.73	BE
Czechia	0.98	0.86	DE, IT	Hungary	0.96	0.82	DE, IT
Demark	0.95	0.83	BE, DE	Malta	0.79	0.62	UK
Ger-many	1.00	0.96	IT, UK	Netherlands	1.00	0.97	BE, IT
Estonia	1.00	0.90	IT	Austria	1.00	0.86	BE, UK
Ireland	0.97	0.81	DE, IT, UK	Poland	1.00	0.86	DE, IT
Greece	0.81	0.67	IT, UK	Portugal	0.93	0.74	BE, DE
Spain	0.90	0.86	DE, UK	Romania	1.00	0.53	BE
France	1.00	0.94	BE, IT, UK	Slovenia	0.96	0.90	IT
Croatia	1.00	0.78	DE	Slovakia	0.81	0.77	DE
Italy	1.00	1.00		Finland	1.00	0.82	IT
Cyprus	0.86	0.65	IT, UK	Sweden	1.00	0.86	IT
Latvia	1.00	0.82	DE, UK	UK	1.00	1.00	

Table 3 Results of Model 1 and Model 2 for EU countries

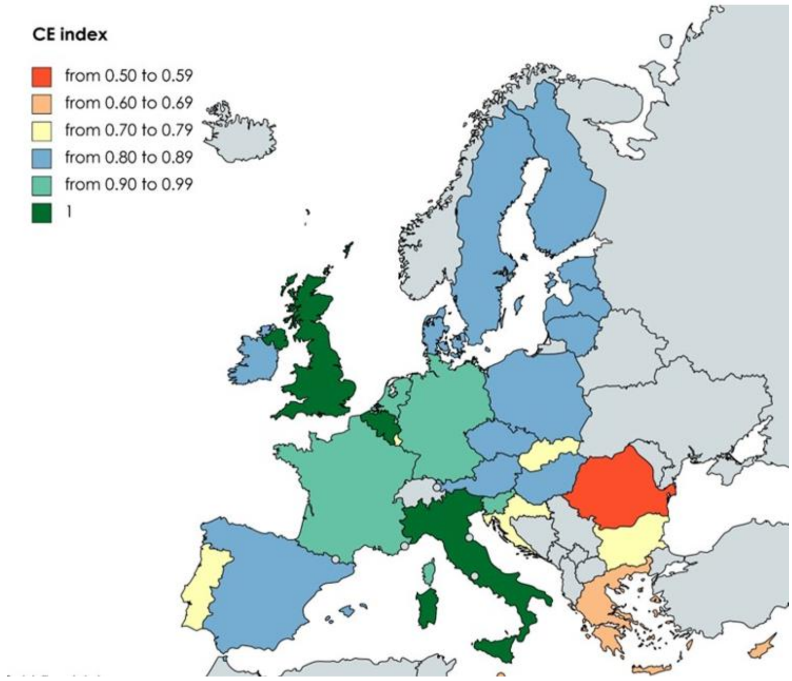


Figure 2 Values of 2LBoD circularity index in EU countries. Source: Own computations, mapchart.net

Belgium, Italy, and United Kingdom were the best performing countries in our 2LBoD model. On the other end of the ranking, we identified Romania (scoring 0.53), Malta (0.62), and Cyprus (0.65). The last two mentioned are island countries and this fact may be one of the most important factors limiting transition to circularity. There are also some more potential reasons for lagging behind. The countries from the bottom of the ranking usually lack comprehensive policies and regulations promoting circular economy practices. Limited investment in research and development for circular technologies and sustainable solutions contributes to a slower adoption of innovative practices. Public awareness and education about the benefits of circular practices may be insufficient, leading to a lack of citizen engagement and participation. Last, but not least, inadequate infrastructure for waste collection, sorting, and recycling can impede effective waste management, leading to higher disposal rates.

Efforts to improve circularity often involve a combination of policy changes, public engagement, technological advancements, and international collaboration. It's crucial to recognize that the assessment of circularity is multi-faceted and context-specific. Addressing the challenges in each country requires a tailored approach that considers their unique socio-economic, political, and environmental conditions. We hope that analyses like the one presented in this paper will help to find potential areas for improvement and identify peer countries for the benchmarking purposes.

4 Discussion and conclusions

This article is not the sole study addressing the construction of the Circularity Composite Index for EU countries using the Data Envelopment Analysis methodology. To the best of our knowledge, there are five other studies tackling a similar subject, albeit within different contexts. Notably, none of these studies employs a multilayer structure like we do. The uniqueness of our approach also lies in its foundation on standardized data collected using the same methodology, thereby preventing distortions in comparisons over time or between countries. This is not the case in the other studies that do not utilize data from Eurostat database of Circular Economy Monitoring Framework (Eurostat, 2023). Table 4 provides a comprehensive overview of these studies, including the (sub)indicators utilized in constructing the CI.

Study	Indicators
Milanović et al. (2022)	- Municipal waste recycling rate, Recycling rate of total waste excluding mineral waste, Recycling rates for specific waste streams
Banjerdpaiboon and Limleamthong (2023)	- Generation of municipal waste per capita, Recycling rate of municipal waste, Generation of waste excluding major mineral wastes per GDP, Recycling rate of packaging waste, Recycling of biowaste per capita, Circular material uses rate
Giannakitsidou et al. (2020)	- MSW generated, Recycling rate of MSW, Basic human needs, Circular material use rate, Foundations of wellbeing opportunity
Temerbulatova et al. (2021)	- Generation of municipal waste per capita, Recycling rate of municipal waste, Water exploitation index, Circular material uses rate, Final energy consumption, Social Progress Index
Halkos and Petrou (2019)	- Labor force, GDP, Investment, Population density, Waste

Table 4 Comparison with other CE indexes in EU

We present the correlations of our results with those of other authors in Table 5.

	2LDEA	MIL2019	BAN2023	GIA2020	TEM2021	HAL2019
2LDEA	1.00	0.66	0.33	0.76	0.55	-0.32
MIL2019	0.66	1.00	0.37	0.75	0.32	-0.41
BAN2023	0.33	0.37	1.00	0.54	0.45	-0.12
GIA2020	0.76	0.75	0.54	1.00	0.64	-0.13
TEM2021	0.55	0.32	0.45	0.64	1.00	0.27
HAL2019	-0.32	-0.41	-0.12	-0.13	0.27	1.00

Table 5 Correlations with other CE indexes in EU

The strongest positive relationship can be observed between our index and CI constructed by Giannakitsidou et al. (2020). On the other side, very different results were obtained by Halkos and Petrou (2019); the correlations of their CI with other indices is even negative in most cases. This is probably caused by the opposite interpretation of the DEA efficiency scores and using different indicators for its construction.

To conclude, measurement of the circularity at the national level is crucial for reaching sustainable development growth. It provides valuable information to governments that play a vital role in shaping and implementing policies facilitating the transition from linear economies to CE-based systems. Our study contributes to this effort by constructing a composite CE index for European countries using approach that has many advantages. Application of DEA helps to overcome two significant issues of basic aggregation methods in index construction: subjective weighting used for aggregation and the the undesirable reliance on preliminary normalization of sub-indicators. Additionally, flexible weighting promotes buy-in from relevant stakeholders, making the final results more widely accepted. Lastly, DEA analysis offers valuable insights into the relative performance of the evaluated units, such as identifying target values and peer units for those that are inefficient.

References

Banjerdpaiboon A, Limleamthong P (2023) Assessment of national circular economy performance using super-efficiency dual data envelopment analysis and Malmquist productivity index: Case study of 27 European countries. Heliyon 9:e16584. https://doi.org/10.1016/j.heliyon.2023.e16584

Cherchye L, Moesen W, Rogge N, et al (2008) Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index. Journal of the Operational Research Society 59:239–251. https://doi.org/10.1057/palgrave.jors.2602445

Cherchye L, Moesen W, Rogge N, Puyenbroeck TV (2007) An Introduction to 'Benefit of the Doubt' Composite Indicators. Soc Indic Res 82:111–145. https://doi.org/10.1007/s11205-006-9029-7

Cherchye L, Moesen W, Van Puyenbroeck T (2004) Legitimately Diverse, yet Comparable: On Synthesizing Social Inclusion Performance in the EU*. J of Common Market Studies 42:919–955. https://doi.org/10.1111/j.0021-9886.2004.00535.x

Giannakitsidou O, Giannikos I, Chondrou A (2020) Ranking European countries on the basis of their environmental and circular economy performance: A DEA application in MSW. Waste Management 109:181–191. https://doi.org/10.1016/j.wasman.2020.04.055

Greco S, Ishizaka A, Tasiou M, Torrisi G (2019) On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness. Soc Indic Res 141:61–94. https://doi.org/10.1007/s11205-017-1832-9

Halkos G, Petrou KN (2019) Assessing 28 EU member states' environmental efficiency in national waste generation with DEA. Journal of Cleaner Production 208:509–521. https://doi.org/10.1016/j.jclepro.2018.10.145

Mahlberg B, Obersteiner M (2001) Remeasuring the HDI by Data Envelopement Analysis

Milanović T, Savić G, Martić M, et al (2022) Development of the Waste Management Composite Index Using DEA Method as Circular Economy Indicator: The Case of European Union Countries. Pol J Environ Stud 31:771–784. https://doi.org/10.15244/pjoes/139896

Shen Y, Hermans E, Brijs T, Wets G (2013) Data Envelopment Analysis for Composite Indicators: A Multiple Layer Model. Soc Indic Res 114:739–756. https://doi.org/10.1007/s11205-012-0171-0

Temerbulatova ZhS, Zhidebekkyzy A, Grabowska M (2021) Assessment of the Effectiveness of the European Union Countries Transition to a Circular Economy: Data Envelopment Analysis. jour 16:142–151. https://doi.org/10.51176/1997-9967-2021-3-142-151

Eurostat (n.d.) Circular Economy - Overview. https://ec.europa.eu/eurostat/web/circular-economy/overview. Accessed 25 May 2024a

Explanatory Memorandum to COM(2018)29 - Monitoring framework for the circular economy - EU monitor. https://www.eumonitor.eu/9353000/1/j4nvhdfdk3hydzq_j9vvik7m1c3gyxp/vkl2lgcz44w4. Accessed 25 May 2024b

Dynamic Portfolio Optimization Under Robust Second Order Stochastich Dominance Model

Jakub Neugebauer[1]

Abstract. In this paper, different dynamic portfolio optimization strategies are examined under the usage of the robust second order stochastic dominance (RSSD) model, with a particular focus on its application to the stocks that are included in the S&P 500 index. In contrary to traditional investment approaches, which often rely on assumptions about historical returns and volatility, the RSSD model is recognized for its advanced handling of uncertainties and the dynamic characteristics of financial markets, proposing a good base for constructing portfolios. By subjecting various dynamic investing strategies to simulation and comparing their performances to the S&P 500 index benchmark, the potential of the RSSD model in augmenting portfolio returns and enhancing risk management is evaluated. Comprehensive analysis is employed with the aim of demonstrating how the RSSD model can be utilized by investors to make more informed and effective investment decisions, thereby optimizing the performance of their investment portfolios in the midst of market fluctuations. The outcomes of this research are intended to contribute significant insights for both academic investigation and practical implementation in the field of finance.

Keywords: lorem, ipsum, dolor, sit, amet

JEL Classification: C44, G11
AMS Classification: 90C15, 90-10

1 Introduction

In the world of finance, portfolio optimization is one of the core tools for both individual and company investors looking to gain returns from their investment while minimizing underlying risk. Traditional methods of portfolio optimization, like the Mean-Variance model proposed by Markowitz, 1952, have been the base of the modern portfolio theory. Nevertheless, these models usually stands on the assumptions of normally distributed returns of investment, which are typically in reality not met. As a response to these limitation, many advanced models have been proposed, such as Mean-Semivariance model, Mean-VaR model and others.

In this paper, we will focus more on the method, which is does not belong into the family of mean-risk models. This method is derived from concept of stochastic dominance and its typical form used lately in portfolio optimization, second order stochastic dominance (SSD) shall be the main objective of this research together with it´s robust version, robust second order stochastic dominance (RSSD). SSD has a benefit of not looking only on one measure of a risk, but rather comparing whole distributions of returns of a given benchmark and desired portfolio and thus provides sophisticated framework for portfolio optimization.

The goal of this paper is to analyze different strategies of dynamic portfolio optimization. We will be looking at two factors. One of them is the selected portfolio optimization model and the other is the frequency of regular investments. This paper proposes a methodology to compare different investment frequencies, ensuring fair conditions for comparison. Simulation will be performed on historical data of S&P 500 index and it´s components. S&P will be chosen as a benchmark to which simulated portfolio will be build against. Key questions in this paper are: Is RSSD more effective than SSD model? Does frequency of investment effect returns in long-term? Is some frequency more effective than others? Different performance indicators will than be used to make conclusion.

By demonstrating the practical benefits of the (R)SSD model and dynamic reevaluation of portfolio structure, this paper aims to provide valuable guidance for investors looking to make more informed and effective investment decisions in the evolving landscape of financial markets.

[1] Prague University of Economics and Business, Faculty of Informatics and Statistics, nám. Winstona Churchilla 1938/4, 120 00 Praha 3-Žižkov, jakub.neugebauer@vse.cz

2 Second Order Stochastic Dominance

The idea of stochastic dominance is a useful decision-making technique that compares two random variables using their complete probability distributions. Compared to standard statistics like mean, variance, VaR, and so on, this offers more reliable information since it compares all possible outcomes of random variables rather than just one chosen number to determine which is "better". The notion of stochastic dominance has its origins in Hadar and Russell 1969 , who provided the foundation for the now popularized term. Definiton of stochastic dominance is following:

Let X and Y be random variables with cumulative distribution functions $F_X(x)$, respectively $F_Y(x)$, Let $D_X^1(x) \equiv F_X(x)$ and:

$$D_X^s(x) = \int_{-\infty}^{x} D_X^{s-1}(t)dt, \quad s \in \mathbb{Z}^+, \quad s \geq 2$$

then we say that the random variable X stochastically dominates random variable Y at order s if $D_X^s(x) \leq D_Y^s(x)$ for all x. This is a case of weak stochastic dominance. In the case of strong stochastic dominance, we would need to add a condition, that $D_X^s(x) < D_Y^s(x)$ for at least one x. In this paper, only the case of weak stochastic dominance will be considered.

Dentcheva and Ruszczyński 2006 explains how stochastic dominance came to be used in portfolio optimization. They developed an expected shortfall-based linear programming (LP) issue and demonstrated that it is equal to a discrete distribution stochastic dominance problem. Although it is a method that is simply computed by any solver, its drawback is that it is limited to historical data. LP model is following:

$$max \sum_{t=1}^{T} \sum_{i=1}^{n} r_{it} w_i p_t \tag{1}$$

subject to:

$$d_{tk} \geq b_k - \sum_{i=1}^{n} r_{it} w_i, \quad t = 1, \cdots, T, \quad k = 1, \cdots, T, \tag{2}$$

$$\sum_{t=1}^{T} d_{tk} p_t \leq v_k, \quad k = 1, \cdots, T, \tag{3}$$

$$\sum_{i=1}^{n} w_i = 1, \tag{4}$$

$$d_{tk} \geq 0, w_i \geq 0, \quad i = 1, \cdots, n, \quad t, k = 1, \cdots, T. \tag{5}$$

In this case, T represents the total number of past observations of returns r_{it} for every asset i. A generalizing probability of a discrete return scenario derived at time t is denoted by p_t. We can set equal probability of each scenario, which will be used in this study, by setting $p_t = 1/T, \quad t = 1, \cdots, T$. Than $v_k = E((b_k - B)^+)$, where B is the random benchmark returns and b_k is the return of benchmark at time $k = 1, \cdots, T$. Lastly, there are auxiliary variables $d_{t,k}$.

By expressing asset returns as uncertain parameters and modifying them in defined intervals, Sehgal and Mehra 2020 presented a robust second order stochastic dominance (RSSD) linear programming model (LPRSSD) and demonstrated that the robust method beats SSD on chosen markets. We assume that the random return $\tilde{r}_{i,t}$ in the LPRSSD model can variate in the range $[r_{i,t} - \tilde{r}_{i,t}, r_{i,t} + \tilde{r}_{i,t}]$, where the perturbation to the nominal return of asset i at time t is denoted by $\hat{r}_{i,t}$. Robustness in terms of returns is attained in this manner. In addition, parameter Γ was introduced by Sehgal and Mehra 2020 as the total number of coefficients w_i. This parameter is unknown due to parametric uncertainty, as it is unlikely that all of the coefficients w_i are uncertain. The formula for LPRSSD is as follows:

$$max \sum_{t=1}^{T} \sum_{i=1}^{n} r_{it} w_i p_t - \sum_{i=1}^{n} s_i - \Gamma z_0 \tag{6}$$

subject to:

$$b_k - d_{tk} - \sum_{i=1}^{n} r_{it} w_i, + \sum_{i=1}^{n} s_{it} + \Gamma z_t \le 0 \quad t = 1, \cdots, T, \quad k = 1, \cdots, T, \tag{7}$$

$$z_0 + s_i \ge \sum_{t=1}^{T} \hat{r}_{i,t} w_i p_t, \quad i = 1, \cdots, n \tag{8}$$

$$z_t + s_{i,t} \ge \hat{r}_{i,t} w_i, \quad i = 1, \cdots, n, \quad t = 1, \cdots, T \tag{9}$$

$$(3) - (5)$$

$$z_0, z_t, s_{i,t}, s_i \ge 0, \quad i = 1, \cdots, n, \quad t = 1, \cdots, T. \tag{10}$$

Where $z_0, z_t, s_{i,t}, s_i$ are new auxiliary variables. Even while LPRSSD is still an LP model, it has far more variables and restrictions than LPSSD, which makes it even more difficult. A cutting plane algorithm was proposed by Sehgal and Mehra 2020 to address the complexity of this model.

Both of these presented models will be evaluated and compared in the simulation of different dynamic investment strategies, which are presented in the following section.

3 Dynamic Investment Strategies

In this study, several dynamic portfolio optimization strategies are proposed to evaluate their performance under different conditions. By dynamic portfolio optimization, regular and frequent investments are meant, on the contrary to one-time or infrequent allocations. The focus is on two main factors: the selected portfolio optimization model and the frequency of regular investments. The following strategies have been designed and will be tested:

- **Strategy 1:** Only buying new assets is allowed while the old portfolio is selected as a benchmark. The model used for optimization is the Second Order Stochastic Dominance (SSD) model.
- **Strategy 2:** Reallocation of the entire budget is allowed while the old portfolio is selected as a benchmark. The model used for optimization is the SSD model.
- **Strategy 3:** Reallocation of the entire budget is allowed while the S&P 500 index is selected as a benchmark. The model used for optimization is the SSD model.
- **Strategy 4:** Reallocation of the entire budget is allowed while the S&P 500 index is selected as a benchmark. The model used for optimization is the Robust Second Order Stochastic Dominance (RSSD) model.
- **Strategy 5:** Dynamic investments are made solely into the S&P 500 index.
- **Strategy 6:** Dynamic investments are made solely into an SSD portfolio created at the beginning with the S&P 500 index as the benchmark.
- **Strategy 7:** Dynamic investments are made solely into an RSSD portfolio created at the beginning with the S&P 500 index as the benchmark.

These strategies are designed to provide a comprehensive comparison of different approaches to dynamic portfolio optimization. By incorporating various optimization models and benchmarks, the strategies aim to assess the potential benefits of each method in terms of portfolio returns and risk management.

The reason for not including more combinations is that no feasible solutions were found for additional strategies during the simulation process. The methodology proposed in this paper ensures fair conditions for comparison, taking into account both the selected optimization models and the frequency of investments.

The primary objective is to demonstrate how the RSSD model can be utilized by investors to make more informed and effective investment decisions. By subjecting these dynamic investing strategies to simulation and comparing their performances against the S&P 500 index benchmark, the potential of the RSSD model in augmenting portfolio returns and enhancing risk management will be evaluated. The outcomes of this research are intended to contribute significant insights for both academic investigation and practical implementation in the field of finance.

4 Simulation results

In this section, the results of the simulation are presented. The simulation is conducted using daily closing prices of stocks from the S&P 500 index and the index itself, as discussed in the previous section. The data spans from the year 2010 to April 2024. A training period of 100 days is always used to train the models. Different frequencies of investments are observed, specifically every 5 (weekly), 10 (biweekly), 20 (monthly), 60 (quarterly), and 250 (yearly) trading days.

The initial budget is set to 1, and dynamic investments are determined by the formula $(\text{frequency}/250) \times 0.5$. This means that half of the initial budget is invested every year, and this investment is split according to each frequency to ensure a valid comparison across different strategies. Interest rates and taxes are excluded from all strategies to maintain consistency.

The objective of these simulations is to assess the performance of various dynamic portfolio optimization strategies under different investment frequencies. By analyzing the results, the potential benefits and drawbacks of each strategy are highlighted, providing valuable insights into their effectiveness in maximizing returns and managing risks in dynamic market conditions.

To evaluate the performance of each strategy, the following Key Performance Indicators (KPIs) are observed:

- **Budget at the end of the simulation:** This KPI measures the total portfolio value at the end of the simulation period, providing a direct comparison of the overall performance of each strategy.
- **Average daily return:** This KPI is calculated as the mean of the daily returns, offering insight into the average performance of the portfolio on a daily basis.
- **Standard deviation (Std) of daily returns:** This KPI measures the volatility of the portfolio, indicating the risk associated with the daily returns.
- **Value at Risk (VaR) at 5% of daily returns:** This KPI estimates the maximum potential loss with a 5% probability on any given day, providing a risk assessment metric.

It is important to note that daily returns are cleaned from observations at times of investments to ensure that the effects of the investment timing do not influence the results. This approach allows for a fair and unbiased comparison of the different strategies based on their performance

	Budget at the end					**Average daily return**				
	Investment frequency					Investment frequency				
	5	10	20	60	250	5	10	20	60	250
S1	39.58*	39.50	39.44	39.27	37.68	0.14*	0.13	0.12	0.12	0.12
S2	53.37’	55.84’	59.00’	63.39*’	48.15’	0.16*’	0.14’	0.14’	0.14’	0.12’
S3	28.08	27.21	30.05	35.06*	32.29	0.13*	0.11	0.11	0.11	0.11
S4	27.97	26.91	30.66	36.33*	32.12	0.13*	0.11	0.11	0.11	0.11
S5	20.73	20.70	20.66	21.20*	19.30	0.11*	0.10	0.10	0.10	0.10
S6	39.57*	39.51	39.39	38.79	36.16	0.14*	0.12	0.12	0.12	0.11
S7	39.96*	39.89	39.77	39.14	36.53	0.14*	0.12	0.12	0.12	0.11
	Std of daily returns					**$VaR_{5\%}$ of daily returns**				
	Investment frequency					Investment frequency				
	5	10	20	60	250	5	10	20	60	250
S1	1.33	1.27	1.24*	1.25	1.48	-1.94	-1.81	-1.80	-1.79*	-1.81
S2	1.88	1.78	1.70	1.68	1.64*	-2.83	-2.75	-2.56	-2.54	-2.19*
S3	1.33	1.30	1.29*	1.31	1.59	-2.02	-1.96	-1.97	-1.86*	-2.05
S4	1.33	1.30	1.28*	1.30	1.59	-1.99	-1.92	-1.94	-1.89*	-2.04
S5	1.19’	1.16’	1.14*’	1.19’	1.49	-1.79’	-1.67’	-1.64*’	-1.65’	-1.64*’
S6	1.37	1.30	1.26*	1.28	1.48’	-1.99	-1.86	-1.85	-1.85	-1.84*
S7	1.38	1.31	1.27*	1.29	1.48’	-2.00	-1.89	-1.86*	-1.86*	-1.86*

Note: * indicates the best value in the row, ’ indicates the best value in the column

Table 1 Simulation results for different investment strategies (S) and different investment frequencies

The results of the simulation indicate several key findings (see Table 1). Strategy 2, which allows for the full reallocation of the budget while using the old portfolio as a benchmark and employing the SSD model, shows the highest returns in terms of both the budget at the end of the simulation and the average daily return. This is likely because this strategy permits frequent restructuring of the portfolio, adapting to market changes effectively. However, it also bears higher risk, as indicated by the higher standard deviation and VaR values.

On the other hand, the safest strategy, as expected, is investing directly into the S&P 500 index (Strategy 5). This strategy yields the lowest standard deviation and VaR, reflecting its lower risk profile.

Interestingly, strategies 2 and 3, as well as strategies 6 and 7, yield similar results. This similarity suggests that there is no significant difference between using SSD and RSSD models in this setup. Moreover, frequent portfolio reallocations using the S&P 500 as a benchmark do not significantly differ from a one-time construction of the (R)SSD model portfolio followed by consistent investments in it. In fact, holding the same portfolio structure appears to result in higher returns in some cases.

The most significant outlier is Strategy 2, which involves restructuring the entire portfolio while using the previous portfolio as a benchmark. This strategy achieves significantly higher returns but also incurs higher risk, making it potentially suitable for risk-loving investors.

Examining the differences in investment frequencies across all strategies, it is observed that weekly investments generally yield the highest returns, while monthly investments tend to offer the lowest risk. This is evident from the lower standard deviation and VaR values in most cases for monthly investments.

Overall, the simulation results provide valuable insights into the performance of various dynamic portfolio optimization strategies, highlighting the trade-offs between return and risk under different investment frequencies.

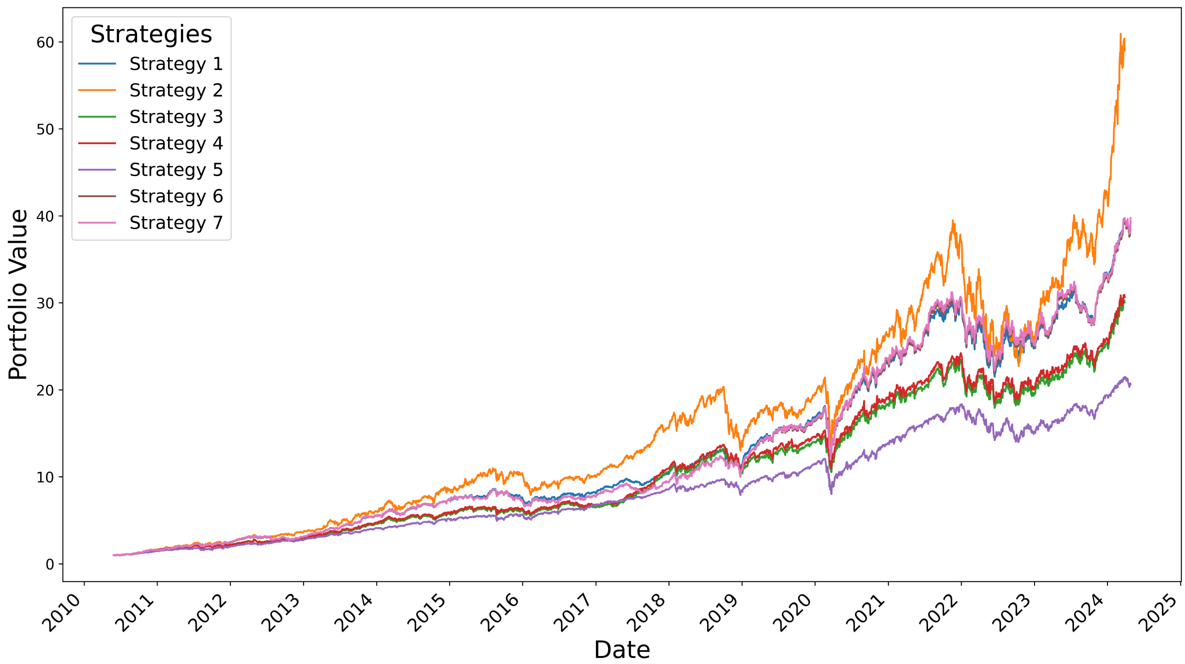


Figure 1 Simulation result for different investment strategies and frequency of 20 days

Figure 1 shows the evolution of portfolio values over time for different investment strategies with a monthly investment frequency, which bears the lowest risk. The following observations can be made from the figure:

- All strategies outperform the strategy of investing solely in the S&P 500 index (Strategy 5), which is depicted by the lowest line.
- Three distinct clusters of similar behavior can be observed:
  - The lowest cluster includes Strategies 3 and 4.
  - The middle cluster includes Strategies 1, 6, and 7.
  - The highest return is achieved by Strategy 2, which stands out significantly from the others.

The similarity in behavior among the strategies can be explained as follows:

- Strategies 3 and 4 both involve reallocating the entire budget while using the S&P 500 index as a benchmark, with Strategy 3 using SSD and Strategy 4 using RSSD for optimization. This results in similar performance.
- Strategies 1, 6, and 7 are grouped together because Strategy 1 allows only buying new assets with the old

portfolio as a benchmark, while Strategies 6 and 7 involve dynamically investing into portfolios created at the beginning using SSD and RSSD with the S&P 500 as a benchmark. Their approach of maintaining or gradually adjusting the portfolio contributes to their similar performance.

- Strategy 2 allows full reallocation of the budget while using the old portfolio as a benchmark, enabling it to adapt more dynamically to market changes and resulting in the highest returns.

This visualization effectively highlights the relative performance and risk profiles of the different investment strategies over time.

5 Conclusion

In this study, various dynamic portfolio optimization strategies were evaluated using daily closing prices of stocks from the S&P 500 index. The strategies differed in terms of the selected optimization model (SSD and RSSD) and the frequency of investments. The simulations, spanning from 2010 to April 2024, revealed that strategies allowing for frequent reallocation of the entire budget while using the old portfolio as a benchmark yielded the highest returns, albeit with higher risk. Conversely, the safest strategy involved investing directly into the S&P 500 index, reflecting a lower risk profile.

It was observed that there is no significant difference between the SSD and RSSD models in this setup. Moreover, strategies that involved dynamic investments into portfolios created at the beginning (using the S&P 500 as a benchmark) showed similar performance, indicating that maintaining or gradually adjusting the portfolio could be effective.

The results highlight the trade-offs between return and risk, with weekly investments generally yielding higher returns and monthly investments offering lower risk. These findings provide valuable insights for investors looking to optimize their portfolios in dynamic market conditions, demonstrating the practical benefits of (R)SSD models and regular portfolio reevaluation.

Acknowledgements

The research project was supported by Grant No. IGA F4/18/2024 of the Internal Grant Agency, Faculty of Informatics and Statistics, Prague University of Economics and Business.

References

Dentcheva D, Ruszczyński A (2006) Portfolio optimization with stochastic dominance constraints. Journal of Banking Finance 30(2):433–451. https://doi.org/https://doi.org/10.1016/j.jbankfin.2005.04.024, https://www.sciencedirect.com/science/article/pii/S0378426605000725, risk Management and Optimization in Finance

Hadar J, Russell WR (1969) Rules for ordering uncertain prospects. The American Economic Review 59(1):25–34. http://www.jstor.org/stable/1811090

Markowitz H (1952) Portfolio selection. The Journal of Finance 7(1):77–91. http://www.jstor.org/stable/2975974

Sehgal R, Mehra A (2020) Robust portfolio optimization with second order stochastic dominance constraints. Computers Industrial Engineering 144:106396. https://doi.org/https://doi.org/10.1016/j.cie.2020.106396, https://www.sciencedirect.com/science/article/pii/S0360835220301303

Level of efficiency of the energy industry in EU countries

Ondřej Novák[1], Veronika Blašková[2]

Abstract. This article compares the efficiency of companies in the energy sector in EU countries. The efficiency calculation is performed through data envelopment analysis method. The dataset includes accounting data (annual frequency) for a total of 3893 companies. The output variables represent the turnover of the company and also the net income of the company. The inputs include variables representing the capital factor and the labor factor.
Empirical results show that the efficiency of this sector in EU countries is at a relatively high level (the average is around 75%). Countries such as Italy and Bulgaria have a significant share of efficient companies. Denmark, on the other hand, has the least. According to the average (and median values), companies in Estonia and Denmark performed very well. Both of these countries have a high share of renewable energy, which may have had a positive impact on their performance.

Keywords: accounting data, company, data envelopment analysis, efficiency

JEL Classification: C44, D24, Q40
AMS Classification: 90B50, 90C08

1 Introduction

Energy is an area that connects all EU countries. The first seed was the European Coal and Steel Community, which included six countries. The interconnection in coal and steel production was not only to prevent another war conflict, but also to ensure a certain solidarity in production. This created a common market where goods could move freely (without taxes) and where countries were forbidden to introduce unfair competitive or discriminatory practices. All these changes were intended to lead to a more efficient market that will benefit all countries involved.

Today, all 27 EU countries are governed by a common energy policy that focuses mainly on security, sustainability and integration. The main topic discussed is the use of renewable and non-renewable energy sources. The EU's long-term target is to have a 45% share of renewables. The use of non-renewable resources poses two problems. The first is that sooner or later we will run out of these resources. The second problem is the production of carbon dioxide, which is released when fossil fuels are burned, leading to atmospheric pollution and other associated problems. That's why the EU is trying to incentivize emission reductions and boost renewable energy.

Although all Member States have the same goal, research to date has shown that the level of the energy sector varies across the EU. For example, Varvařovská and Staňková (2021) examined the productivity of selected EU countries and demonstrated different performance across countries. According to their research, the differences are mainly due to the type of energy sources. In particular, the use of renewable energy sources has a negative impact on the total product of the energy production sector. However, the conclusions of Varvařovská and Staňková (2021) are based only on production analyses. Alternatively, instead of comparing the estimated production coefficients, it is possible to use a different approach and focus on a comparative efficiency assessment approach.

Yang and Li (2017) or Zhou et al. (2013) used a non-parametric data envelopment analysis (DEA) method to evaluate energy efficiency in China. However, such large-scale studies are still lacking for the EU area. Focusing on the EU area, studies such as Shahnazi and Shabani (2020) or Wan et al. (2015) can be mentioned, but they used a parametric approach and mainly assessed productivity. In contrast to conventional analyses where the production function is parameterized, in the case of the DEA method it is not necessary to introduce assumptions about the functional form of the production frontier. The companies (generally speaking, decision-making units) that best transform inputs into outputs form the frontier of efficiency. This group of companies is referred to by the model as efficient units. Other (inefficient) companies lie in the set of admissible solutions. Hollingsworth (2003) referred to the DEA method as the most widely used method for evaluating efficiency regardless of the industry chosen.

[1] Mendel University in Brno, Department of Statistics and Operation Analysis, Zemědělská 1, 613 00 Brno, Czech Republic, xnovak18@node.mendelu.cz.
[2] Mendel University in Brno, Department of Statistics and Operation Analysis, Zemědělská 1, 613 00 Brno, Czech Republic, veronika.blaskova@mendelu.cz.

Recent studies using the DEA method to assess the efficiency of energy-related sectors focusing on the EU region include Bigerna et al. (2019), Mardani et al. (2017) and Simeonovski et al. (2021).

The main aim of this article is to compare the efficiency of the energy sector in EU countries. Unlike the Bigerna et al. (2019), Mardani et al. (2017) and Simeonovski et al. (2021) studies, in this article we focus on the accounting data of individual companies. This gives us a unique micro view of the sector. Instead of using aggregated data, the efficiency calculation is performed at the individual company level and then aggregated by company geographical affiliation. This is the procedure proposed by Varvařovská and Staňková (2021) with the difference that instead of the production analysis, the efficiency of the company and thus the efficiency of the country will be calculated.

2 Material and Methods

The required accounting data (in annual frequency) was obtained from the Orbis database. Within this database, companies falling under the category Electricity, gas, steam and air conditioning supply (i.e., NACE code 35) were filtered. Given the availability of data, two input and two output variables were selected, see Table 1.

Variable	Type	Unit	Other studies using this variable
Turnover	Output	ths EUR	Krejčí and Staňková (2022), Křetínská and Staňková (2021)
Net income	Output	ths EUR	Adnan et al. (2021), Belkhaoui et al. (2020), Wu et al. (2016)
Total assets	Input	ths EUR	Aguilar and Portilla (2024), Staňková et al. (2022)
Costs of employes	Input	ths EUR	Varvařovská and Staňková (2021), Staňková and Hampel (2019)

Table 1 Summary of variables used, including references to studies that used the variable

The output of the company represents the turnover and net income of the company. This combination of variables allows to consider not only the total volume of activities (in financial expressions) but also the different levels of taxation in each country. The input variables represent both labor and capital factors. Capital is expressed here in its physical form through the accounting item total assets of the company. In the case of the labor factor, we decided to take into account the different wage level and instead of the commonly used variable number of employees we used a cost concept in the form of the cost of employees.

For the selected period (2018 and 2019), it was possible to obtain accounting data for a total of 3,910 companies from 24 EU countries. Data could not be obtained for companies from Cyprus, Malta and the Netherlands. For the sake of the representativeness of the results, we decided to include in the analysis only countries where we have more than two dozen representatives. Because of this rule, we have also excluded Greece, Luxembourg, Lithuania and Ireland. The efficiency assessment therefore covers 20 countries based on data from 3,893 companies. See Table 2 for the basic characteristics of the dataset.

Variable	Year	Min	Median	Average	Max
Turnover	2019	1	1,162	208,424	80,134,000
Net income		1	118	14,409	8,498,000
Total assets		1	2,858	430,417	303,284,000
Costs of employes		1	58	15,921	13,516,000
Turnover	2018	2	1,103	201,433	76,016,000
Net income		1	114	11,902	4,978,000
Total assets		1	2,840	409,339	283,169,000
Costs of employes		3	50	14,923	13,364,000

Table 2 Descriptive statistics of the dataset (all variables expressed in thousands of EUR)

Given that there are companies of different sizes in the dataset (as evidenced by the values in Table 2), we have chosen the BCC model to evaluate efficiency, which allows us to work with variable returns to scale. The BCC input-oriented model compiled for unit *H*, which is one of the *p* units, is as follows:

$$\max \quad E_H = \sum_{j=1}^{n} v_{jH} y_{jH} + \mu,$$

$$\text{subject to} \quad \sum_{i=1}^{m} u_{iH} x_{iH} = 1,$$

$$\sum_{j=1}^{n} v_{jH} y_{jk} + \mu \leq \sum_{i=1}^{m} u_{iH} x_{ik}, \forall k = 1,2,\dots,p,$$

$$u_{iH} \geq \varepsilon, \forall i = 1,2,\dots,m,$$

$$v_{jH} \geq \varepsilon, \forall j = 1,2,\dots,n,$$

$$\mu \text{ free},$$

where input variable is arranged in matrix $X = \{x_{ik}, i = 1, 2, \dots, m, j = 1, 2, \dots, p\}$. And output variable is arranged in matrix $Y = \{y_{ik}, i = 1, 2, \dots, n, j = 1, 2, \dots, p\}$. The ε is the so-called infinitesimal constant, and μ represents the magnitude of the deviation from constant returns to scale. The calculations were performed in in DEA SolverPro.

3 Results

According to the results of the BCC model, the efficiency of companies in the energy sector is at a relatively high level. The average efficiency is around 75% in both years. The median efficiency is equal to 100%. In 2018, a total of 2,567 companies were identified as 100% efficient by the model. In 2019, 2,543 companies. A detailed overview of the 100% efficient companies grouped by country is recorded in Table 3. Table 3 not only reports the share of efficient companies in each country with respect to the total number of efficient companies in a given year, but also the conversion to the share of efficient companies with respect to the number of companies in a particular country, given the different representation of country representatives in the dataset.

	% of efficient companies		**% of efficient companies based on the number of companies in the dataset**	
Country	**2019**	**2018**	**2019**	**2018**
Austria	0.20	0.23	17.24	20.69
Belgium	0.31	0.31	17.78	17.78
Bulgaria	15.38	15.35	89.07	89.75
Croatia	4.29	4.29	87.20	88.00
Czech Republic	4.56	4.48	76.32	75.66
Denmark	0.12	0.19	8.11	13.51
Estonia	0.55	0.55	73.68	73.68
Finland	1.97	1.99	25.64	26.15
France	11.36	11.34	81.41	81.97
Germany	0.59	0.58	6.64	6.64
Hungary	5.94	5.88	84.83	84.83
Italy	25.80	25.67	68.62	68.93
Latvia	0.31	0.31	66.67	66.67
Poland	7.39	7.40	81.74	82.61
Portugal	0.28	0.39	26.92	38.46
Romania	8.45	8.41	85.32	85.71
Slovakia	2.44	2.38	72.94	71.76
Slovenia	4.40	4.36	83.58	83.58
Spain	1.06	1.32	18.62	23.45
Sweden	4.60	4.56	46.25	46.25

Table 3 % of efficient companies in a country in relation to the total number of efficient companies and % of efficient companies in a country in relation to the number of companies in the dataset

It can be concluded that the efficiency frontier is mainly formed by companies from Italy, Bulgaria and France. Other countries have a share of efficient companies of less than 10%. However, if we take into account the number of representatives in each country in the dataset, a large share of efficient companies is for example also in Croatia and Slovenia.

Table 4 below shows the average and median values for each country, as well as the derived ranking from these two characteristics. Since in several countries more than half of the companies have been identified as efficient by the model, the median value of one puts these countries in the same position. In the case of average values (although it can be assumed that the data do not have a normal distribution), a more detailed ranking can be constructed.

	2019				**2018**			
Country	**Mean**	**Rank**	**Median**	**Rank**	**Mean**	**Rank**	**Median**	**Rank**
Austria	0.4699	14	0.3835	14	0.5052	14	0.3517	14
Belgium	0.2846	18	0.1396	18	0.2768	19	0.1444	17
Bulgaria	0.9061	4	1.0000	1-12	0.9118	4	1.0000	1-12
Croatia	0.6877	12	1.0000	1-12	0.6903	12	1.0000	1-12
Czech Republic	0.8833	5	1.0000	1-12	0.8850	5	1.0000	1-12
Denmark	0.9242	2	1.0000	1-12	0.9232	2	1.0000	1-12
Estonia	0.9340	1	1.0000	1-12	0.9379	1	1.0000	1-12
Finland	0.3375	17	0.1522	17	0.3274	17	0.1245	18
France	0.8655	7	1.0000	1-12	0.8701	6	1.0000	1-12
Germany	0.1927	20	0.0932	19	0.1910	20	0.0908	19
Hungary	0.8693	6	1.0000	1-12	0.8661	7	1.0000	1-12
Italy	0.7991	10	1.0000	1-12	0.8053	9	1.0000	1-12
Latvia	0.7766	11	1.0000	1-12	0.7782	11	1.0000	1-12
Poland	0.5729	13	0.5323	13	0.5654	13	0.5213	13
Portugal	0.2452	19	0.0741	20	0.2911	18	0.0718	20
Romania	0.8020	9	1.0000	1-12	0.7990	10	1.0000	1-12
Slovakia	0.8207	8	1.0000	1-12	0.8311	8	1.0000	1-12
Slovenia	0.9100	3	1.0000	1-12	0.9201	3	1.0000	1-12
Spain	0.4668	15	0.3625	15	0.4865	15	0.3350	15
Sweden	0.4125	16	0.1985	16	0.4094	16	0.1945	16

Table 4 Average and median values for each country, including derived rankings

If the average efficiency is considered, the first three places are taken by Estonia, Denmark, and Slovenia. These positions are the same for both 2019 and 2018. At the other end of the rankings are Germany, Portugal, and Belgium. These countries also rank last in terms of median values. However, the specific position in the ranking varies from year to year (and also according to the selected indicator). In general, however, the changes in the ranking are not dramatic.

4 Discussion

At present, we know little about the factors influencing the efficiency of the energy sector. According Simeonovski et al. (2021), who examined EU countries based on aggregate data, they report the length of EU membership as a factor influencing efficiency. At the micro-level, however, this link does not appear to be relevant to a country's level of efficiency. For example, German companies perform noticeably worse in terms of efficiency compared to, for example, French companies (both countries can be described as founding members).

Zhou et al. (2013), who examined the efficiency of the power sector in different provinces in China, identified the share of electricity generated in coal-fired power plants as the main factor influencing efficiency there. A similar conclusion was reached by Bigerna et al. (2019) when analyzing the total factor productivity of energy producing companies in EU countries. This factor also seems to be relevant in the case of our results. In terms of average values, Denmark and Estonia took the top two places. Both of these countries are known for their positive thinking on the use of renewable energy. According to Bigerna et al. (2019), countries with a higher share of renewables should have a higher total productivity factor and, according to our empirical results, a higher efficiency value. Unfortunately, the study by Bigerna et al. (2019) constructed Cobb-Douglas production functions for only five selected countries, so it is not possible to make a direct comparison between their results on total factor productivity and our efficiency results.

In the case of analyses at the individual company level, future research may add a further unique perspective and consider, for example, the concentration of companies in a sector or the size of the company. It is possible that the structure of companies in the market could also be an important factor in assessing a country's efficiency. This factor has been shown to be significant, for example, in the automotive sector, see Kovárník and Staňková (2023). In this article, we have tried to account for different company sizes at least through variable returns to scale. However, it would be possible to calculate efficiency separately by company size or to check whether company size has a statistically significant effect on the level of efficiency (for example, by using an accompanying regression analysis).

5 Conclusion

This article deals with a quantitative assessment of the efficiency of the energy sector in EU countries between 2018 and 2019. Unlike most studies that include international comparisons, this one did not use aggregated data for individual countries, but instead used directly the accounting data of individual companies falling under NACE code 35. With information from 3,893 companies from 24 EU countries, it was possible to compile a ranking of EU countries.

In general, companies in the energy sector perform well, as the average efficiency is around 75%. The efficiency frontier is mainly formed by companies from Italy, Bulgaria and France. If we focus on the efficiency of the whole country (derived from the average and median value of all companies in a given country), the best positions are occupied by Estonia, Denmark and Slovenia. All the countries mentioned have more than half of the companies 100% efficient and the average efficiency of the whole country is above 90%.

Acknowledgements

This article was supported by grant No. PEF-TP-23-001 of the Grant Agency IGA PEF MENDELU.

References

Adnan K M M, Sarker S A, Tama R A Z, Pooja P (2021) Profit efficiency and influencing factors for the inefficiency of maize production in Bangladesh. Journal of Agriculture and Food Research 5:100161. https://doi.org/10.1016/j.jafr.2021.100161

Aguilar G, Portilla J (2024) Market power, social welfare, and efficiency in the Peruvian microfinance. Economia Politica. https://doi.org/10.1007/s40888-023-00321-y

Belkhaoui S, Alsagr N, van Hemmen S F (2020) Financing modes, risk, efficiency and profitability in Islamic banks: Modeling for the GCC countries. Cogent Economics & Finance 8(1):1-25. https://doi.org/10.1080/23322039.2020.1750258

Bigerna S, D ’Errico M C, Polinori P (2019) Environmental and Energy Efficiency of EU Electricity Industry: An Almost Spatial Two Stages DEA Approach. The Energy Journal 40(1):29–54. https://doi.org/10.5547/01956574.40.1.sbig

Hollingsworth B (2003) Non-parametric and parametric applications measuring efficiency in health care. Health Care Management Science 6(4):203–218. https://doi.org/10.1023/A:1026255523228

Kovárník R, Staňková M (2023) Efficiency of the Automotive Industry in the Visegrad Group. LOGI – Scientific Journal on Transport and Logistics 1(1):12–23. https://doi.org/10.2478/logi-2023-0002

Krejčí M, Staňková M (2022) The Position of the Czech Republic within the Metallurgical Sector. Mathematical Methods in Economics 2022: Conference Proceedings. Jihlava: Vysoká škola polytechnická Jihlava, pp. 193–198.

Křetínská M, Staňková M (2021) Evaluation of the Construction Sector: a Data Envelopment Analysis Approach. Mathematical Methods in Economics 2021: Conference Proceedings. Praha: Česká zemědělská univerzita v Praze, pp. 287–292

Mardani A, Zavadskas E K, Streimikiene D, Jusoh A, Khoshnoudi M (2017) A comprehensive review of data envelopment analysis (DEA) approach in energy efficiency. Renewable and Sustainable Energy Reviews 70:1298–1322. https://doi.org/10.1016/j.rser.2016.12.030

Shahnazi R, Shabani Z D (2020) Do Renewable Energy Production Spillovers Matter in the EU?. Renewable Energy 150(6):786–796. https://doi.org/10.1016/j.renene.2019.12.123

Simeonovski K, Kaftandzieva T, Brock G (2021) Energy Efficiency Management across EU Countries: A DEA Approach. Energies 14(9):2619. https://doi.org/10.3390/en14092619

Staňková M, Hampel D (2019) Efficiency of the building project sector in the Czech Republic – Stochastic Frontier Analysis approach. In: International Conference of Numerical Analysis and Applied Mathematics (ICNAAM 2018). Melville, American Institute of Physics (AIP), 2116(1):400007

Staňková M, Hampel D, Janová J. (2022). Micro-data efficiency evaluation of forest companies: The case of Central Europe. Croatian Journal of Forest Engineering, 43(2):441–456. https://doi.org/10.5552/crojfe.2022.1541

Varvařovská V, Staňková M (2021) Does the Involvement of "Green Energy" Increase the Productivity of Companies in the Production of the Electricity Sector?. European journal of business science and technology 7(2):152–164. https://doi.org/10.11118/ejobsat.2021.012

Wan J, Baylis K, Mulder P (2015) Trade Facilitated Technology Spillovers in Energy Productivity Convergence Processes Across EU Countries. Energy Economics 48(2):253–264. https://doi.org/10.1016/j.eneco.2014.12.014

Wu Y E, Ting I W K, Lu W M, Nourani M, Kweh Q L (2016) The impact of earnings management on the performance of ASEAN banks. Economic Modelling 53:156–165. https://doi.org/10.1016/j.econmod.2015.11.023

Yang W, Li L (2017) Analysis of total factor efficiency of water resource and energy in China: A study based on DEA-SBM model. Sustainability 9(8):1–21. https://doi.org/10.3390/su9081316

Zhou Y, Xing X, Fang K, Liang D, Xu, C (2013) Environmental efficiency analysis of power industry in China based on an entropy SBM model. Energy Policy 57(6):68–75. https://doi.org/10.1016/j.enpol.2012.09.060

Level of efficiency of the tertiary education sector in the EU

Ondřej Novák[1], Michaela Staňková[2]

Abstract. Despite the common legislative framework of the European Union, there are differences in the education system among its members. Given that education is funded from public budgets, it should be monitored whether these funds are used efficiently. Using a non-parametric method of data envelopment analysis, we evaluate the tertiary education systems of EU countries based on the number of graduates with respect to the inputs used. In addition to the staff, we have also included public resources (namely R&D expenditure and public expenditure on tertiary education) as inputs. Empirical results show that some developed countries (such as Germany and Austria) do not have an efficient education system. Ireland, on the other hand, was among the top countries due to the high number of graduates relative to the inputs used.

Keywords: data envelopment analysis, education, EU countries, efficiency

JEL Classification: C44, H52, I21
AMS Classification: 90B50, 90C08

1 Introduction

The education sector has a specific position compared to other economic sectors (such as manufacturing or construction). The quality of services provided in the education sector directly affects the quality of work in downstream sectors. This is because the output of the education sector is people who are subsequently employed in jobs in other sectors (or do work directly in the education sector). Evaluating the efficiency of the education system is therefore crucial for all employers in the market as well as for national governments. Another reason for the need to control efficiency is the very fact that the educational process is publicly funded. According to Glushak et al. (2015), for the proper functioning of the education sector it is necessary to increase funding not only from public sources, but also from entrepreneurs themselves.

The education sector has undergone dramatic changes in recent years. Among the current challenges, we can undoubtedly mention the Covid-19 pandemic, which caused forced changes in the education system for which some entities were not prepared (Munastiwi and Puryono, 2021). Teachers had to switch to online teaching practically from one day to the next, which was a problem for many schools/universities given the lack of technical facilities (Francom et al. 2021). However, even before 2020, the education system faced some challenges. Some problems are the same across all levels of education. Examples of these are the lack of funding or the reluctance of pupils/students to learn (Šmejkalová, 2016). For the efficiency and quality of primary education, see Demitriadou et al. (2020) and Maazzen et al. (2020) studies. In contrast, Thierry and Emmanuel (2023) and Staňková and Stojanová (2021) studies focus on secondary education. In the context of this article, the focus will be on tertiary education.

In the case of tertiary education, funding gaps can be addressed through grants, which are typically tied to the construction of new and modern curricula or directly to the results of primary or applied research. However, obtaining such grants is not an easy task, which requires having enough quality teaching and research staff at the university. In tertiary education, in addition to funding problems and reduced student motivation, we can also speak of a certain inequality between disciplines. According to Šnýdrová et al (2017), Stojanová and Blašková (2014a and 2014b)] interest in some degree programs is disproportionately high compared to the number of positions available on the labor market. There is an increasing number of cases where even a person with a university degree has problems finding a job on the market. According to Úlovec (2014), this condition is compounded by the absence of good work habits. Students are not sufficiently motivated to high performance during their studies, which translates to low motivation even when pursuing their profession. Due to the characteristics of Generation Y and Z (i.e., those who have entered the workforce in recent years), too high an expectation of financial reward from graduates is also a problem. Many graduates from these generations think that a focus on natural or technical

[1] Mendel University in Brno, Department of Statistics and Operation Analysis, Zemědělská 1, 613 00 Brno, Czech Republic, xnovak18@node.mendelu.cz.
[2] Mendel University in Brno, Department of Statistics and Operation Analysis, Zemědělská 1, 613 00 Brno, Czech Republic, michaela.stankova@mendelu.cz.

sciences will get them a job in the market, see Stojanová et al. (2015). However, employers are always looking for the best quality workforce, both in terms of hard and soft skills. According to the Blašková and Staňková (2023) results, the emphasis on the quality of students in terms of their future employability is key to the efficient functioning of the education sector.

The main aim of this article is to assess how efficiently public resources (specifically R&D expenditure and public expenditure on tertiary education) are used in tertiary education in each EU country. With regard to the chosen period, a sub-objective is also to evaluate the efficiency of tertiary education during the Covid-19 pandemic.

2 Data and Methods

We deal with publicly available data from the EUROSTAT database (annual aggregated data for individual EU countries). Our analysis covers the period from 2019 to 2021. This will make it possible to observe the changes associated with the Covid-19 pandemic that began in Europe in early 2020. Our model is built based on four variables, see Table 1. The output variable is the number of graduates (in numbers) in the tertiary education sector. We have included both labour and capital factors in the inputs. The labour factor is represented in the model by the number of academic staff (in thousands). The capital of universities is represented by two variables namely R&D expenditure and public expenditure on tertiary education (both in EUR million). In this article, tertiary education is defined as categories 5 to 8 of the ISCED classification. Unfortunately, the EUROSTAT database only provides information in aggregate for categories 5 to 8, so it is not possible to take into account different education systems.

Variable	Type	Unit	Other studies using this variable
No. of graduates	Output	Numbers	Ma and Li (20221), Demitriadou et al. (2020)
No. of staff	Input		Andersson and Sund (2022), Mašková and Blašková (2021)
R&D expenditure	Input	EUR million	Blašková and Staňková (2023), Dumitrescu et al. (2020)
Public expenditure	Input		Staňková and Stojanová (2022), Mašková and Blašková (2021)

Table 1 Summary of variables used, including references to studies that used the variable

A non-parametric data envelopment analysis (DEA) method is chosen to estimate the efficiency of tertiary education in EU countries. Similarly, to Staňková and Stojanová (2022) or Blašková and Staňková (2023), we chose a radial measurement approach to calculate (in)efficiency. Given the aggregate data, we chose a model operating with constant returns to scale. The assessment of the EU countries is therefore derived from the CCR input-oriented model, which can be written as follows:

$$
\begin{aligned}
\text{max} \qquad & E_H = \sum_{j=1}^{n} v_{jH} y_{jH}, \\
\text{subject to} \qquad & \sum_{i=1}^{m} u_{iH} x_{iH} = 1, \\
& -\sum_{i=1}^{m} u_{iH} x_{ik} + \sum_{j=1}^{n} v_{jH} y_{jk} \leq 0, \forall k = 1,2,\dots,p, \\
& u_{iH} \geq \varepsilon, \forall i = 1,2,\dots,m, \\
& v_{jH} \geq \varepsilon, \forall j = 1,2,\dots,n,
\end{aligned}
$$

where input variable is arranged in matrix $X = \{x_{ik}, i = 1, 2, \dots, m, j = 1, 2, \dots, p\}$. And output variable is arranged in matrix $Y = \{y_{ik}, i = 1, 2, \dots, n, j = 1, 2, \dots, p\}$. The efficiency E is calculated for country H, which is one of the p rated countries. The ε is the so-called infinitesimal constant.

In order to fully assess changes in efficiency over time, we also focused on the calculation of the Malmquist index, which incorporates information on the individual change in efficiency as well as the change in the production frontier. If the result of the indicator is equal to one, there is no change over time. Values greater than one indicate an improvement in the situation and values less than one indicate a deterioration. In order to be able to compute the Malmquist index for two periods, we need to solve four linear programming problems (in our case, four CCR models). The Malmquist index (*MI*) can be defined as the geometric mean of two efficiency ratios (*E*), where one is the efficiency change measures by the period 1 technology and the other is the efficiency change measured by the period 2 technology:

$$MI = \left[\frac{E^1((x_H, y_H)^2)}{E^1((x_H, y_H)^1)} * \frac{E^2((x_H, y_H)^2)}{E^2((x_H, y_H)^1)} \right]^{1/2}.$$

The calculation of the Malmquist index as well as the individual CCR models was performed in DEA SolverPro (version 16f).

3 Results

The resulting CCR model efficiency scores for each country are recorded in Table 2. The efficient ones are Bulgaria, Ireland and Romania, which have scores equal to one in all three periods. Ireland has reached the efficiency frontier mainly due to its low number of academic staff compared to the number of graduates. For example, the ratio of graduates to academic staff for the EU region averages 3.20 in 2019, 3.41 in 2020 and 3.29 in 2021. However, Ireland scores roughly triple on this indicator (9.73 in 2019, 9.71 in 2020 and 10.17 in 2021). In other words, there were 10.17 graduates per teacher in Ireland in 2021, but only 3.29 graduates per teacher on average across the EU.

In contrast, Bulgaria and Romania have achieved efficiency mainly thanks to a small amount of public funding. For example, if we calculate the ratio between graduates and the volume of public expenditure on tertiary education, we get a value of 1334.80 for Bulgaria and 1256.61 for Romania in 2021. In all other EU countries this ratio will not exceed 205. Even in the other years examined, the values of this indicator for Bulgaria and Romania are several times higher than in other EU countries.

The countries with the lowest efficiency scores are Luxembourg, Germany and Austria. These countries, on the other hand, provide a relatively large amount of funding to universities compared to other countries (both in terms of R&D expenditure and public expenditure on tertiary education).

	2019		2020		2021			2019		2020		2021	
Country	**Eff.**	**Rank**	**Eff.**	**Rank**	**Eff.**	**Rank**	**Country**	**Eff.**	**Rank**	**Eff.**	**Rank**	**Eff.**	**Rank**
Austria	0.242	26	0.187	25	0.254	25	Italy	0.696	10	0.584	7	0.723	10
Belgium	0.469	20	0.454	10	0.490	17	Latvia	0.669	13	0.334	19	0.700	12
Bulgaria	1.000	1	1.000	1	1.000	1	Lithuania	0.721	9	0.338	17	0.653	14
Croatia	0.746	6	0.414	13	0.807	9	Luxembourg	0.176	27	0.156	27	0.188	27
Cyprus	0.661	14	0.410	14	0.832	8	Malta	0.472	18	0.308	20	0.489	18
Czechia	0.676	12	0.486	9	0.618	16	Netherlands	0.310	23	0.269	23	0.343	23
Denmark	0.373	22	0.344	16	0.360	22	Poland	1.000	1	0.587	6	0.915	6
Estonia	0.434	21	0.296	21	0.447	20	Portugal	0.584	15	0.338	18	0.707	11
Finland	0.490	17	0.449	11	0.451	19	Romania	1.000	1	1.000	1	1.000	1
France	0.853	5	0.814	5	0.924	5	Slovakia	0.733	7	0.446	12	0.693	13
Germany	0.262	25	0.173	26	0.249	26	Slovenia	0.470	19	0.286	22	0.444	21
Greece	0.729	8	0.550	8	0.931	4	Spain	0.534	16	0.385	15	0.639	15
Hungary	0.694	11	1.000	1	0.843	7	Sweden	0.275	24	0.256	24	0.284	24
Ireland	1.000	1	1.000	1	1.000	1							

Table 2 Efficiency scores (Eff.) for each year, including derived country ranking

In addition, we analyzed in detail the changes in results caused by the Covid-19 pandemic. Table 3 reports the results of the change in efficiency (the Catch-up effect), the change in the production frontier (the Frontier shift) and the overall change in the form of Malmquist index across periods. Only in the case of Ireland, which has values equal to 1 or very close to 1 throughout the period, is there no negative effect of the pandemic. For the other countries, we can see that the "production" of graduates in 2020/2021 has declined as the value of the Frontier shift is less than one. However, if we also take into account changes in efficiency and therefore examine the overall Malmquist index, several countries have improved their situation in the 2020/2021 period. For example, Cyprus, Estonia, Germany, France, etc. have index values greater than one. However, if we want to assess the EU as a whole, the situation has not improved on average, as the average Malmquist index for the whole EU is 0.977.

	Catch-up		Frontier		Malmquist			Catch-up		Frontier		Malmquist	
Country	**19/20**	**20/21**	**19/20**	**20/21**	**19/20**	**20/21**	**Country**	**19/20**	**20/21**	**19/20**	**20/21**	**19/20**	**20/21**
Austria	0.773	1.358	1.368	0.737	1.058	1.000	Italy	0.839	1.238	1.267	0.784	1.064	0.970
Belgium	0.968	1.079	1.129	0.882	1.092	0.952	Latvia	0.500	2.095	1.923	0.446	0.961	0.934
Bulgaria	1.000	1.000	0.987	0.916	0.987	0.916	Lithuania	0.469	1.930	1.896	0.453	0.889	0.873
Croatia	0.555	1.949	1.740	0.477	0.966	0.929	Luxembourg	0.882	1.210	1.160	0.852	1.022	1.032
Cyprus	0.620	2.030	1.614	0.569	1.000	1.155	Malta	0.653	1.585	1.373	0.755	0.896	1.197
Czechia	0.719	1.272	1.372	0.725	0.986	0.923	Netherlands	0.870	1.273	1.202	0.824	1.045	1.049
Denmark	0.923	1.046	1.068	0.937	0.985	0.980	Poland	0.587	1.559	1.502	0.607	0.881	0.946
Estonia	0.680	1.512	1.458	0.687	0.992	1.040	Portugal	0.579	2.089	1.796	0.494	1.040	1.032
Finland	0.916	1.006	1.129	0.865	1.035	0.870	Romania	1.000	1.000	1.154	0.932	1.154	0.932
France	0.954	1.135	1.103	0.900	1.052	1.022	Slovakia	0.608	1.554	1.537	0.657	0.935	1.022
Germany	0.661	1.440	1.381	0.724	0.912	1.043	Slovenia	0.608	1.553	1.486	0.686	0.904	1.065
Greece	0.754	1.692	1.472	0.633	1.110	1.071	Spain	0.720	1.661	1.490	0.641	1.073	1.064
Hungary	1.440	0.843	1.997	0.444	2.876	0.374	Sweden	0.932	1.111	1.123	0.888	1.046	0.986
Ireland	1.000	1.000	1.006	1.007	1.006	1.007							

Table 3 Results of changes in efficiency (Catch-up), production possibilities frontier (Frontier) and Malmquist index (Malmquist) individually for the period 2019/2020 (19/20) and 2020/2021 (20/21)

4 Summary and Discussion

The empirical results show that there are countries that are efficient because of low numbers of academics (Ireland), but also countries that are efficient because of low use of public resources (Bulgaria and Romania). In the case of Ireland, the ratio of graduates to academic staff is roughly three times the EU average. The converse view is that, relative to inputs, Ireland can produce many more graduates. The popularity of students interested in studying in this country can be at least partly explained by its geographical location and also by its politics. Ireland is very close to the UK, otherwise it is practically surrounded by water. Students there therefore have to travel longer distances if they want to take advantage of studying abroad, which can be a barrier for some students. The second fact is that tuition fees in the UK (specifically England, Wales and Northern Ireland) are three times higher, see for example Blašková and Staňková (2023). So, there is a logical assumption that some students from the UK will prefer to study in "cheaper" Ireland. A similar phenomenon has already been documented, for example, in the Czech Republic by Zámková and Blašková (2013).

The efficiency of Ireland has also been demonstrated in the research by Mašková and Blašková (2021), where the efficiency of tertiary education with regard to the (un)employment of graduates was examined between 2016 and 2018. Although the assessment here was conducted on different variables and over a different time period, the results of Mašková and Blašková (2021) are consistent with ours. Other efficient countries according to our analysis are Romania and Bulgaria. Unfortunately, in the study by Mašková and Blašková (2021), Bulgaria was excluded from the analyses due to data unavailability. But in the case of Romania, the results are consistent, as this country is identified as efficient in both studies. According to the cluster results of Blašková and Staňková (2023), these two countries are very similar in terms of their education systems. It is therefore logical that they are also similar in the results of the efficiency analysis.

According to EUROSTAT, Bulgaria and Romania are among the countries with the lowest GDP per capita. Moreover, these countries give lower amounts as a proportion of GDP than is typical in the EU. For example, in 2020, EU countries provided finance to the tertiary sector averaging 5.02% of GDP. However, in the case of Bulgaria it was only 4.5%, in the case of Romania only 3.14% (the lowest in the EU). However, given such an underfunded education sector, these two countries are able to produce a large number of graduates, which, according to the model, makes them efficient. However, the question remains what is the quality of the knowledge of these students. For example, Agasisti and Zoido (2020) sought to capture this factor through the results of PISA tests administered to students aged 15. In the case of tertiary education, we do not yet have a quantified variable that we can include in the model to represent the quality of graduates across countries. According to the low results of the 2018 PISA tests (see European Commission, 2020), the skills of secondary school pupils in Bulgaria and Germany are not high.

Countries such as Luxembourg, Germany and Austria were the worst performers in our country rankings. These countries, on the other hand, are among the advanced and rich economies that can afford to put more money into education, which, however, puts them at the lowest efficiency score in our model. It is possible that different education systems play a role here, which is reflected in the different levels of R&D expenditure and public expenditure on tertiary education.

Last but not least, we analyzed the impact of the Covid-19 pandemic on the education sector. Within the 2019/2020 period, an upward shift of the frontier has been observed for all countries except Bulgaria (on average for the whole EU to around 1.4 times the previous period). In contrast, all countries except Ireland show a decline in this indicator in the 2020/2021 period. This is an average decrease of less than 30 percentage points. According to Tang (2023), however, not all study programmes were affected equally by the pandemic. The declines in graduates were particularly in medicine and dentistry, where clinical rotations were suspended and medical licensing examinations were discontinued. Similarly, for natural science fields (e.g., geophysics or biology) where it was a problem to conduct field research. On the other hand, students of social and economic studies had only possible technical problems with distance learning, see Romana and Plopeanu (2021). However, in light of this change, some countries got closer to the ideal situation (i.e., to the frontier) and their individual efficiency increased, which had a positive impact on the overall value of the Malmquist index. Nevertheless, the question remains what the impact of the pandemic will be on the skills and overall quality of graduates.

5 Conclusion

This article deals with a quantitative assessment of the efficiency of the tertiary education sector in EU countries. The analysis covers the period from 2019 to 2021, making it possible to see the negative impact of the Covid-19 pandemic that hit Europe in the first half of 2020. Ireland was the best performer throughout the monitoring period. This country has efficiency scores of one in all three years. But in addition, it is the only country to maintain positive results even during a pandemic (both an increase in the overall Malmquist index and an increase in the frontier). Countries such as Luxembourg, Germany and Austria ranked worst in our model. These developed countries are able and willing to put more money into education, which puts them at the bottom of the ranking in the model.

Acknowledgements

This article was supported by grant No. PEF-TP-23-001 of the Grant Agency IGA PEF MENDELU.

References

Agasisti T, Zoido P (2019) The efficiency of schools in developing countries, analysed through PISA 2012 data. Socio-Economic Planning Sciences 68:100711. https://doi.org/10.1016/j.seps.2019.05.002

Andersson C, Sund K (2022) Technical Efficiency and Productivity of Higher Education Institutions in the Nordic Countries. International Journal of Public Administration 45(2):107–120. https://doi.org/10.1080/01900692.2020.1868508

Blašková V, Staňková M (2023) Graduate Employability as a Key to the Efficiency of Tertiary Education. Journal on Efficiency and Responsibility in Education and Science 16(4):262–274. https://doi.org/10.7160/eriesj.2023.160401

Demitriadou E, Stavroulia K E, Lanitis A (2020) Comparative evaluation of virtual and augmented reality for teaching mathematics in primary education. Education and Information Technologies 25:381–401. https://doi.org/10.1007/s10639-019-09973-5

European Commission (2020) Directorate-General for Education, Youth, Sport and Culture: Education and training monitor 2020: country analysis, Luxembourg: Publications Office of the European Union. https://dx.doi.org/10.2766/739096

Francom G M, Lee S J, Pinkney H (2021) Technologies, Challenges and Needs of K-12 Teachers in the Transition to Distance Learning during the COVID-19 Pandemic. TechTrends 65:589–601. https://doi.org/10.1007/s11528-021-00625-5

Glushak N, Katkow Y, Glushak O, Katkowa E, Kovaleva N (2015) Contemporary Economic Aspects of Education Quality Management at the University. Procedia – Social and Behavioral Sciences 214:252–260. https://doi.org/10.1016/j.sbspro.2015.11.643

Ma D L, Li X F (2021) Allocation Efficiency of Higher Education Resources in China. International Journal of Emerging Technologies in Learning (iJET) 1611:59–71. https://doi.org/10.3991/ijet.v16i11.23315

Maazzen N, Ashrafian T, Yilmaz Z, Karagüler M E (2020) A multi-criteria approach to affordable energy-efficient retrofit of primary school buildings. Applied Energy 268:115046. https://doi.org/10.1016/j.apenergy.2020.115046

Mašková K, Blašková V (2021) Efficiency of tertiary education in EU countries. In: Mathematical Methods in Economics 2021: Conference Proceedings, Prague, pp. 312–316

Monica R, Plopeanu A-P (2021) The effectiveness of the emergency eLearning during COVID-19 pandemic. The case of higher education in economics in Romania. International Review of Economics Education 37:100218. https://doi.org/10.1016/j.iree.2021.100218

Munastiwi E, Puryono S (2021) Unprepared management decreases education performance in kindergartens during Covid-19 pandemic. Heliyon 7:e07138. https://doi.org/10.1016/j.heliyon.2021.e07138

Staňková B, Stojanová H (2021) Secondary Education Efficiency in Selected Countries. In: Mathematical Methods in Economics 2022: Conference Proceedings. Jihlava: Vysoká škola polytechnická Jihlava, pp. 336–341

Stojanová H, Blašková V (2014a) The role of graduates' field of study and its impact on the transition to working life. Procedia Economics and Finance 12:636–643. https://doi.org/10.1016/S2212-5671(14)00388-8

Stojanová H, Blašková V (2014b) The Significance of the Chosen Field of Study Depending on the Difficulty of Finding a Job. In: Inted2014: 8th International Technology, Education and Development Conference, INTED Proceedings, pp. 4002–4012

Stojanová H, Blašková V, Tomšík P, Tesařová E (2015) Specification and Characteristic of Generation Y in the Sphere of Work Attitude. In: DIEM 2015: Innovation, Leadership & Entrepreneurship, pp. 565–579.

Šmejkalová J (2016) Proč se zabývat kvalitou vysokoškolského vzdělávání? Ekonomické listy 2:42–51.

Šnýdrová M, Šnýdrová I, Vnoučková L (2017) Vnímání příčin uplatnitelnosti absolventů a jejich závislosti a specifika u dílčích skupin. Ekonomické listy 8(2):40–55.

Tang K H D (2023) Impacts of COVID-19 on primary, secondary and tertiary education: a comprehensive review and recommendations for educational practices. Educational Research for Policy and Practice 22:23–61. https://doi.org/10.1007/s10671-022-09319-y

Thierry M A, Emmanuel O N B (2023) Does Financial Development Increase Education Level? Empirical Evidence from Sub-Saharan Africa. Journal of the Knowledge Economy 14:3878–3903. https://doi.org/10.1007/s13132-022-01020-y

Úlovec M (2014) Potřeby zaměstnavatelů a připravenost absolventů škol – komparační analýza. Praha: NÚV.

Zámková M, Blašková V (2013) Comparing the Views on Tuition Fee Introduction of Brno University Students. In: Proceedings of the 10th International Conference – Efficiency and Responsibility in Education, Prague, pp. 671–679

Profit Allocation in a Multi-Echelon Closed-Loop Supply Chain: A Cooperative Game Theoretical Approach

Petr Pokorny[1]

Abstract. This paper provides insight into how surplus profits from cooperation in a closed-loop supply chain (CLSC) can be shared among a manufacturer, a third-party collector, and a retailer so that the players are willing to cooperate. The manufacturer uses both new and recycled materials to make products that are sold by a retailer. A third-party collector collects the end-of-life products and sells them to the manufacturer for remanufacturing. In contrast to the dominant non-cooperative research, we use the cooperative game theory approach to study the stability of the coalition structures. We start with a non-cooperative Stackelberg solution led by the producer and then form all possible coalitions to coordinate the CLSC. By analyzing the core and its stability conditions, we prove that the core is not empty and that the end customer can benefit from better net product prices when coalitions are formed. The fairness of the surplus profit distribution is tested using the Shapley value and is shown to be in the core. The return rate all increases under the cooperative approach, with one exception when the third party collector is not part of any coalition.

Keywords: Closed-Loop Supply Chain, Cooperative, Game Theory, Nash Equilibrium, Core, Shapley

JEL Classification: C70 C71
AMS Classification: 91A06 91A10 91A12

1 Introduction

Environmental sustainability and social responsibility have gained increasing attention in supply chain management, leading to the concept of Closed-Loop Supply Chains (CLSC). In addition to the traditional forward flow of goods, a CLSC also considers the reverse flow of goods, representing the products at the end of their use/life that are returned to be recycled, refurbished or remanufactured. Our three-echelon CLSC consists of a manufacturer (M) that produces new products and sells them through a retailer (R). The manufacturer can save a unit of manufacturing cost by using a recycled material. A third party collector (C) is responsible for collection, processing the returns to sell the recycled material to M. Multi-echelon CLSC are not as widely studied as 2-player games, but they provide an opportunity to apply cooperative game theory with its solution concepts and to compare the advantages with the non-cooperative approach, which is often based on Stackelberg leader-follower models. The coalition strategies are tested for their stability and the incentives of all players not to leave the coalition, which are linked to the distribution of profits according to their individual expectations. This paper provides insight into how the surplus profits generated by cooperation can be distributed in a way that players are willing to cooperate. It also presents other interesting findings on the overall performance of CLSCs under different coalition structures.

1.1 Cooperation in CLSC

The body of literature on cooperative game theory applied to CLSCs is not as large as in the case of non-cooperative firms. Most of the literature focuses on different coalition structures in two- or multi-level CLSCs and the different ways of allocating surplus profits. (Mingming Leng and Parlar 2009; Leng and Parlar 2010) analyze the effect of information sharing strategy in a three-echelon SC and its impact on total cost performance. (Cheng-Tang and Liu 2013) also apply the Shapley value as a means to coordinate the cooperative game based on a characteristic function. (Fiala 2016) analyzes the profit allocation in a sequence of a non-cooperative and cooperative games applied on two-level SC with random price dependent demand. (Zheng et al. 2019) apply the cooperative approach to a supply chain structure with a manufacturer, distributor and retailer, where the distributor is concerned about the fair distribution of profits between it and the retailer. Finally, (Jena and Sarmah 2022) analyze a CLSC with two competing manufacturers and a retailer, and their profits under a coalition scheme, using a modified Shapley value to coordinate the CLSC.

[1] Prague University of Economics and Business, Faculty of Informatics and Statistics, Department of Econometrics, nám. W. Churchilla 1938/4, 130 67 Praha 3, pokornyp@vse.cz.

1.2 Cooperative Game Theory

As per (Gillies 1997; Peleg and Sudhölter 2003), a cooperative game in characteristic function form is an ordered pair (N, v), where $N = \{1,2, \dots, n\}$ is a finite set of players, i.e. $N = \{M, C, R\}$ in our case, and $v: 2N \rightarrow R+$ is a function that assigns a nonnegative real number to every subset $S \subseteq N$, where $v(\emptyset) = 0$. The function $v(S)$ denotes the profit the subset S can achieve as a separate coalition. A coalition consisting of all the players is the grand coalition. We let $v(N)$ denote the profit of the grand coalition which includes all players. If there exists a division of $v(N)$ such that no subset of players can improve its total payoff by departing from the coalition, then that division is fair. The set of all fair pay-off divisions is the core of a game, CO. Formally,

$$CO(v) = \{x \in R_+^n | \sum_{i \in N} x_i = v(N); \sum_{i \in S} x_i \geq v(S); \forall S \subseteq N\}. \quad (1)$$

x_i is called the payoff or the imputation of a player $i \in N$. A game with a nonempty core for all instances is balanced. A characteristic function is called super additive if the following property holds for any two arbitrary coalitions $S, T \subseteq N$.

$$v(S \cup T) \geq v(S) + v(T), \qquad \forall S, T \subseteq N, \qquad S \cap T = \emptyset. \quad (2)$$

The core of a game can be empty. The Shapley value, (Shapley 1952), of a game is a payoff vector $\Phi(v)$ assigned to each player i and calculates an average expectation that each player should receive based on his marginal contribution to all the coalitions he joins.

$$\Phi(v)_i = \sum_{S \subset N} [(n-k)!\,(k-1)!/n!]\ [v(S) - v(S - \{i\})], \quad (3)$$

where k is the number of players in the coalition S without the player i.

2 Model

We consider three players: the manufacturer M, the retailer R and a third party collector C. R buys the products from M and sells them further down the supply chain and actively advertises, while C endeavors to collect the used products and rewards the customer for each return.

2.1 Notation and Assumptions

CLSC parameters and decision variables are as follows

q	Demand quantity for new products
w	Manufacturer's wholesale price
p	Retail price for customers
s	R's sales effort
r	Quantity of returned product acquired by C
e	Acquisition price offered for the returns to customers
b	Buy-back price offered for the returns by M
a	3rd party's acquisition efforts
c_n/c_r	Unit production cost of products produced from new /recycled raw materials
c_a/c_s	Unit acquisition/sales effort cost
p_e	Reference market price point
n_p	Minimum net price for a customer as $n_p = p - e$
τ	Return rate calculated as $\tau = r/q$

π_j^i is a profit function of agent $j, j \in \{M, C, R, SC\}$ in model, $i \in \{C, MS, MR, MC, CR\}$, $\pi_{SC}^i = \pi_M^i + \pi_R^i + \pi_C^i$. The Stackelberg model MS refers to the fully decentralized CLSC with a dominant leader's role of the manufacturer. The model C stands for the centralized control model or the grand coalition. The remaining threesome of coalitions represents the coalition structure model, i.e., the manufacturer with the retailer (MR), the manufacturer and the 3rd party collector (MC), and finally the retailer with the 3rd party collector (CR).

2.2 Demand and Supply functions

We must first define both the demand and supply functions for the forward and reverse channel respectively. For the purpose of comparing the non-cooperative approach with the cooperative one we will use the similar forms as in (Taleizadeh et al. 2017) and (Pokorny 2022). Hence, the demand function is defined as

$$q^i = \alpha - \beta p^i + \gamma e^i + \delta s^i - \theta(p^i - p_e). \quad (4)$$

With parameters $\alpha, \beta, \gamma, \delta, \theta, p_e \geq 0$, where α is the total market size for products made of new and raw materials, β is the consumer price elasticity, γ is the customer's sensitivity to the acquisition price of the collector, δ captures the sensitivity to the sales effort (e.g. advertising), and θ represents the so-called reference price effect. The supply function in the reverse chain takes the following form

$$r^i = r_0 + r_1 e^i + r_2 a^i \quad (5)$$

where $r_0, r_1, r_2 \geq 0$, and depends on the price incentives, e, and customer's sensitivity to acquisition (collection) efforts, a.

3 Profit functions and Optimization models

In this section we present the profit functions and optimization model for the respective models mentioned above.

3.1 Centralized Model C

If a CLST is one legal entity or controlled by one decision maker, we consider it to be integrated or centralized. The optimization problem for the integrated CLSC is as follows

$$\max_{e,s,p,r} \pi_{SC}^C = (p^C - c_n)q^C + (c_n - c_r - e^C)r^C - 0.5c_s(s^C)^2 - 0.5c_a(a^C)^2, \quad (6)$$

3.2 Decentralized Model MS

Profit functions of *M, R,* and *C* for the M-Led Stackelberg optimization model are as follows

$$\begin{gathered} \max_{w,b} \pi_M^{MS} = (w^{MS} - c_n)q^{MS} + (c_n - c_r - b^{MS})r^{MS} \\ \begin{cases} s.t. \max\limits_{e,a} \pi_C^{MS} = (b^{MS} - e^{MS})r^{MS} - 0.5c_a(a^{MS})^2 \\ \left\{ s.t. \max\limits_{s,p} \pi_R^{MS} = (p^{MS} - w^{MS})q^{MS} - 0.5c_s(s^{MS})^2 \right. \end{cases} \end{gathered} \quad (7)$$

The game is played starting with M making the decision on the wholesale price w for R and the buy-back price b for C. Then C decides on the acquisition price e and collection efforts a and R sets the end consumer price p along with the decision on the sales efforts s.

3.3 Coalitional Model MC

In this model, we follow with the leadership of M when M and C form a coalition. Thus, the margin on buy-back price b is eliminated and the optimization model is

$$\begin{gathered} \max_{w,e,a} \pi_{MC}^{MC} = (w^{MC} - c_n)q^{MS} + (c_n - c_r - e^{MS})r^{MS} - 0.5c_a(a^{MS})^2 \\ \left\{ s.t. \max\limits_{s,p} \pi_R^{MS} = (p^{MS} - w^{MS})q^{MS} - 0.5c_s(s^{MS})^2 \right. \end{gathered} \quad (8)$$

The game is played starting with MC partners making the decision on the wholesale price w for R, on the acquisition price e and collection efforts a and R sets the end consumer price p along with the decision on the sales efforts s.

3.4 Coalitional Model MR

In this model, we continue with M's leadership of M this time teamed up with R. Thus, the wholesale price w is eliminated and the optimization model is

$$\max_{p,b,s} \pi_{MR}^{MR} = (p^{MR} - c_n)q^{MR} + (c_n - c_r - b^{MR})r^{MR} - 0.5c_s(s^{MR})^2 \quad (9)$$

$$\{s.t. \max_{e,a} \pi_C^{MR} = (b^{MR} - e^{MR})r^{MR} - 0.5c_a(a^{MR})^2$$

The game is played starting with MR partners making the decision on buy-back price b for C, end consumer price p, and on the sales efforts s. C decides on the acquisition price e and the collection efforts a.

3.5 Coalitional Model CR

C and R form a CR coalition against the market leader M. This model resembles the MS one reduced to a 2-echelon CLSC case. This model is thus played in two steps as opposed to a three-staged MS model.

$$\max_{w,b} \pi_M^{CR} = (w^{CR} - c_n)q^{CR} + (c_n - c_r - b^{CR})r^{CR}$$
$$\{s.t. \max_{p,e,s,a} \pi_{CR}^{CR} = (p^{CR} - w^{CR})q^{CR} + (b^{CR} - e^{CR})r^{CR} - 0.5c_a(a^{CR})^2 - 0.5c_s(s^{CR})^2 \tag{10}$$

4 Numerical Example

Table 1 shows the input parameters for the numerical simulation.

α	β	γ	δ	θ	p_e	r_0	r_1	r_2	c_s	c_a	c_r	c_n
1100	10	3	6	1	90	0	5	6	500	300	50	80

Table 1 Model parameters

The values are chosen arbitrarily to satisfy the main requirement of the CLSC, i.e. $c_n \geq c_r$, which motivates the manufacturer to use the recycled material and the whole CLSC to collect the used products. We can also observe that the sensitivity of the return function is equally distributed between the acquisition price and the sales effort, assuming that the initial size of the market returns is zero. The reference end-user price is set at 90 units, and we can see in Table 2 that the resulting equilibrium price is not far from it. The demand function is more sensitive to the retail price than to the acquisition price. The sensitivity to the acquisition price in the demand function serves as an inducement for the customer to prefer the eco-product to the disposable one.

4.1 Equilibrium Results

The performance of all the models is summarized in the Table 2.

Variable	$C = MCR$	MS	MC	MR	CR
q^*	192	87	94	176	96
r^*	101	43	88	50	51
w^*	-	95	90	-	94
p^*	91	103	105	96	96
e^*	20	9	18	10	10
b^*	-	16	-	20	15
s^*	0.02	0.02	0.02	0.02	0.01
a^*	2.09	0.95	1.02	1.92	1.05
π_M^*	-	1,897	-	-	2,109
π_R^*	-	689	825	-	-
π_C^*	-	330	-	491	-
π_{MC}^*	-	-	2,638	-	-
π_{MR}^*	-	-	-	3,216	-
π_{CR}^*	-	-	-	-	1,054
π_{SC}^*	4,217	2,916	3,463	3,707	3,163
n_p	70.77	94.50	87.46	86.10	85.92
τ [%]	52.81	50.04	93.74	28.20	52.81

Table 2 Equilibrium results

The key performance measures are the players' profits, the net price to the consumer n_p and the total product return rate τ. We can conclude that all coalitional models generate a higher total CLSC profit π_{SC}^* than the decentralised

MS model. On their own, all players can make higher profits in the coalitional models than in the MS model. The net consumer price n_p is the lowest in model C and again it shows that the customer benefits more from the cooperative coalition setup than from a fully decentralised CLSC structure in MS. Last but not least, the interesting finding is the product return rate τ, which reaches 93.74% when M and C form a coalition and R remains decentralised. It is the result of the double margin elimination when the repurchase price b is eliminated between M and C. τ is also very low when M and R form a coalition. Contrary to the MC model, b is not eliminated and remains an important source of revenue for C, and b is very high, compensating for the low total quantity returned r. On the other hand, in the MR model, the cooperating partners M and R can focus on selling new products, since the wholesale price w is eliminated, favouring the consumer price p and increasing demand q.

4.2 Coalition structure performance and core stability analysis

To analyze the game in a cooperative approach we must construct the characteristic function as per eq. (1). We will scale all pay-offs to the grand coalition (C or MCR) and set $\pi^*_{SC} = 4{,}217$ as 100 %. By doing so, we obtain the following **Table 3** of relative pay-offs $v(S)$ for all the possible coalitions S in lexicographical order.

Coalition S	M	C	R	MC	MR	CR	MCR
$v(S)$	0.450	0.078	0.163	0.626	0.763	0.250	1.000
x^S	0.750	0.078	0.172	0.828	0.922	0.250	0.750
$\Phi(v)_S$	0.591	0.149	0.260	-	-	-	-

Table 3 Characteristic function, the Core and the Shapley Value

Checking the condition from eq.(2), we find out that the game is additive and balanced, thus the core should be non-empty since there exists a solution to the problem from eq.(1) and the respective so-called imputations x^S, $\forall S \subseteq N$, satisfy the condition of $CO(v)$ as per eq.(1). The last row of **Table 3** then shows the Shapley value per eq.(3) which shows the players' expectation of the fair share of profits based on their average marginal contribution to all coalitions. The structure of the core can be illustrated as shown in Figure 1. Each vertex represents one player. The core is represented as the grey shaded area in the triangle using barycentric coordinates and it contains the Shapley value which is not always guaranteed.

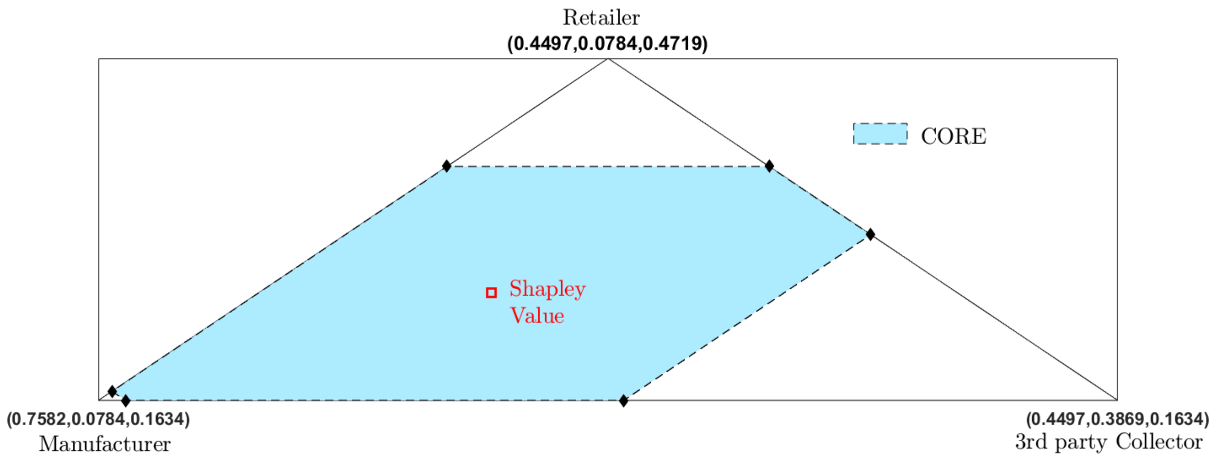


Figure 1 Core of the 3-echelon CLSC

Looking closely at the core in Figure 1, it is worth noting that the closer the shaded area is to one of the vertices of the triangle, the greater the potential share of profit that player can achieve. This characteristics proves the leadership and negotiation power of M over R and C. C is the least influential negotiator in this case. The Shapley vector values also skew the expected gain more towards M. It is clear that the structure of the core would change under the different leadership structures.

5 Conclusion

In this paper, we have analyzed the three-echelon CLSC under manufacturer leadership. We started with the non-cooperative Stackelberg solution and demonstrated the performance of a fully decentralized CLSC. We then constructed coalitional structures among the players, which allowed us to analyze the benefits of cooperation. We concluded that the cooperative approach is beneficial for all players from the point of view of the surplus profits generated and also from the point of view of the consumer, since the net price turns out to be better. Moreover, the results proved to be stable because the core of the game is not empty and the expected fair share of the players' profits, as measured by the Shapley value solution concept, is also within the core. Therefore, the paper concludes that forming coalitions and cooperating is beneficial for all CLSC agents. There are, of course, other solution concepts besides the Shapley value that could be tested, such as the Nucleolus or the Equal Satisfaction measure. These could be analyzed in future research.

Acknowledgements

The research project was supported by Grant No. F4/18/2024 of the Internal Grant Agency, Faculty of Informatics and Statistics, Prague University of Economics and Business.

References

Cheng-Tang Z, Liu L (2013) Research on coordination mechanism in three-level green supply chain under non-cooperative game. Applied Mathematical Modelling 37:3369–3379. https://doi.org/10.1016/j.apm.2012.08.006

Fiala P (2016) Profit allocation games in supply chains. Central European Journal of Operations Research: CEJOR 24:267–281. https://10.1007/s10100-015-0423-6

Gillies DB (1997) Solutions to General Non-Zero-Sum Games. In: Dimand MA, Dimand RW (eds) The foundations of game theory. Volume 2. Elgar Reference Collection., pp 562–600

Jena SK, Sarmah S (2022) Price competition and co-operation in a duopoly closed-loop supply chain. International Journal of Production Economics

Leng M, Parlar M (2010) Analytic Solution for the Nucleolus of a Three-Player Cooperative Game. Naval Research Logistics - NAV RES LOG 57:. https://doi.org/10.1002/nav.20429

Mingming Leng, Parlar M (2009) Allocation of Cost Savings in a Three-Level Supply Chain with Demand Information Sharing: A Cooperative-Game Approach: Operations Research. Operations Research 57:200–213. https://doi.org/10.1287/opre.1080.0528

Peleg B, Sudhölter P (2003) Introduction to the theory of cooperative games. Ringgold Inc
Pokorny P (2022) Reverse Channel Competition in a Dual Sustainable Closed-Loop Supply Chain. In: 40th International Conference Mathematical Methods in Economics 2022. College of Polytechnics Jihlava, Tolstého 1556/16, Jihlava, Czech Republic, College of Polytechnics Jihlava, pp 293–299

Shapley LS (1952) A Value for N-Person Games. RAND Corporation. Available at: https://www.rand.org/pubs/papers/P295.html

Taleizadeh A, Moshtagh MS, Moon I (2017) Optimal decisions of price, quality, effort level and return policy in a three-level closed-loop supply chain based on different game theory approaches. European J of In-dustrial Engineering 11:486

Zheng X-X, Liu Z, Li KW, et al (2019) Cooperative game approaches to coordinating a three-echelon closed-loop supply chain with fairness concerns. International Journal of Production Economics 212:92–110. https://doi.org/10.1016/j.ijpe.2019.01.011

Discrete-Time Dynamic Exchange Rate Overshooting Model

Pavel Pražák [1]

Abstract. The paper deals with a discrete-time version of the overshooting model. Insights into short-run currency market behaviour are often provided by the exchange rate overshooting principle, which illustrates how equilibrium levels are initially quickly overshooted before a slower adjustment occurs. The model is formulated as a system of difference equations using lagarithmic transformation. Within this framework, transient exchange rate deviations within discrete time intervals are also analysed. In addition, the model can be used to formulate and discuss a qualitative analysis of the various possible policy options.

Keywords: exchange rate, difference equations, discrete dynamical systems, stationary solution, qualitative analysis, stability

JEL Classification: C61
AMS Classification: 39A60

1 Introduction

The paper (Dornbusch 1976) deals with the principle of exchange rate overshooting in a short period of time in a small open economy. The original model is formulated as a system of ordinary differenntial equations. Here we introduce a model with difference equations. The model combines aspects of monetarist and Keynesian models. An important aspect of the model is the behaviour of prices, where in the short run prices are assumed to be more or less fixed, but in the long run prices are assumed to be flexible. Such price behaviour has an impact on the behaviour of the exchange rate of the economy. If prices are inelastic in the short run, the market for goods and services does not respond to changes in exogenous variables. The change in these exogenous variables is thus absorbed mainly by the change in the exchange rate, which is in fact the price of the foreign exchange market. In the longer term, prices may then adjust to the original change and the exchange rate is again corrected. This shows that the exchange rate can react more dynamically in the short run than a change in the longer run, which is more consistent with the equilibrium level. Thus, the exchange rate overshoots the long-run equilibrium level, which is mainly due to short-term maladjusted prices. The following text and interpretation is mainly based on the notes in (Azariadis 2000; Zhang 2006).

2 The Model

The model is proposed within the context of the IS-LM model for a small open economy with sticky output prices and perfect forecasts. Given that the focus is on short-run phenomena, the level of output can be considered as constant. In economic dynamics, natural logarithms of variables are frequently employed, (de la Fuente 2000); Shone 2003; Prazak 2017). The principal advantage of this logarithmic approach to formulating economic dynamics problems is that certain special non-linear relationships can be expressed as linear relationships. In particular, if X represents a given positive variable, then it is posible to consider its logarithmic tranformation

$$x = \ln X, \quad X > 0. \tag{1}$$

This implies that variables denoted by lowercase letters are the natural logarithms of variables denoted by uppercase letters. This notation is used frequently in the following text, although it is not specified for all particular cases. Given the frequent occurrence of short-run changes in economic dynamics, it is possible to utilise the following approximation

$$\ln(X + 1) \approx X, \quad X \to 0, \tag{2}$$

which allows for a straightforward transition from variables in logarithms to real variables and vice versa.

Here, we shortly introduce the essentials variables, constants and definition that are used, for more details (Blanchard 2000; Dornbusch 1976),

- The dynamics and the model itself is described at discrete **time periods** t, $t \in \mathbb{N}_0$.

[1] University of Hradec Králové, Faculty of Informatics and Management, Department of Informatics and Quantitative Methods, Rokitanského 62, Hradec Králové, Czech Republic, pavel.prazak@uhk.cz

- The **home price level** at period t is denoted as P_t and the **rate of inflation** in this period is
$$\pi_t = \frac{P_t - P_{t-1}}{P_{t-1}}.$$
Using the notation (1) and the aproximation (2) we can observe that
$$\pi_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1 \approx \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln P_t - \ln P_{t-1} = p_t - p_{t-1}. \tag{3}$$
- The **nominal exchange rate** E_t at time period t is also considered, (Fronckova and Prazak 2020). This variable represents the price of a unit of foreign currency in terms of the domestic currency. Thus an increase in E_t means a devaluation of the home currency and vice versa. Then the **real foreignn** exchange rate is the price of foreign goods in terms of home goods. It is given by
$$\epsilon_t = E_t \frac{P^*}{P_t},$$
where P^* is the foreign price level and P_t is the home price level at the time period t. Again, it is usefull to consider natural logarithm of this relation, which can be formulated as follows
$$\ln \epsilon_t = \log E_t + \ln P^* - \ln P_t = e_t + p^* - p_t, \tag{4}$$
where the notation (1) was used.
- The **home nominal interest rate** at time period t is R_t

The description of the model is divided into four parts - relation for goods market, relation for money market, relation for price adjustement and presentation of the concept of uncovered interest rate parity. This description will be followed by a synthesis of the given relations into a system of two nonlinear differential equations.

2.1 Goods market

A small open economy in a short period is considered thus the output supply can be fixed at some constant, exogenous level $Y_t = Y^s$ for all t. Thus, the logarithmic form of this equation is
$$y_t = y^s, \tag{5}$$
where the notation (1) was used. On the other hand the demand for domestic output is proportional to the real exange rate ϵ_t and it is indiretly proportional to ex-ante real interest rate $R_t - \pi_{t+1}$. Thus
$$Y_t^d = \frac{\epsilon_t^\delta}{(1 + R_t - \pi_{t+1})^\sigma}, \quad \delta > 0, \sigma > 0,$$
where δ is elasticity of aggregate demand with respect to the real exange rate and σ is elasticity of aggregate demand with respect to the domestic real interest rate respectively. Using (2) the logarithmic form of the given relation is
$$\ln Y_t^d = \ln \frac{\epsilon_t^\delta}{(1 + R_t + \pi_{t+1})^\sigma} = \delta \ln \epsilon_t - \sigma \ln(1 + R_t - \pi_{t+1}) \approx \delta \ln \epsilon_t - \sigma(R_t - \pi_{t+1}).$$
Using the notation (1) and substituting (2) and (4) into the latest equation we finally have
$$y_t^d = \delta(e_t + p^* - p_t) - \sigma(R_t - p_{t+1} + p_t). \tag{6}$$

2.2 Money market

The real money demand at period t is directly proportional to the real output and it is indirectly proportinal to the nominal interes rate as
$$\frac{M}{P_t} = \frac{Y^\phi}{(1 + R_t)^\lambda}, \quad \phi > 0, \lambda > 0,$$
where M is nominal demand for money at period t, P_t is price level in the economy and R_t is nominal interest rate. Moreover, ϕ can be considered as elasticity of liquidity preference with respect to real income and λ is elasticity of of liquidity preference with respect to nominal rate of interest. Taking the natural logarithm of the given equation allows us to express the equation for money demand as
$$m - p_t = \phi y - \lambda R_t, \tag{7}$$
where the notation (1) was used.

2.3 Perfect foresight and price adjustment

A key feature of the model is that prices and the price level are sticky and adjust slowly. When aggregate demand is above potential output, prices rise. On the other hand, if aggregate demand is below potential output, prices fall. This dynamic can be described by the following relationship

$$1 + \pi_{t+1} = \left(\frac{Y_t^d}{Y}\right)^{\alpha}, \quad \alpha > 0,$$

where α is the elasticity of the home inflation rate with respect to excess aggregate demand. If aggregate demand Y_t^d in period t is greater than real output Y_t, then the expression on the right-hand side is greater than one and the rate of inflation π_{t+1} in period $t+1$ must be positive for the expression on the left-hand side to be greater than one. The above given dynamics can be alternatively described as geometric adjustment, (Kodera 2001). With the help of (2) the logarithmic form of the introduced equation can be considered. If the notation (1) is used and (2), (3) are applied, we get

$$p_{t+1} - p_t = \alpha(y_t^d - y). \tag{8}$$

2.4 Interest rate and uncovered interest rate parity

The concept of uncovered interest rate parity is a financial theory that considers the difference in nominal interest rates between domestic and foreign countries to be equal to the relative changes in the foreign exchange rate over the same time period

$$1 + R_t - R^* = \frac{E_{t+1}}{E_t},$$

where R^* is the foreign nominal interest rate, which is considered to be an exogenous constant value in this context. Once more, the logarithmic form of the given equation can be considered with the help of equation (2). If the notation (1) is employed and the approximation (2) is applied, then after rearrangement, the following equation is obtained

$$R_t = R^* + e_{t+1} - e_t. \tag{9}$$

2.5 The system of difference equations

The system of linear difference equations (6)–(9) describes the dynamics of the model. If the endogenous variables y_t^d and R_t are eliminated the given system can be reduced to the system of two difference equations

$$p_{t+1} - p_t = \frac{\alpha}{1-\alpha\delta}\left(\sigma(e_t + p^* - p_t) - \frac{\delta}{\lambda}(p_t - m + \phi y) - y\right) \tag{10}$$

$$e_{t+1} - e_t = \frac{1}{\lambda}(p_t - m + \phi y) - R^*. \tag{11}$$

Equation (11) was obtained by combination of (9) and (7) and equation (10) was obtained by combination of (6), (8) and (7).

3 Results

First the steady state and the phase diagram for the system (10)-(11) is found and constructed. Then eingenvalues and eigenvectors of a the system are computed.

3.1 Stationary solution

The system (6)–(9) is autnonomous (time invariant), thus a stationary solution (or steady state) can be considered. Having system (10)-(11) its stationary solution is $p_t = p^\circ$ and $e_t = e^\circ$ for all t. This constant solution can be found by the following system of linear algebraic equations

$$0 = \frac{\alpha}{1-\alpha\delta}\left(\sigma(e^\circ + p^* - p^\circ) - \frac{\delta}{\lambda}(p^\circ - m + \phi y) - y\right) \tag{12}$$

$$0 = \frac{1}{\lambda}(p^\circ - m + \phi y) - R^*. \tag{13}$$

Its solution is

$$
\begin{aligned}
p^{\circ} &= \lambda R^* + m - \phi y && (14)\\
e^{\circ} &= p^{\circ} - p^* + \frac{1}{\sigma}(y + \delta R^*) = \left(\lambda + \frac{\delta}{\sigma}\right) R^* + m - p^* + \left(\frac{1}{\sigma} - \phi\right) y. && (15)
\end{aligned}
$$

In the stationary state the relation (8) and (9) can be also applied and we can observe that

$$
\begin{aligned}
y_t^d &= y && (16)\\
R_t &= R^*. && (17)
\end{aligned}
$$

3.2 Qualitative analysis and phase diagram

The model can be analyzed in terms of a phase space diagram for variables price level p_t and exchange rate e_t as is shown in Figure 1. The diagram consists of two nullclines that represent the process $p_{t+1} - p_t = 0$ and the process $e_{t+1} - e_t = 0$ respectively. Their intersection represents the steady state (p°, e°) of the system that is given by (14) and (15). The two nullclines of (6)–(9) correspond to the equations

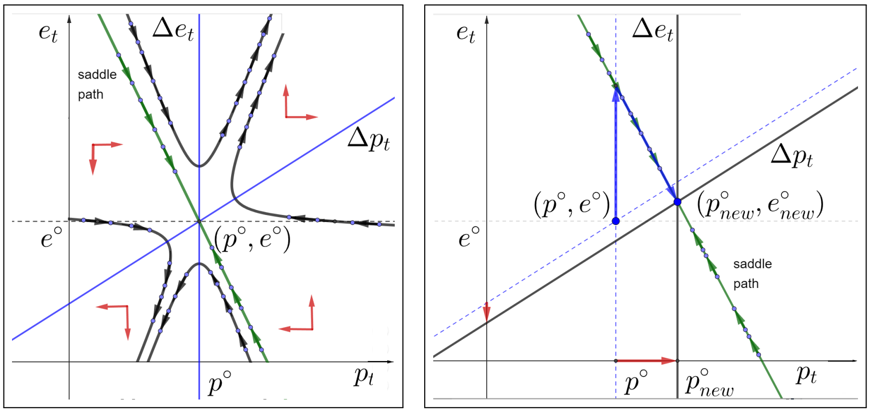


Figure 1 On the left panel there is an illustration of the phase diagram of the system (10)-(11). On the right panel there is an illustration of adjustment trajectory in response to an increse of monney supply.

$$
\begin{aligned}
\text{nullcline } p_{t+1} - p_t : \quad e_t &= \left(1 + \frac{\delta}{\lambda\sigma}\right) p_t + \frac{\delta}{\lambda\sigma}(\phi y - m) + \frac{y}{\sigma} - p^*\\
\text{nullcline } e_{t+1} - e_t : \quad p_t &= m - \phi y + \lambda R^*.
\end{aligned}
$$

Using the second equation it is clear that the nullcline $e_{t+1} - e_t = 0$ is a vertical line representing the constant value $p_t = p^{\circ}$. Furthermore, directly from (11) it is

$$
\frac{\partial(e_{t+1} - e_t)}{\partial p_t} = \frac{1}{\lambda} > 0.
$$

Instead of infinitesimal change of coordinate p_t we can consider the differece $\Delta p_t = p_{t+1} - p^{\circ} > 0$ for some t, then $p_{t+1} > p^{\circ}$ and $\Delta e_t = e_{t+1} - e_t > 0$. This observation means that e_t is an increasing sequence and the exchange rate appreciates. On the other side if $\Delta p_t = p_{t+1} - p^{\circ} > 0$ for some t then $p_t < p^{\circ}$ and $\Delta e_t = e_{t+1} - e_t < 0$. It means that e_t is decreasing sequence and the exchange rate depreciates. The nullcline $p_{t+1} - p_t = 0$ is an increasing function of p_t as its slope is positive and greater than one, in particular the slope is

$$
1 + \frac{\delta}{\lambda\sigma} > 1.
$$

Let us futher consider (10) and then the following partial derivative

$$
\frac{\partial(p_{t+1} - p_t)}{\partial p_t} = -\frac{\alpha(\sigma + \delta/\lambda)}{1 - \alpha\delta}.
$$

Now it is obvious that the sign of this partial derivative is not generally given. However, provided that $1 - \alpha\delta > 0$ the sign of the later partial derivative can be specified. This assumption, which can be reformulated and completed as

$$
0 < \alpha < \frac{1}{\delta}, \tag{18}
$$

can be interpreted as follows. The parameter α represents the speed of price adjustment. Thus, the given assumption implies that prices adjust at a relatively slow pace. In this case we have

$$\frac{\partial(p_{t+1}-p_t)}{\partial p_t} < 0.$$

Now we consider the coordinate p_0 of a point on the nullcline $p_{t+1}-p_t = 0$. If $p_t < p_0$ for some t, then the region above the nullcline is considered. In this case $p_t - p_0 < 0$, which implies that $p_{t+1} - p_t > 0$. This shows that the sequence p_t is increasing in the region considered. A similar argument can be made for the region below the nullcline $p_{t+1} - p_t = 0$, in which case the sequence p_t is decreasing. All these results are shown in figure 1. This analysis suggest that the stationary point is a saddle point. The more detailed analysis is given in the following subsection.

3.3 Eigenvalues and eingenvectors

Using the steady-state solution, the variables representing the steady-state deviations can be introduced and the autonomous system (10)-(10) can be transformed into an autonomous homogeneous system of difference equations in a matrix form

$$\begin{pmatrix} p_{t+1}-p^\circ \\ e_{t+1}-e^\circ \end{pmatrix} = \mathbf{A}\begin{pmatrix} p_t-p^\circ \\ e_t-e^\circ \end{pmatrix} \tag{19}$$

where

$$\mathbf{A} = \begin{pmatrix} 1-\dfrac{\alpha(\sigma+\delta/\lambda)}{1-\alpha\delta} & \dfrac{\alpha\sigma}{1-\alpha\delta} \\ \dfrac{1}{\lambda} & 1 \end{pmatrix}.$$

The behavior of the solution to the system (19) depends on the eigenvalues of matrix **A**, (Zhang 2006; de la Fuente 2000), and the values of eingenvalues depend on values of parameters. For further analysis it is useful to suppose that the price adjustment is sufficiently slow. In particular, it is considered that (18) is valid. The characteristic polynomial, (de la Fuente 2000), of matrix **A** with eigenvalue μ is

$$P(\mu) = (\mu-\mu_1)(\mu-\mu_2) = \mu^2 - \mathrm{tr}\mathbf{A}\,\mu + \det\mathbf{A},$$

where tr**A** is the trace of matrix **A** and det **A** is its determinat. Under the condition (18) it can be observed that

$$\begin{aligned} \mathrm{tr}\mathbf{A} &= \mu_1+\mu_2 = 2-\frac{\alpha(\sigma+\delta/\lambda)}{1-\alpha\delta} < 2 \\ \det\mathbf{A} &= \mu_1\mu_2 = 1-\frac{\alpha(\sigma+\delta/\lambda)}{1-\alpha\delta}-\frac{\alpha\sigma}{\lambda(1-\alpha\delta)} = 1-\frac{\alpha}{1-\alpha\delta}\left(\sigma+\frac{\delta}{\lambda}+\frac{\sigma}{\lambda}\right) < 1, \end{aligned}$$

thus the discriminant D of the characterictic polynomial $P(\mu)$ is

$$D = (\mathrm{tr}\mathbf{A})^2 - 4\det\mathbf{A} = \left(\frac{\alpha}{1-\alpha\delta}\left(\sigma+\frac{\delta}{\lambda}\right)\right)^2 + \frac{4\alpha\sigma}{\lambda(1-\alpha\delta)} > 0.$$

Now it follows that both roots of characteristic polynomial μ_1 and μ_2 are real numbers and their sum is less than 2 and their product is less than 1. From the relation

$$P(1) = (1-\mu_1)(1-\mu_2) = 1-\mathrm{tr}\mathbf{A}+\det\mathbf{A} = -\frac{\alpha\sigma}{\lambda(1-\alpha\delta)} < 0,$$

it can be observed that eigenvalues μ_1 and μ_2 lie on the opposite side of 1 - it means that one of them is greater than 1 and one of them is less than 1. Similarly, one can calculate that

$$P(-1) = (1+\mu_1)(1+\mu_2) = 1+\mathrm{tr}\mathbf{A}+\det\mathbf{A} = 4-\frac{\alpha}{1-\alpha\delta}\left(2\left(\sigma+\frac{\delta}{\lambda}\right)+\frac{\sigma}{\lambda}\right).$$

For the stationary point (p°, e°) to be a saddle point, both eigenvalues should be greater than (-1), (Kelley and Peterson 2001). The additional condition

$$\alpha < \frac{4\lambda}{(2\lambda+1)(\sigma+2\delta)}, \tag{20}$$

which reinforces the assumption that the speed of price adjustment is slow, as stated in (18), implies that

$$P(-1) > 0.$$

This implies that both eigenvalues are either greater than or less than -1. The latter possibility is inconsistent with the observation that one of the eigenvalues is greater than 1. Therefore, it can be concluded that under the assumptions (18) and (20), there are two different real eigenvalues with the following properties

$$|\mu_1| < 1 \text{ and } \mu_2 > 1.$$

Due to the lack of space, the discussion of eigenvectors is not included here.

3.4 General solution to the system

The general solution of the system (10), (11) is of the form (Kelley and Peterson 2001),

$$\begin{aligned} p_t &= p^\circ + C_1 h_1 \mu_1^t + C_2 h_2 \mu_2^t \\ e_t &= e^\circ + C_1 \mu_1^t + C_2 \mu_2^t, \end{aligned}$$

where C_1 and C_2 are real constants whose values must be found by initial conditions and $(h_1, 1)^\top$ and $(h_2, 1)^\top$ are eigenvectors associated with eigenvalues μ_1 and μ_2 respectively.

4 Discussion and Conclusion

There is a special stable saddle path $e_t = e^\circ + h_1(p_t - p^\circ)$ that has negative slope, $h_1 < 0$, (Azariadis 2000). This path can be observed in Figure 1. Now, let us consider that there is a shift in exogenous parameter money supply and istead of m there is $m_{\text{new}} > m$. Then the nullcline $e_{t+1} - e_t = 0$ shifts to the right and the nullcline $p_{t+1} - p_t = 0$ shifts downwards, see the right hand panel in Figure 1. The old steady state (p°, e°) moves to the new steady state $(p^\circ_{\text{new}}, e^\circ_{\text{new}})$ with a new stable saddle path. As the price level cannot immediately respond to the new situation, it initially remains at the old steady-state level. In contrast, the exchange rate can adjust instantly, jumping to a value that places the system on the new saddle path. This value typically exceeds the new steady state level, indicating an overshooting of the exchange rate. It means that the immediate response of the considered economy is an exchange rate depreciation followed by an expected appreciation, see Figure 1. The model can be analysed in more detail, in particular the saddle path should be decribed in more detail. This will be the subject of future work.

Acknowledgements

The support of the 2024 Specific Research Project No. 2105 of the Faculty of Informatics and Management of the University of Hradec Králové is kindly acknowledged. Special thanks to prof. Antonín Slabý and Ing. Daniel Schmid for their valuable suggestions.

References

Azariadis C (2000) Intertemporal Macroeconomics. Blackwell Publishers Ltd, Oxford

Blanchard O (2000) Macroeconomics. Prentice-Hall, Inc., New Jersey

Dornbusch R (1976) Expectations and exchange rate dynamics. Journal of Political Economy 84:1161–1176

Fronckova K, Prazak P (2020) Predicting Exchange Rates Using the Kalman Filter. In: Hradec Economic Days 2020, vol 10, pt 1. University of Hradec Kralove, Hradec Kralove, pp 161–168

de la Fuente A (2000) Mathematical Methods and Models for Economics. Cambridge University Press, Cambridge

Kelley WG, Peterson AC (2001) Difference Equations, An Introduction with Appliocations. Academic Press, New York

Kodera J (2001) Monetary Analysis. Melandrium, Slany

Prazak P (2017) Dynamic linear IS-LM model. In: Hradec Economic Days, vol 7(1). University of Hradec Kralove, Hradec Kralove, pp 715–722

Shone R (2003) An Introduction to Economic Dynamics. Cambridge University Press, New York

Zhang WB (2006) Discrete Dynamical Systems, Bifurcations and Chaos in Economics. Elsevier B.V., Amsterdam

Choking Hazard: Surviving the Heat of Competitive Counter-Strike

Jan Rejthar[1]

Abstract. This article explores choking under pressure among competitive Counter-Strike players and identifies pressure determinants in professional competitions. Using a generalized additive model on complete competitive Counter-Strike history from the hltv.org archive, the study parallels traditional sports psychology, showing players choke under pressure. A key finding is the positive moderating effect of experience on players' susceptibility to choking. This research enhances our understanding of psychological intricacies in esports and offers actionable insights for player development and performance optimization in competitive gaming contexts.

Keywords: Counter-Strike, Choking under pressure, Esports, Economics of sports, Generalized Additive Models

JEL Classification: C20, L83, Z20, M51, M54
AMS Classification: 62P20

1 Introduction

In the last decade, esports has transformed from niche entertainment to a global phenomenon (Meng-Lewis et al., 2022). Defined as professional or semi-professional competitive gaming, esports have gained immense popularity, supporting numerous players and high-stakes tournaments (Nagorsky & Wiemeyer, 2020). The number of active players, spectators, and prize pools continues to rise (Watanabe et al., 2021), with the current generation increasingly favoring esports over traditional sports (Hamari & Sjöblom, 2017).

The rise of high-stakes tournaments and fierce competition in esports has made performance-related issues a focal point of inquiry (Hallmann & Giel, 2018). Despite early skepticism about its legitimacy as a sport, esports is now widely recognized in the scientific literature (Flegr & Schmidt, 2022). Additionally, the extensive data generated by esports has been facilitated across various fields such as machine learning and big data (e.g., Xenopulous et al., 2020), economics (e.g., Parshakov et al., 2022; Rejthar, 2024; Shengjie, 2019), or media studies (Hamari & Sjöblom; Weber, 2009).

While studies have explored various facets of esports (Flegr & Schmidt, 2022), there is a paucity of research on the psychological aspects of player performance, particularly choking under pressure, referring to players failing to perform at critical moments despite extensive training (Hill et al., 2010). The primary objective of this paper is to explore choking under pressure in Counter-Strike (CS).

2 Data and methodology

2.1 Tools for the analysis

All the analysis was done using Python v3.12.1 and R v4.0, utilizing the R econometric packages gratia v0.9.0 (Simpson, 2024) and mgcv v1.9.1 (Wood, 2001), and a Python data manipulation package Polars v0.20.7 (Ritchie Vink et al., 2024).

2.2 Data sources

HLTV

HLTV, founded in 2002, is a crucial platform for professional CS news, analysis, statistics, tournament results, and betting. It publishes a weekly ranking of top CS teams and players, widely recognized as the primary ranking system in the CS professional scene and validated by Rejthar and Kotrba (2023). The exact calculations for HLTV ratings and rankings are proprietary (HLTV.org, 2024). Additionally, HLTV maintains a comprehensive archive

[1] Prague University of Economics and Business, jan.rejthar@vse.cz.

of tournament data, including prize pools, participating teams, dates, locations, match results, and detailed player statistics.

Liquipedia

Liquipedia, established in 2008 by an esports organization, Team Liquid, is a volunteer-driven wiki covering over 40 esports titles. It provides tournament results and is a comprehensive esports wiki supported by dedicated staff and multiple organizations (Winkler, 2016). Additionally, it categorizes tournaments into different tiers based on tournaments' prestige.

HLTV was the primary data source for all matches in all tournaments. Due to differences in databases and naming conventions, Liquipedia was used to assign tournament prestige through a semi-automatic process. Missing data and errors in the HLTV archive were cross-referenced with Liquipedia and the official tournament websites. If data were unavailable, the tournament was discarded. Some smaller, older tournaments had errors or outdated tournament pages.

2.3 Variables and Data Summary

The analysis uses data from all CS: Global Offensive[2] (CS: GO) matches in the HLTV archive from August 25, 2012, to October 22, 2023. After excluding non-competitive or erroneous tournaments and those played in CS 1.6 or CS 2, the dataset includes 57,975 matches from 4,759 tournaments. HLTV rankings, available since October 1, 2015, were used, resulting in a final dataset of 42,672 matches from 3,708 tournaments. Data from the 1,051 tournaments before October 1, 2015, were used to assess player experience on October 1, 2015.

Some variables and their estimations must be explained in depth, starting with Tier and Major. Tier categorizes tournaments into four levels of prestige based on Liquipedia's classifications: S-Tier, A-Tier, B-Tier, and C-Tier. S-Tier tournaments are the most prestigious, featuring outstanding prize pools and the best teams globally, and played on LAN[34]. A-Tier tournaments have large prize pools and numerous top-tier teams. B-Tier tournaments include smaller LAN and larger online events with top-tier teams. C-Tier tournaments are the lowest ranked, usually played online with no top teams (Liquipedia, 2024).

Major is a dichotomous variable indicating whether the tournament is a Major Championship sponsored by Valve, the developer of CS: GO. Majors are highly prestigious, occurring two to three times a year. All Majors are S-Tier events, but not all S-Tier events are Majors.

Subsequently, a variable Stage is defined to assess how deep a match is played in the tournament. Stage is defined in the following way to make the finals the highest stage and ensure consistency amongst different tournament formats. Let T be the set of tournaments in the dataset, and for each $t \in T$, let R_m be the set of all the possible numbers of rounds $r \in R_m$ a team might need to go through starting from match m to win the tournament t. Then, Stage is calculated as

$$\text{Stage}_{mt} = \max_{t \in T}\left(\max_{r \in R_m}(r)\right) - \max_{r \in R_m}(r) \tag{1}$$

Appraising players' performance and form is necessary to assess whether players choke. The performance is assessed via the players' HLTV rating, while their form is assessed by long-term HLTV rating.

Subsequently, to inspect a moderation effect of experience on performance under pressure, experience must be assessed. At match t, the experience of a player i is defined using an indicator function $1(A)$ as

$$\text{Experience}_{it} = \log\left(\sum_{j=1}^{t-1} 1\big(\forall x \in \{\text{Tier,LAN,Major,Stage}\},\ x_{ij} \geq x_{it}\big)\right) \tag{2}$$

where S-Tier $>$ A-Tier $>$ B-Tier $>$ C-Tier, Major $>$ Not Major, and LAN $>$ online.

Lastly, assessing the strength difference between the teams is necessary to account for the heterogeneity in opposition. This article follows Klaassen & Magnus (2003) and estimates the relative skill difference between teams i and j as

[2] Global Offensive is a version of CS that was played as the major professional CS title from 2012 to 2023 replacing CS: Source and CS 1.6, and was gradually replaced by CS: 2 in autumn of 2023.
3 A tournament played at a venue (abbreviation of Local Area Network).
4 This condition was slightly relaxed during the Covid-19 pandemic.

$$\text{Strength\_difference}_{ij} = \log(\text{Rank}_j) - \log(\text{Rank}_i) \quad (3)$$

where Rank_i and Rank_j are the HLTV ranks of the teams, with i being the team in question and j being the opponent. This accounts for the non-linear nature of rank differences.

These variables help assess performance under pressure, hypothesizing that higher prestige, larger prize pools, or how deep in the tournament the match is played increases stress, leading to choking. It is further expected that experience mitigates this effect.

All prizes are adjusted for inflation to August 2012 USD using dat from the U.S. Bureau of Labor Statistics. Summary statistics of the whole dataset can be found in Table 1.

Statistic	N	Mean	St. Dev.	Min	Max
Stage	42,672	21.91	3.72	1.00	26.00
Rank difference	42,672	40.32	43.10	1.00	377.00
Tournament count	42,672	58.60	46.54	1.00	267.00
Experience	42,672	1.20	0.63	1.00	16.00
Prize pool (positive) (in $1 000)	2,214	68.03	186.35	0.21	2,345.31
Without prize pool	1,494				
Major	12				
Form	133,031	1.04	0.33	0.07	3.07
	S	A	B	C	
Tier	177	360	950	2,221	
	Online		LAN		
Type	34,756		7,916		
	Best-of-1		Best-of-3		Best-of-5
Format	15,329		27,106		237

Table 1 Summary statistics of the whole dataset

2.4 Model

To assess whether and under which circumstances players choke, the following semiparametric generalized additive model is estimated:

$$\begin{aligned}
\text{Performance}_{iklmpt} &= \beta_0 + \beta_1\text{Major}_t + \beta_2\text{S}_t + \beta_3\text{A}_t + \beta_4\text{B}_t + \beta_5\text{Elimination_match}_{mt} \\
&+\beta_6\text{Elimination_map}_{mpt} + \beta_7\text{Playoffs}_{mt} + \beta_8\text{Log_prize_pool}_t + \beta_9\text{Log_stage}_{mt} + \beta_{10}\text{Type}_{mt} \\
&+\beta_{11}\text{Form}_{it} + \beta_{12}\text{Strength_difference}_{mklt} + \beta_{13}\text{For_money}_t + \beta_{14}\text{Rating_2.0}_{impt} \\
&+\sum_{j=1}^{9}\beta_{i+14}\text{Map_name}_{jmpt} + f(\text{Experience}_{imt}) + f(\text{Log_stage}_{mt}) + f(\text{Log_prize_pool}_t) \\
&+f(\text{Experience}_{imt}, \text{Major}_t) + f(\text{Experience}_{imt}, \text{Tier}_t) + f(\text{Experience}_{imt}, \text{Elimination_match}_{mt}) \\
&+f(\text{Experience}_{imt}, \text{Elimination_map}_{mpt}) + f(\text{Experience}_{imt}, \text{Playoffs}_{mt}) + f(\text{Experience}_{imt}, \text{Type}_{mt}) \\
&+\text{u}_{impt}
\end{aligned} \quad (4)$$

for player i on map p in match m between teams k and l in tournament t. The smooths are estimated via restricted maximum likelihood (REML) due to efficiency and numeric stability (Wood, 2017). All the smoother functions are modeled as penalized thin plate regression splines (Marra & Wood, 2011) with a basis complexity of 8 since Winter and Wieling (2016) recommend using at most half of the unique values of Experience. This should allow both modeling complex behavior of the variables but preventing excessive *wiggliness* or *fitting the noise*.

3 Results

Results of the regression can be found in Table 2. It consists of coefficient estimates of the parametric fixed effects and estimates of complexity of the smoothing terms. All the k-indices for the smoothers are >= 0.98, suggesting that the maximum complexity of the smoothers picked was not too strict.

	Dependent variable:	Performance
	Estimates	Standard errors
Major	-0.005^{***}	(0.003)
S-Tier	-0.021^{***}	(0.002)
A-Tier	-0.022^{***}	(0.001)
B-Tier	-0.008^{***}	(0.001)
Elimination_match	0.007^{***}	(0.001)
Elimination_map	-0.001	(0.001)
Playoffs	-0.012^{***}	(0.001)
LAN	-0.005^{***}	(0.001)
	edf	F
f(Log_stage)	4.181^{***}	(10.076)
f(Log_prize_pool)	1.324^{***}	(6.901)
f(Experience)	0.693^{*}	(0.332)
f(Experience):Major	0.869^{**}	(3.297)
f(Experience):Not_Major	0.047^{*}	(0.010)
f(Experience):S-Tier	2.858^{***}	(6.511)
f(Experience):A-Tier	0.070	(0.360)
f(Experience):B-Tier	0.014	(0.663)
f(Experience):C-Tier	0.059	(0.300)
f(Experience):Elimination_match	2.527^{***}	(46.178)
f(Experience):Not_elimination_match	0.025^{*}	(0.004)
f(Experience):Elimination_map	2.730^{***}	(6.916)
f(Experience):Not_elimination_map	0.002^{*}	(0.000)
f(Experience):Playoffs	2.284^{***}	(19.809)
f(Experience):Not_playoffs	0.000^{*}	(0.000)
f(Experience):LAN	0.025	(0.001)
f(Experience):Online	0.012^{***}	(0.000)
Control variables and constant		Included
Deviance explained		12.8%
Note: * p < 0.05; ** p < 0.01; *** p < 0.001		

Table 2 Estimates of the Generalized Additive Model

The shrinkage is successful for most of the smooth, suggesting that only some dominate the model, which was expected due to Experience being a part of most of them. The model explained 12.8% of the deviance, which suggests that either important variables are missing or that player performance in CS on a single map has a significant random component. The latter seems likely and in line with, e.g., Abelson (1985), indicating that skill differences between players manifest over an extended number of maps and matches.

The partial effects of the variables can be found in Figures 1 and 2. All the fixed effects are negative and significant, as expected, except for the Elimination_match, which is positive, and the Elimination_map, which is not significant at 5%. The smooth term for the Log prize pool is close to linear and decreasing, indicating that a higher prize pool negatively influences performance.

The smooth term for the log of the stage initially increases for over 99% of the sample but starts decreasing for the most experienced players. However, the confidence intervals (CI) are large for these highly experienced players due to the small sample size, and zero is within the CI, suggesting uncertainty in this effect.

The shape of the interaction smooths is similar for Playoffs, Elimination map, and Elimination match, where experience initially has a positive moderating effect but decreases later. These smooths also exhibit large CIs, akin to the smooth for the Log_stage.

Conversely, the interaction smooth for S-Tier shows an initial decrease followed by a rise, with large CIs, but zero does not pass through the interval, indicating a more robust effect.

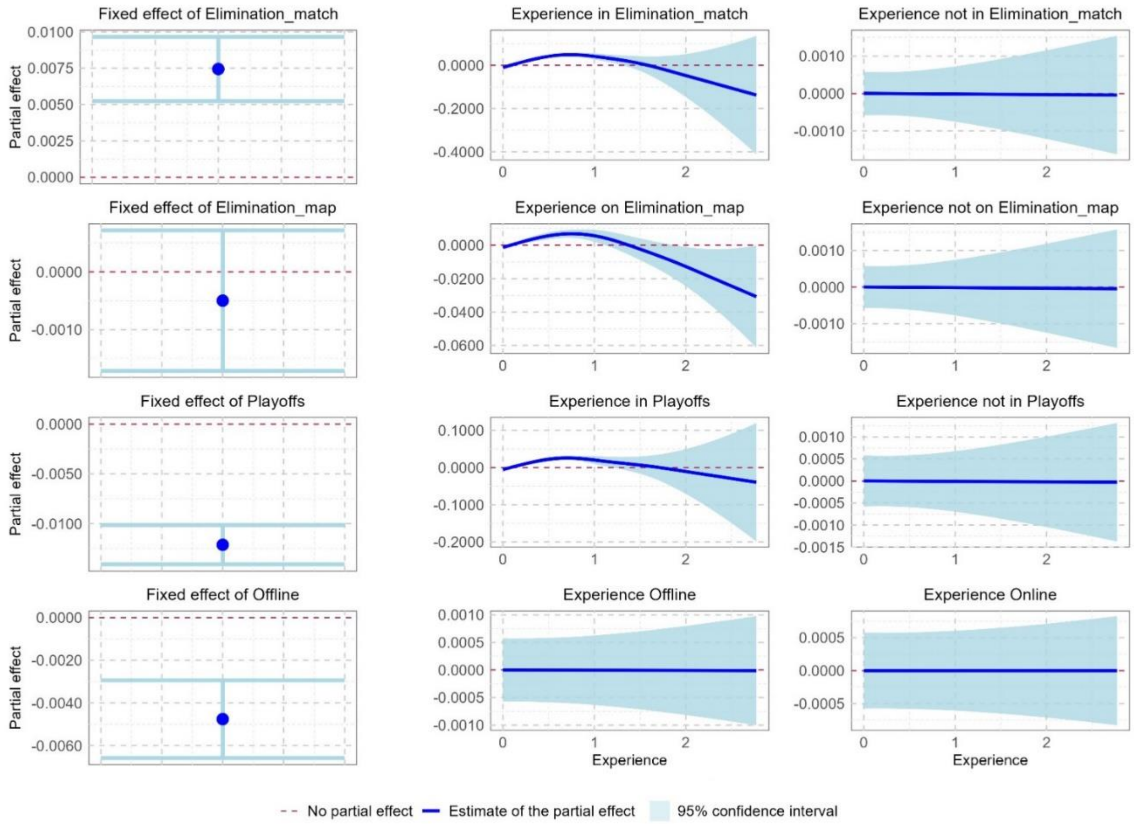


Figure 1 Partial effects of the stressors and their interactions with Experience (Part 1/2)

The interaction smooth for Majors is strictly increasing, demonstrating that experience has a consistently positive moderating effect on performance in Majors. In contrast, experience does not significantly affect performance on LAN and A-Tier tournaments.

3.1 Discussion

Players perform worse once the amount of stress increases, particularly in higher prestige tournaments, LAN tournaments, and playoffs, and players perform worse further in the tournaments. Experience mitigates performance deficits, especially in high-pressure situations such as facing elimination. This study confirms that pressure significantly negatively affects player performance in CS tournaments, thus firmly suggesting that players choke under pressure, which is in line with the sports economics literature (e.g., Dohmen, 2008; Piskorska et al., 2016).

The size of a tournament's prize pool negatively impacts player performance, aligning partially with Shengjie (2019), who, rather than choking under pressure, suggests a crowding-in effect that diminishes intrinsic motivation. The findings support the importance of mental strength in elite gaming (Wagstaff & Leach, 2015) and emphasize the need for psychological interventions and specialized training.

4 Conclusion

This study examines choking under pressure in professional esports, specifically Counter-Strike. Analyzing the complete competitive history of Counter-Strike: Global Offensive via a generalized additive model, it finds a clear presence of choking under pressure induced by various tournament and player features. It also finds a non-linear moderation effect of experience for at least some features.

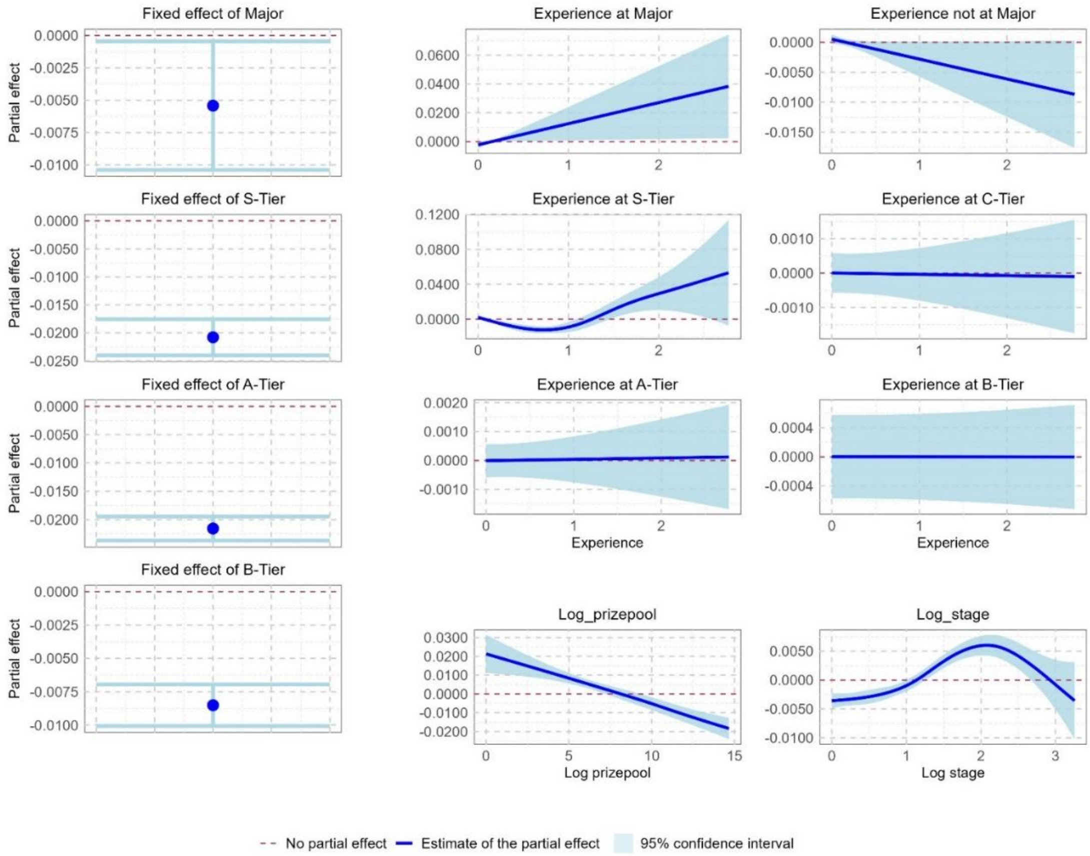


Figure 2 Partial effects of the stressors and their interactions with Experience (Part 2/2)

As esports gain mainstream recognition, understanding performance dynamics is crucial. This study lays the ground for future research on esports economics both as an extension of sports economics to the realm of esports and as a field delving into esports-specific facets.

Acknowledgements

This work was supported by The Internal Grant Agency of the Prague University of Economics and Business under Grant VŠE IGS F4/52/2023.

References

Abelson, R. P. (1985). A variance explanation paradox: When a little is a lot. Psychological Bulletin, 97(1), 129–133. https://doi.org/10.1037/0033-2909.97.1.129

Dohmen, T. J. (2008). Do professionals choke under pressure? Journal of Economic Behavior & Organization, 65(3), 636–653. https://doi.org/10.1016/j.jebo.2005.12.004

Flegr, S., & Schmidt, S. L. (2022). Strategic management in eSports – a systematic review of the literature. Sport Management Review, 25(4), 631–655. https://doi.org/10.1080/14413523.2021.1974222

Hallmann, K., & Giel, T. (2018). eSports – Competitive sports or recreational activity? Sport Management Review, 21(1), 14–20. https://doi.org/10.1016/j.smr.2017.07.011

Hamari, J., & Sjöblom, M. (2017). What is eSports and why do people watch it? Internet Research, 27(2), 211–232. https://doi.org/10.1108/IntR-04-2016-0085

Hill, D. M., Hanton, S., Matthews, N., & Fleming, S. (2010). Choking in sport: A review. International Review of Sport and Exercise Psychology, 3(1), 24–39. https://doi.org/10.1080/17509840903301199

HLTV.org. (2024). HLTV.org—The home of competitive Counter-Strike. HLTV.Org. https://www.hltv.org/

Klaassen, F. J. G. M., & Magnus, J. R. (2003). Forecasting the winner of a tennis match. European Journal of Operational Research, 148(2), 257–267. https://doi.org/10.1016/S0377-2217(02)00682-3

Kolev, G. I., Pina, G., & Todeschini, F. (2015). Decision Making and Underperformance in Competitive Environments: Evidence from the National Hockey League. Kyklos, 68(1), 65–80. https://doi.org/10.1111/kykl.12073

Liquipedia. (2024). Portal:Tournaments. https://liquipedia.net/counterstrike/Portal:Tournaments

Marra, G., & Wood, S. N. (2011). Practical variable selection for generalized additive models. Computational Statistics & Data Analysis, 55(7), 2372–2387. https://doi.org/10.1016/j.csda.2011.02.004

Meng-Lewis, Y., Wong, D., Zhao, Y., & Lewis, G. (2022). Understanding complexity and dynamics in the career development of eSports athletes. Sport Management Review, 25(1), 106–133. https://doi.org/10.1016/j.smr.2020.08.003

Nagorsky, E., & Wiemeyer, J. (2020). The structure of performance and training in esports. PloS One, 15(8), e0237584.

Parshakov, P., Shakina, E., & Barajas, A. (2022). Testing the contestable market theory in eSports. Economic Research-Ekonomska Istraživanja, 35(1), 122–142. https://doi.org/10.1080/1331677X.2021.1889389

Piskorska, E., Mieszkowski, J., Kochanowicz, A., Wedrowska, E., Niespodziński, B., & Borkowska, A. (2016). Mental skills in combat sports – review of methods anxiety evaluation. Archives of Budo, 12, 301–313.

Reitman, J. G., Anderson-Coto, M. J., Wu, M., Lee, J. S., & Steinkuehler, C. (2020). Esports Research: A Literature Review. Games and Culture, 15(1), 32–50. https://doi.org/10.1177/1555412019840892

Rejthar, J. (2024). Unlocking the Untapped Potential of Video Game Data: A Case Study of Aim Trainers. In L. Vale Costa, N. Zagalo, A. I. Veloso, E. Clua, S. Arnab, M. Vairinhos, & D. Gomes (Eds.), Videogame Sciences and Arts (Vol. 1984, pp. 315–321). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-51452-4_22

Rejthar, J., & Kotrba, V. (2023). Confirmation of the validity of the HLTV ranking. ICGA Journal, 45(1), 2–15. https://doi.org/10.3233/ICG-230229

Ritchie Vink, Stijn de Gooijer, Alexander Beedie, Marco Edward Gorelli, Weijie Guo, J van Zundert, Gert Hulselmans, universalmind303, Orson Peters, Marshall, chielP, nameexhaustion, Matteo Santamaria, Daniël Heres, Josh Magarick, ibENPC, Moritz Wilksch, Jorge Leitao, Jonas Haag, … Adrián Gallego Castellanos. (2024). pola-rs/polars: Python Polars 0.20.10 (py-0.20.10) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.10679411

Shengjie, Y. (2019). For Love of the Game: A Study of Tournament Theory and Intrinsic Motivation in Dota 2 [PhD Thesis]. Duke University Durham.

Simpson, G. L. (2024). gratia: Graceful ggplot-Based Graphics and Other Functions for GAMs Fitted using mgcv [Computer software]. https://gavinsimpson.github.io/gratia/

Wagstaff, C. R. D., & Leach, J. (2015). The value of strength-based approaches in SERE and sport psychology. Military Psychology, 27(2), 65–84. https://doi.org/10.1037/mil0000066

Watanabe, N. M., Xue, H., Newman, J. I., & Yan, G. (2021). The Attention Economy and Esports: An Econometric Analysis of Twitch Viewership. Journal of Sport Management, 36(2), 145–158. https://doi.org/10.1123/jsm.2020-0383

Weber, R., Behr, K.-M., Tamborini, R., Ritterfeld, U., & Mathiak, K. (2009). What Do We Really Know about First-Person-Shooter Games? an Event-Related, High-Resolution Content Analysis. Journal of Computer-Mediated Communication 14(4), 1016–1037. https://doi.org/10.1111/j.1083-6101.2009.01479.x

Winkler, A., 2016. A history of Liquipedia [WWW Document]. URL https://tl.net/blogs/505910-a-history-of-liquipedia (accessed 7.31.21).

Winter, B., & Wieling, M. (2016). How to analyze linguistic change using mixed models, Growth Curve Analysis and Generalized Additive Modeling. Journal of Language Evolution, 1(1), 7–18. https://doi.org/10.1093/jole/lzv003

Wood, S. N. (2001). mgcv: GAMs and generalized ridge regression for R. R News, 1(2), 20–25.

Wood, S. N. (2017). Generalized Additive Models: An Introduction with R (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279

Xenopoulos, P., Doraiswamy, H., & Silva, C. (2020). Valuing Player Actions in Counter-Strike: Global Offensive, In: 2020 IEEE International Conference on Big Data (Big Data). Atlanta, GA, USA, pp 1283–1292. https://doi.org/10.1109/BigData50022.2020.9378154

Impact of Asset Price History on Price Volatility

Michaela Sedláková[1]

Abstract. The effect of the initial price history on asset price volatility is studied using a Learning to Forecast experiment. Participants are tasked with predicting the future prices of three distinct risky assets over many consecutive periods. Unlike previous Learning to Forecast experiments that focused on a single risky asset, this research allows participants to compare the price trajectories of individual assets. One asset is characterized by a very stable initial price development compared to the other two assets. Given that all risky assets share the same fundamental value, we can investigate the effect of different initial price history on the overall price dynamics. Our conjecture is that the asset with a stable initial price history will exhibit lower volatility compared to the other two assets. To validate this, we employ statistical tests based on selected volatility measures, including relative absolute deviation from the fundamental value and variance. From the results, it is clear that the asset characterized by a stable initial price history in most cases demonstrates reduced price volatility.

Keywords: experimental economics, expectations, asset pricing

JEL Classification: C92, D84, G12, G41
AMS Classification: 91A90, 91G30

1 Introduction

Expectations have a significant impact on investment decision making, as well as on aggregate market dynamics. Through the Learning to Forecast (hereinafter LtF) experiment it is possible to study how people constitute their expectations and how these expectations affect financial markets or macroeconomic. In behavioral finance, LtF experiments are often used to test various hypotheses about how people form their expectations and how these expectations may differ from rational ones. Participants in these experiments usually play the role of professional forecasters and their task is to set their expectations about various economic indicators, for example asset price, inflation rate etc.

We conducted a LtF experiment at the VSB – Technical University of Ostrava, where students of the Faculty of Economics were in the role of financial forecasters for a pension fund. The experimental framework is based on Anufriev et al. (2022), where the participants' task is to predict the future price of a risky asset. Unlike previous LtF experiments focused on a single risky asset, our experiment extended the investment options to three risky assets. For these risky assets, the initial price history was also provided, where one asset was characterized by a stable historical price development compared to the other two assets. Contrasting previous experiments involving a single risky asset, participants can compare the price development of each individual asset with one another.

The goal of this paper is to study the effect of the initial price history on asset price volatility, measured by a relative absolute deviation and a variance of the asset price. The conjecture is that the asset with a stable initial price history will be less volatile, and this will be verified via selected two-sample statistical tests.

2 Literature review

The formation of expectations plays an important role in modern financial and macroeconomic modelling. The concept of rational expectation hypothesis was first introduced by Muth (1961) and later extended by Lucas (1972). The fundamental idea behind rational expectations is that individuals make decisions based on all available information, including their own expectations of future events. The rational expectations hypothesis is incorporated into many models of market economy functioning based on the theory of general equilibrium and accounting for factors like uncertainty and imperfect information.

Due to the complexity of the real world, it is mostly difficult to perform a clean test on the rational expectation hypothesis using empirical data. As a result, experimental economists often turn to LtF experiments to study expectation formation and learning in financial markets and macroeconomics. Learning to Forecast experiment was

[1] VSB – Technical University of Ostrava, Department of Finance, Sokolská třída 33, 702 00 Ostrava 1, Czech Republic, michaela.sedlakova@vsb.cz

first introduced in Marimon et al. (1993). In LtF experiments, human subjects provide their expectations for economic variables (such as market prices or inflation rates). Once all individual predictions are collected, the realized variable is calculated using a computer algorithm, and this is repeated for many consecutive periods.

Traditionally, LtF experiments are primarily focused on questions related to asset pricing where participants play the role of a professional forecaster. Hommes et al. (2008) explored expectation formation in a stationary asset pricing experiment, where the rational expectations fundamental price was constant. Despite this, significant price bubbles occurred in most of their experiments, sometimes exceeding 16 times the fundamental value. These bubbles were mainly driven by positive expectations feedback. Additionally, the participants were highly coordinated and followed a common prediction strategy. For a comprehensive review of LtF experiments, refer to Bao et al. (2021).

In this contribution, the focus is on the impact of initial price history on asset price volatility. Few papers have explored this area. Hennequin (2022) conducted a two-stage LtF experiment. In the first stage, the market consisted of one participant and five robots, ensuring a specific price development. Participants experienced either a stable asset market with small deviations from the fundamental value or a bubbly asset market. In the second stage, the market included only human subjects. The study results revealed that the first stage significantly influenced aggregate price dynamics in the second stage. The convergence to the fundamental value was much faster in the groups where all subjects experienced the stable market in the first stage. In contrast, the occurrence and emergence of bubbles was typical for groups where participants experienced very significant fluctuations in the market during the first stage of the experiment. The effect of investment horizon on asset price volatility is studied in Anufriev et al. (2022). At the beginning of their experiment, the participants received the price history of the risky asset, which was characterized by a very stable price development or a more volatile historical development. The treatments of the experiment therefore differed in terms of investment horizon and asset price history. It is clear from their results that, regardless of the length of the investment horizon, low volatility of the asset price was demonstrated if the initial price history was stable.

3 Experimental design

Six experiments were conducted at the VSB – Technical University of Ostrava during October and November 2022 and also April and May 2023. A total of 75 students from the Faculty of Economics participated in these experiments and the group sizes ranged from 9 to 17 students. The experiments took place in a computer classroom, where all participants operated within the same market throughout the entire experiment. At the beginning of each session, participants received detailed instructions, including printed copies, and were familiarized with their task. No communication was allowed between participants during the experiment. After completing the experiment, students filled out a questionnaire and received payment based on their ranking. The payment amounts ranged from 50 to 700 CZK and the ranking was determined by their average prediction error.

3.1 General information

The experimental design is based on the experiment in Anufriev et al. (2022). Participants are introduced to the experiment in the following manner. Participants play the role of a financial forecaster for a pension fund that needs to optimally invest a large amount of money for one period. The pension fund has several investment options: risk-free asset and three risky assets. The instructions explicitly state that the risky assets are not correlated. In the case of risk-free asset, the money is invested into a government bond which pays a fixed interest rate of 5%. Alternatively, a pension fund can allocate funds to shares of indefinitely lived risky assets. These risky assets are associated with uncertainty regarding future prices and dividends. The dividends are independently and identically distributed with a mean of $10 per period. The participants' task is explained to make a prediction of future asset prices, based on which the pension funds will make trading decisions. The instructions do not specify the exact pricing equation in accordance with the standard practice of LtF experiments. However, some quality characteristics of the market are described. For instance, a higher price forecast leads to increased demand for assets, and several funds influence total demand. As mentioned above, we extend our research to the number of risky assets which are distinguished by different initial price history. With increased number of risky assets, participants can compare price developments of all assets in the market. Specifically, assets A1 and A2 exhibit greater volatility, while asset A3 demonstrates a stable price development around its fundamental value (as depicted in Figure 1, where the first 10 periods represent the initial price history). Notably, all these assets share the same fundamental value of $200.

During the experiment, participant's available information for the price prediction of the period $t+1$ in period t consists of:

- past realized prices up to period $t-1$,
- participant's previous price predictions up to period t,
- The total average error of the participant, as well as the average errors for particular assets.

Once all predictions from all participants for period $t+1$ were received, the realized price of assets for the current period t was determined according to equation (1) and this was repeated for all 50 consecutive periods.

3.2 Price generating mechanism

The experimental design is based on the typical LtF experimental set up, incorporating the present value model of asset pricing. Mean-variance investors divide their wealth into risk free and risky assets. The gross return of risk free asset is $R=1+r>1$ and all risky assets pay an IID dividend with mean $\bar{y}$ each period. The market-clearing price $p_{t,a}$ of an asset a in a time period t is defined according to Brocks and Hommes (1998) as follows,

$$p_{t,a} = p_a^f + \frac{1}{(1+r)}\left(\bar{p}_{t+1,a}^e - p_a^f - \varepsilon_{t,a}\right), \tag{1}$$

where p_a^f is fundamental value of particular asset a which is calculated as a present value of dividend payment $p_a^f = \bar{y}_a / r$, r is discount factor, $\bar{p}_{t+1,a}^e$ are average expectations about price in the period $t+1$ for the asset a, and $\varepsilon_{t,a}$ is a small random outside supply of the asset from noise traders.

It is obvious from equation (1) that the market price $p_{t,a}$ is a weighted average of the fundamental value and average expectations for the period $t+1$. If an increase in price is expected in the future, it increases the demand in the current period as well as the price. This is called positive expectations feedback. The rational equilibrium is given by the fundamental value of the asset.

3.3 Conjecture

In this paper, we focus on examining how initial price history affects subjects' expectations and what impact it has on asset volatility. Our conjecture is that the asset A3 with stable initial price history will be less volatile in comparison with the other two risky assets. Two measures of volatility were chosen, namely a relative absolute deviation and a price variance of the risky asset. The relative absolute deviation (RAD) was proposed by Stöckl et al. (2010) and is calculated as follows,

$$RAD_{t,a} = \frac{\left|p_{t,a} - p_a^f\right|}{p_a^f}. \tag{2}$$

This measure is frequently used for studying bubble formation in LtF experiments. Our conjecture, will be tested using selected two-sample tests. The alternative hypothesis is that the volatility of asset A1 (respectively A2) is higher than the volatility of the asset A3.

4 Results

We start the discussion of the overall market behavior. The realized asset prices for all six experiments are presented in Figure 1. The experiments could be classified into two groups, namely stable markets (EXP1, EXP2, EXP3) and markets with medium-sized bubbles (EXP4, EXP5, EXP6). First, the overall price dynamics of stable markets will be assessed. In the case of EXP1, asset prices exhibit stable small oscillations around the fundamental value regardless of the initial price history. Compared to the other experiments, EXP2 has an atypical development. All asset prices show monotonic convergence to the fundamental value. The development of asset prices in EXP3 demonstrates persistent oscillations, with no apparent convergence to fundamental value. Nevertheless, it is not very simple to assess the effect of initial price history on asset price volatility from the figures in the case of stable markets. In order to evaluate the impact of initial price history, we can further focus on the interquartile range (IQR) of asset prices (see Table 1). In most cases, the IQR of the A3 asset is lower compared to other assets. The exception is EXP1, where asset A1 shows the lowest interquartile range.

Attention will now shift to assessing price developments in medium-sized bubble markets. In the case of EXP4[2], there is consistently stable price development until the 25th period across the entire market. Subsequently, assets A1 and A2 present more pronounced oscillations, but the amplitude gradually decreases and converges to the fundamental value. On the other hand, the A3 asset is characterized by a very stable development close to the fundamental value. In the case of EXP5, it can again be seen that asset A3 exhibits significantly lower deviations from fundamental value compared to the other assets. EXP6 is characterized by persistent oscillations around the fundamental value, where the amplitude of these oscillations is more or less constant for all assets. Based on the graphical analysis, it can be assumed that the initial price history of the asset has an impact on the asset price volatility, as the A3 asset exhibits lower deviations from fundamental value in most cases. This hypothesis will now be tested through two-sample statistical tests.

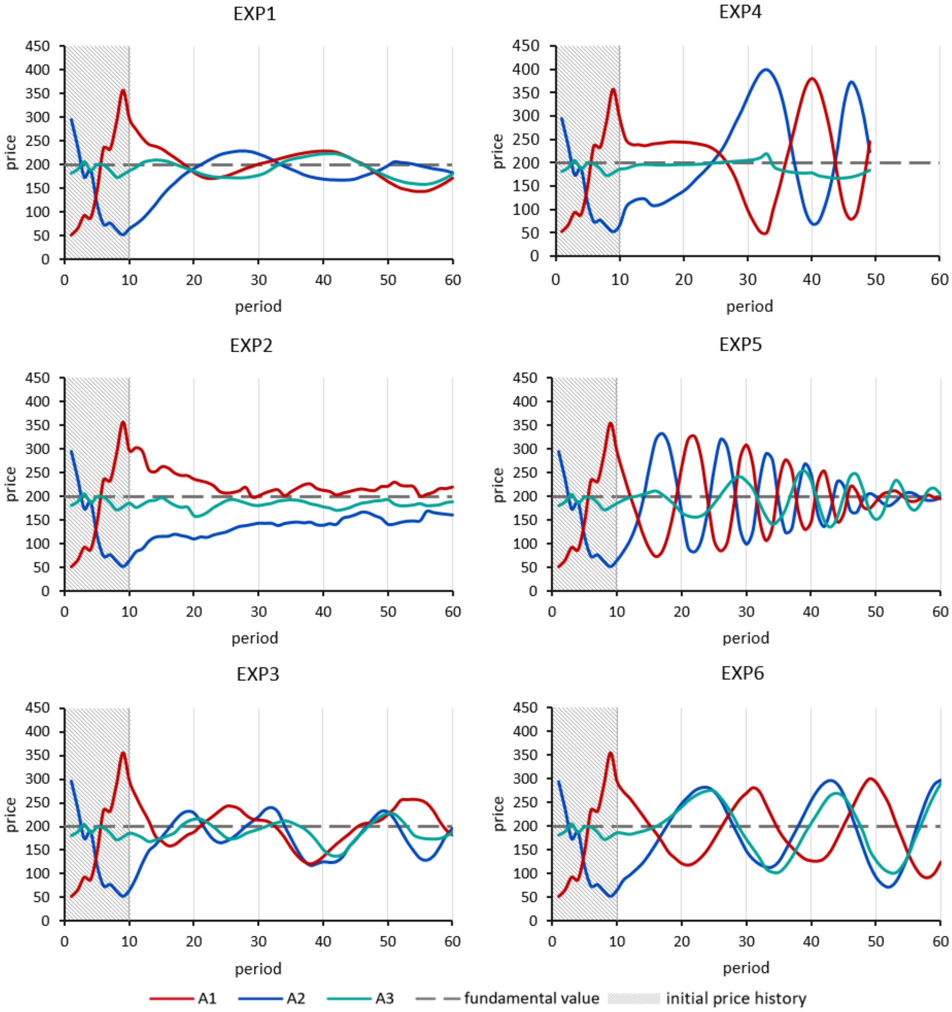


Figure 1 Price dynamics in all sessions for all three assets. Left – stable markets, right: markets with medium-sized bubbles. Please note that period from 1 to 10 represents the initial price history.

As part of the statistical testing, we will always test the hypothesis that asset A3 exhibits a lower level of volatility than assets A1 and A2, respectively. Two measures of volatility have been chosen, namely relative absolute deviation and variance. To test the statistical significance of the relative absolute deviation from the fundamental price, the Mann–Whitney[3] one-tailed test is chosen, as the normality of the data is violated in most experiments. Based on the summary results, it can be concluded that the initial price history of the asset has a statistically significant effect on the asset price volatility at the 5% significance level, as the null hypothesis was rejected in 9 out of 12 cases (see Table 2). Very similar results are obtained by testing the variance of realized prices via Levene's test[4], where the null hypothesis is rejected in 10 out of 12 cases.

[2] We have obtained price predictions only until the period 50 due to technical issues.
[3] Refer to Mann and Whitney (1947).
[4] Alternative formulation of Levene's test statistic based on median used for testing equality of variances (see Brown and Forsythe, 1974).

As mentioned before, the influence of information regarding initial price history was also investigated in Anufriev et al. (2022). It is clear from their results that the tendency for price bubbles to form is lower if the asset price history is stable, regardless of the investment horizon. In our experimental setting, the research is extended to study the effect of initial price history in the market of three uncorrelated assets. Thus, it was possible to investigate how price history influences participants' expectations with respect to the behavior of other assets. Based on the overall results, it can be concluded that similar results were achieved compared to Anufriev et al. (2022).

Experiment	Asset	Mean price	Median price	Standard deviation	IQR	RAD (%)
EXP1	A1	183,83	189,82	34,48	33,76	13,03
	A2	196,13	198,48	31,83	49,59	13,36
	A3	189,19	187,74	20,32	35,08	9,77
EXP2	A1	137,92	142,14	20,05	27,99	31,04
	A2	223,23	216,31	22,04	18,27	11,65
	A3	183,19	183,54	7,99	9,43	8,40
EXP3	A1	176,67	176,52	39,09	62,03	18,47
	A2	200,55	207,05	38,12	57,20	15,70
	A3	188,22	186,91	21,95	29,86	9,92
EXP4	A1	215,26	193,37	104,88	184,75	45,23
	A2	207,91	237,45	83,96	90,00	34,35
	A3	189,75	195,13	13,66	19,97	6,49
EXP5	A1	197,30	193,59	64,59	68,27	25,10
	A2	192,20	194,77	63,04	78,84	24,42
	A3	195,33	199,68	31,37	46,55	13,04
EXP6	A1	190,96	190,39	72,12	137,60	31,87
	A2	190,12	181,19	61,09	111,94	27,13
	A3	195,98	196,50	56,94	92,17	23,88

Table 1 Descriptive statistics for all experiments.

p-values of the one-sided Mann-Whitney test comparing the RAD

	EXP1	EXP2	EXP3	EXP4	EXP5	EXP6
A1	0,3440	0,0001**	0,0002**	0,0001**	0,0042**	0,0085**
A2	0,0282*	0,2763	0,0041**	0,0001**	0,0076**	0,1364

p-values of the Levene's test comparing the variance

	EXP1	EXP2	EXP3	EXP4	EXP5	EXP6
A1	0,1095	0,0001**	0,0001**	0,0001**	0,0003**	0,0147*
A2	0,0018**	0,0021**	0,0005**	0,0001**	0,0004**	0,3704

Table 2 p-values of the one-sided Mann-Whiteney test comparing the RAD (upper table) and price variance of the Levene's test (lower table), pairwise for the asset A3 and A1 or A2, respectively. The alternative hypothesis is that the row asset A3 has a value of the volatility measure lower than the column asset. * and ** denote asset comparisons when the null hypothesis of equality is rejected at the 5% or 1% significance level, respectively.

5 Conclusion

In this contribution, we investigated the impact of the initial price history on asset price dynamics through a Learning to Forecast experiment. Our investigation was based on a straightforward experimental asset pricing model. While previous LtF asset pricing experiments typically focused on predicting the price of a single risky asset over many consecutive periods, we extended our research to markets with three uncorrelated risky assets. This extension allowed us to explore whether participants' expectations were also influenced by the performance of other assets in the market.

The individual assets in our study exhibited varying initial price histories. Notably, asset A3 demonstrated stable historical development around its fundamental value compared to the other assets. Our primary objective was to examine how this initial price history impacted asset price volatility. In most cases, asset A3, with its stable historical development, consistently exhibited lower volatility. Deviations from the fundamental value were significantly smaller for this asset. We then tested the hypothesis that asset A3 has lower volatility than other assets using two measures: relative absolute deviation and variance. The results of the statistical test applied to the relative absolute deviation show that the null hypothesis of equal asset price volatility was rejected in 9 out of 12 cases (in the case of testing price variance in 10 out of 12 cases). In other words, assets with more volatile initial price histories were prone to forming bubbles and experiencing higher volatility. In summary, while all assets shared the same fundamental price, their initial price history played a crucial role in determining their subsequent volatility.

Acknowledgements

The research was financially supported by the SGS projects No. SP2023/019 and SP2024/047 of VSB – Technical University of Ostrava.

References

Anufriev, M., Chernulich, A., & Tuinstra, J. (2022). Asset price volatility and investment horizons: An experimental investigation. *Journal of Economic Behavior & Organization*, *193*, 19–48.

Bao, T., Hommes, C., & Pei, J. (2021). Expectation formation in finance and macroeconomics: A review of new experimental evidence. *Journal of Behavioral and Experimental Finance*, *32*, 100591.

Brock, W. A., & Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, *22*(8–9), 1235–1274.

Brown, M. B., & Forsythe, A. B. (1974). Robust Tests for the Equality of Variances. *Journal of the American Statistical Association*, *69*(346), 364–367.

Hennequin, M. (2021). Experiences and expectations in asset markets: an experimental study. *University of Sussex Working Paper*.

Hommes, C., Sonnemans, J., Tuinstra, J., & Van de Velden, H. (2008). Expectations and bubbles in asset pricing experiments. *Journal of Economic Behavior & Organization*, *67*(1), 116–133.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, *4*(2), 103–124.

Mann, H., & Whitney, D. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Annals of mathematical Statistics*, *18*(1), 50–60.

Marimon, R., Spear, S. E., & Sunder, S. (1993). Expectationally Driven Market Volatility: An Experimental Study. *Journal of Economic Theory*, *61*(1), 74–103.

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, *29*(3), 315–335.

Stöckl, T., Huber, J., & Kirchler, M. (2010). Bubble measures in experimental asset markets. *Experimental Economics*, *13*(3), 284–298.

Knapsack problem for build spare parts stock

Robert Szczyrbak[1]

Abstract. The manufacture of products and services involves the operation of machinery and equipment. In the course of operations, machinery gradually wears out, causing product defects, breakdowns or even to accidents. Product defects and failures strongly affect the OEE rate. Machine failures reduce the availability rate. The role of the maintenance department is to ensure the safe and continuous operation of the machinery fleet. For this purpose, the costs of technical materials and labor costs are incurred.

Ensuring continuous operation of the machinery park is both the implementation of periodic inspections associated with the replacement of worn-out components, but also the minimization of repair time. The main factor causing an increase in repair time is the lack of possession of the required parts in the technical warehouse.

In addition to standard indicators related to logistical aspects, it is worth using mathematical models for this purpose, such as integer programming models. To develop such a model, the research problem should be defined. It can be related both to reducing the value of the warehouse or total reduction of fixed costs, but also to raising the value of the machinery availability index. This paper shows knapsack problem, which can be used in technical warehouse. The main goal of this model is to minimize loses in production due failures

Securing in spare parts of the machinery park is a very important part of the management strategy. Decisions related to whether or not to have a particular component are worth justifying with appropriate arguments formed on the basis of analyses of availability and uniqueness of spare parts. It is also worth remembering that parts in stock are subject to obsolescence and in a few years a given component may not fully perform its functions.

Keywords: maintenance, Operational research, spare parts, knapsack problem

1 Introduction

Main goal of maintenance departments is to keep in good condition all machines in factory. They do this by periodically inspecting machines, replacing wear parts and overhauling machines. But there is a risk, when factory has got many production lines. Good practice is to set machine classifitaction. Santos et. al. [7] prepared a practical and structured method of equipment classification criticality based on its importance for the productive process, which culminated in the equipment classification into three categories (A, B and C). This method included five factors: Quality, Availability, Safety and Environment, Costs and Technological Complexity. When we have got machine classification we can set similar classification for spare parts. Before we start, we need to know what spare parts is. Ramaganesh et al. [5] defined spare parts as a interchangeable parts that are retained in inventory and used for failed unit repair or replacement. Spare parts availability has a huge impact on breakdown time. Maintenance teams need to know which parts they should stock to avoid long downtimes that generate high costs. The unavailability of the needed spare parts will affect the efficiency of the machine and incur unnecessary costs. Yet overstocked replacement parts can also incur costs in terms of inventory cost. Two related logistics operations are the provision of spare parts and scheduled repairs, which must be viewed together to achieve cost-effective and reliable logistics support. Afifi et. al. [1] said that to increase machine availability, several components in the machines should be replaced before failure occurs. Therefore, optimizing spare parts inventory management is a key factor in improving maintenance activities. Hu et. al [2] observed that the cost of spare parts takes a large share of product lifecycle cost. Maintenance departments include the delivery time of the required parts in their machine maintenance plans. In order to avoid a capital freezing situation, parts are delivered several days before the work is started. Several activities in maintenance departments can be analyzed and solved by using operational research. This paper shows one of approach to maintenance problem.

[1] AGH Univeristy of Krakow, rszczyrbak@agh.edu.pl

2 Literature review

Spare parts management process affects the performance of maintenance. Availability of spare parts has major impact for stock building. Rinaldi et. al [6] notice that the spare parts refer to service parts, replacement parts, or repair parts that are used to support maintenance and repair operations. The spare parts supply chain differs from the other supply chains for different reasons, mainly the uncertainty and the variability of the demand, combined with the high service level required.

Throughout the last decades we can use useful technique in the field of operations management, several can be used for managing spare parts, mainly for modelling and evaluating inventory policies.

This section is divided into 2 topics: spare parts management approach and knapsack problem in operational research.

2.1 Spare parts management approach

Texeira et.al. [8] said that the spare parts management has obtained great interest in literature. Various topics are addressed concerning spare parts, such as inventory control, demand forecasting and reliability, and supply chain management. This paper focuses on framework of spare parts management according to operational research and classification methods. They also gave advice to set the stock management policy, it will be important to take into account two aspects: the number of machines for which the spare parts are used and the demand for spare parts. There is better situation where we have got more than one type of machine in factory. If we have got very different machines, we need to buy many spare parts.

Hu et. al [2] marks out the questions in spare parts management about decisions which parts are to be stocked as spare parts, when and how many parts should be ordered to bulid or refresh the stock. Spare parts management should refer to life cycle of machine. In their article, life cycle cost (LCC) is closely connected to investment and management of spare parts inventories. In general, the equipment lifecycle process can be divided into three main phases:

- Initial procurement phase
- Normal operation phase
- End-of-life phase

When the LCC is set we can assess spare parts by some criteria. Each phase has got different criterias. They are determined by different needs at the mentioned phases.. In the first stage, we will need parts, qualified as consumable and parts whose criticality is at a high level. The second phase represents the longest life cycle of the machine. so we should take care of repair kits and optimize EOQ and reorder point. In the long run, we should also watch out for emerging component modifications. It happens that manufacturers withdraw the component we use, and we must then carry out an upgrade. If the number of recalled parts is significant and the manufacturer announces the end of support, we should make a decision: secure in the components still available or prepare to replace the machine. In the last phase, depending on our decision, we should focus on the parts we will not use. They represent frozen capital, which will be very difficult to withdraw. The total cost of decommissioning the machine may be increased by the cost of decommissioning the spare parts in inventory.

Texeira et. al [8] suggest one of approach in classification. In first step they define a criticality of spare parts. The main object is to assign to each spare part a level of criticality using the designation VED:

- Vital: Part failure have great impact on production processes,
- Essential: Part failure have middle impact on production processes,
- Desirable: Part failure pose no risk to the production processes.

In second step criticality is combined with spare parts price and lead time. Also they recommend to create the stock management policy two aspects: the number of machines for which the spare parts are used and the demand for spare parts.

Another approach use Mixed-Integer Linear Programming (MILP) with algorithm. Afifi et. al [1] expand their first model called “Periodic Preventive Maintenance’’. They integrate into PPM model the spare parts inventory policy by proposing a new model as ‘‘the Periodic Preventive Maintenance and the Spare Parts Inventory Problem’’ (PPMSPIP). The proposed model synchronizes both decisions:

- periodic maintenance scheduling,
- spare parts inventory management.

Next step they created hybrid memetic algorithm (HMA) for PPMSPIP. They tested their algorithm on 5 machines. Their Hybrid Memetic Algorithm (HMA) and PPMSPIP were coded in C++ and was run sequentially with the MILP solver.

2.2 Knapsack problem in operational research

The Knapsack problem can be use in resource allocation, cargo loading, project selection, investment decision, and other fields. Li et. al [4] said that this is a typical NP-hard combinatorial optimization problem. He et al [3] noticed many extensions of 0-1 knapsack problem. They listed multidimensional knapsack problem, multiple-choice knapsack problem, multiple knapsack problem, quadratic knapsack problem, set-union knapsack problem, max–min knapsack problem, discounted {0-1} knapsack problems, knapsack problem with setup (KPS) and knapsack problem with single continuous variable. Zhou et al. [9] formulated the goal of this problem as to maximize the total value of the items in the knapsack, while constraints ensure that the sum of the weights is less than or equal to the knapsack's capacity.

There are many algorithms which are based on knapsack problem. Li et. al. [4] compare the Binary quantumbehaved particle swarm optimization (BQPSO) algorithm with quantum-behaved particle swarm optimization algorithm (QPSO) and expand for another algorithms.

He et al [3] said that the evolutionary algorithm (EA), such as genetic algorithm (GA), particle swarm optimization (PSO), differential evolution (DE), ant colony optimization (ACO) and group theory based optimization algorithm (GTOA) and so on, are a kind of important metaheuristics, which have the advantages of simple parameters, strong applicability and generality, and easy to implement.

3 Knapsack problem – case study

3.1 Purchasing spare parts problem – case study

Purpose of formulation new version of Knapsack problem is decision of purchase spare parts to avoid high losses due the breakdowns. According to original version of Knapsack problem prepared Purchasing Spare Parts Problem (PSPP) model. In this case machine's supplier sent the offer with 50 items. All of them cost more than 1,2 million EUR. Amount of our budget for this case is 400 thousand EUR.

We have a set of spare parts described by several parameters:

1. Sets:
   a) i – spare part, $i \in S$;
2. Parameters:
   a) c_i – price of spare part *i;*
   b) s_i – losses if we have a spare part *i* during the breakdown;
   c) l_i – losses if we do not have a spare part *i* during the breakdown;
   d) p_i – probability of breakdown spare part *i*;
3. Decision variable
   a) x_i= 1, if we buy spare part *i*, 0 otherwise

Based on the above parameters and set, the binary programming model of problem is given below:

$$\min losses = \sum_{i \in S} s_i * x_i * p_i + l_i * (1 - x_i) * p_i \qquad (1)$$

s.t:

$$\sum_{i \in S} c_i * x_i \le b \qquad (2)$$

Parameters s_i and l_i include price of item *i*. but rest of their components are different. Parameter s_i describes the value of losses when parts are available in stock. The value of the parameter includes, for example, the cost of spare parts, the cost of storage and the cost of repair. Parameter l_i ignores the cost of warehousing, but can include the value of lost sales and fines. Other costs such as labor of employees or scrap are proportional to lead time of parts. The parameter p_i represents the probability of component failure.

3.2 Results and discussion

The PSPP model was prepared in AMPL and solved by GLPK Integer Optimizer v4.43. The computational experiments were prepared and on a Lenovo IdeaPad computer with an 11th Gen Intel Core i5 processor and 16GB RAM. Model has been successfully generated and processed. Integer optimal solution was found in 0,0 second. During process 0,2 Mb conducted of memory was used.

In optimal solution we can buy 12 items. Amount of objective function is more than 67 thousand EUR. The number of spare parts purchased is determined by the size of the budget. Model does not apply if the total value of the offer will be within the budget.

The main step in this model is to determine what costs will be incurred when the replacement part is in our possession or not. the probability of component failure can be estimated in cooperation with the manufacturer of the machine.

4 Conclusion

According to Li et al. [4] and Zhou et al. [9] knapsack problem can be used to solve any problem. This paper shows using it in spare parts management. Furthermore the PSPP model can be expand by add some new constrains for e.g. exclusion one item by another. We can set parameters according to spare parts classification. Hu et. al [2] noticed that the classification of spare parts permits the identification of the most important spare parts. Then managers can use different inventory strategies for different classes of spare parts, and prioritize the most important items in spare parts management.

Binary decision variable in this problem gives only “yes or no” decision. We do not know how many one part we should buy. This model can be used, when we buy new machine and we decide which parts should be on stock.

Texeira et.al. [8] explained why spare parts management is important for maintain production process. The decision to purchase a spare part should be made at the initial stage of build spare parts stock. Next stages may include models based on economic value of order, reorder points and so on. But there is a risk, when we decide to remove the machine.

According to Hu et. al [2] this paper is about first phase in life cycle machine. As they said in paper when new equipment is introduced, a spare parts replenishment system needs to be established to effectively provide spares supply support.

Maintenance management in any of the approaches whether Time Based Maintenance or Condition Based Maintenance or others includes spare parts management. While we can separate maintenance work into internal and external, we use external sources for spare parts. Proper spare parts management will yield results on the value of assets and the financial result of the company.

References

[1] Afifi Sohaib, Hrouga Mustapha, Mjirda Anis, Allaoui Hamid. *A memetic based algorithm for simultaneous preventive maintenance scheduling and spare-parts inventory management for manufacturing systems*. Applied Soft Computing Journal 151 (2024) 111161

[2] Hu Qiwei, Boylan John E., Chen Huijing, Labib Ashraf. *OR in spare parts management: A review*. European Journal of Operational Research 266 (2018) 395–414

[3] He Yichao, Wang Jinghong, Liu Xuejing, Xizhao Wang, Haibin Ouyang. *Modeling and solving of knapsack problem with setup based on evolutionary algorithm*. Mathematics and Computers in Simulation 219 (2024) 378–403

[4] Li Xiaotong, Fang Wei, Zhu Shuwei. *An improved binary quantum-behaved particle swarm optimization algorithm for knapsack problems*. Information Sciences 648 (2023) 119529

[5] Ramaganesh M., Vikas Varma Ganapathiraju, Rajpradeesh T., Bathrinath S. *Spare parts ordering decisions using age based, block based and condition based replacement policies*. Materials Today: Proceedings 46 (2021) 7854–7859

[6] Rinaldi Marta, Fera Marcello, Macchiarolia Roberto, Bottani Eleonora. *A new procedure for spare parts inventory management in ETO production: a case study*. Procedia Computer Science 217 (2023) 376–385
[7] Santos T, Silva F. J. G., Ramos S. F., Campilho R. D. S. G., Ferreira L. P. *Asset Priority Setting for Maintenance Management in the Food Industry*. Procedia Manufacturing 38 (2019) 1623–1633
[8] Teixeira Catarina, Lopes Isabel, Figueiredo Manuel. *Multi-criteria classification for spare parts management: a case study*. Procedia Manufacturing 11 (2017) 1560 – 1567
[9] Zhou Yongquan, Shi Yan, Wei Yuanfei, Luo Qifang, Tang Zhonghua. *Nature-inspired algorithms for 0-1 knapsack problem: A survey*. Neurocomputing 554 (2023) 126630

Improvement of Methods of Fertility Rates Modelling

Ondřej Šimpach[1]

Abstract. Fluctuations and trends of fertility rates development do not have to be regular or long-term. Knowledge of fertility rates is needed for policy planning and public administration.
Therefore, we focus on the modelling of the fertility rates in the Czech Republic. Particularly, we apply standard Lee-Carter model with time-independent parameters $\mathbf{a}_x$ and $\mathbf{b}_x$ and time-varying index $\mathbf{k}_t$. Hyndman "demography" package in RStudio software is utilized. Because parameter $\mathbf{a}_x$ (the average value of the empirical time series) can be biased, we suggest an approach for its improvement.
We utilize functional data (fertility rate, age, year), where fertility rate is a function of women's age for time period 1950 to 2022. There are 4 models with different $\mathbf{a}_x$ parameter compared: a) standard parameter of the Lee-Carter model, b) median of age-specific fertility rates, c) $\mathbf{a}_x$ calculated on data 1999–2022, and d) $\mathbf{a}_x$ calculated on data 2008–2022. However, no approach was found to be better than the original calculated as simple arithmetic means of fertility rates in specific age which had the lowest mean squared error.
These results are important for subsequent analyses because for working with demographic data about fertility it is important to consider the most recent data, which are not significantly skewed and influenced by a range of factors.

Keywords: fertility rates, Lee Carter model, policy planning, stochastic modelling

JEL Classification: J11, J13
AMS Classification: 91D20

1 Introduction

The number of live births as a positive item in the population balance, a natural currency, is always in the center of interest of all users of demographic data. Related demographic indicator are the fertility rates by age (f_x) – the number of children per woman in specific age calculated as the ratio of the number of live births to women at age x (or in a given five- or ten-year age group) to the average state of women at age x (age group). The changes in fertility depending on the age of mothers are being studied. Total fertility is then related to this indicator. It is the sum of age-specific fertility rates expressing the fertility intensity of a given population in a given time period (calendar year). It indicates the number of children that would be born to one woman during the reproductive period, if the values of the fertility rate did not change. While the reproductive period of a woman is considered to be in the age range of 15–49 years. It is therefore the number of live births per woman aged 15–49 years. A value of 2.08 ensures that the numerical state of the population is maintained. (Langhamrová, Šimpach, 2013 and Šimpach, 2015)

The behavior of the population has changed in terms of family planning. The birth of the first child is being continuously postponed to later ages. There are also fewer children per woman during the reproductive period (15–49 years), i.e. the fertility rates are lower. Besides, total natality is also declining. In 2023, there was born the fewest number of children in the last 20 years in the Czech Republic. Fluctuations and trends of fertility rates development do not have to be regular or long-term. Knowledge of fertility rates is needed for policy planning and public administration.

1.1 Fertility rates development

From Figure 1 that displays total fertility rates in 1950–2022 can be seen that values that ensured stable population were achieved only until year 1965. Then the total fertility rates were lower than 2.08 until year 1973. It increased again later thanks to pro-natal policies, but only until the year 1980. This was the last year when the total fertility rates ensured stable state of the population. A significant decrease in the birth rate happened at the beginning of the nineties in the Czech Republic. The fertility rate decreased to even 1.13 in 1999. The change of regime caused uncertainty at the beginning, but the total fertility rates started to increase slowly again later. They are still low and

[1] Prague University of Economics and Business, Faculty of Informatics and Statistics, Department of Statistics and Probability, W. Churchill sq. 4, 130 67 Prague 3, Czech Republic, ondrej.simpach@vse.cz.

very probably will never be higher than 2.08. An interesting peak was noted in 2021 in Covid-19 pandemic year, when the total fertility rate increased suddenly on 1.83. Finally, the year 2022 experienced a decrease which could be expected also in 2023. The 2023 generation was the least numerous in the last 22-year-period. The number of live births decreased year-on-year by 10.2 thousand, or 10%, to 91.1 thousand. A similar decline could be expected for the total fertility rate. (CZSO, 2024).

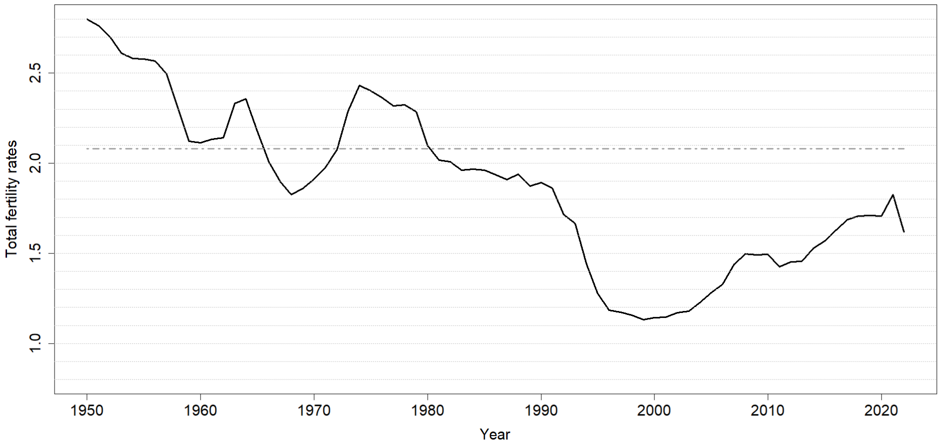


Figure 1 Total fertility rates in 1950–2022; *Source: CZSO*

In the early 90s of the last century, the highest fertility rate was for the age group of 20–24-year-old women, and by the beginning of the 21st century, it had shifted to the older ages of 25–29. The average age of a mother at the birth of a child between 2013 and 2022 increased by 0.6 years from 29.9 years to 30.4 years. Throughout the last decade, within five-year age groups, the highest fertility intensity was always among 30–34-year-old women, which, however, only slightly exceeded the fertility level of women aged 25–29. Thus, the trend of shifting fertility to the older ages, which began in the 1990s, continued.

Fertility modelling can be done by deterministic or stochastic methods. Deterministic approach is for example by cohort-component method based prepared scenarios – usually basic, optimistic and pessimistic variant. Stochastic methods take into account the random nature of events. Our paper focuses on modern stochastic approach firstly elaborated by Lee and Carter (1992).

2 Data and Methods

The aim of the paper is to suggest certain possibilities how to improve the modelling of the fertility rates. We utilized data about age-specific fertility rates in particular year (functional data), where fertility rate is a function of women's age, and included them into standard original Lee-Carter model. We aim on enhancement the model's projection abilities and focus on the improvement of the parameter $\mathbf{a}_x$ which is the average value of the empirical time series of fertility rates. Particularly, we compare the mean squared forecast error (MSE) of four different models. The calculation of MSE is applied on back-projected data.

2.1 Data

Empirical data about total fertility rates and age-specific fertility rates were taken from Czech Statistical Office (CZSO, 2023) for years 1950–2022. The CZSO observes the data about births only for females from 15 years to 45 years. The children born to younger women are added to the category of 15 years old and the children born to women over 45 years are added to the category of 45 years old. Another possible data source – Human fertility database – incorporates the data from 12 to 55 years, but there is not up-to-date data (only until year 2021).

The development of the total fertility rates has been already displayed in Figure 1. Here we visualize the age-specific fertility rates in the rainbow chart in Figure 2. The graph on the left displays the changes of *fx,t* (empirical values of *fx,t* of the Czech females for the period 1925–2022) over time. The graph on the right depicts the changes of *fx,t* within the age groups.

It is possible to see that the number of live births per 1000 *x*-year old females has changed significantly during the examined period 1950 to 2022. The clear trend of postponing childbirth to a later age can be seen from Figure 2 on the left. The curves were originally skewed to the right but changed to an almost normal curve later with a mean value around 30 years.

In the case of fertility projections, it makes sense to use only the current and stable database, which is not affected by past influences. (see Šimpach, Langhamrová, 2014) Otherwise, there might be bias in the projected values. However, we are not projecting the fertility rates to the future, but only backwards. Therefore, it is possible to use the whole time series from 1950 until the latest available data 2022.

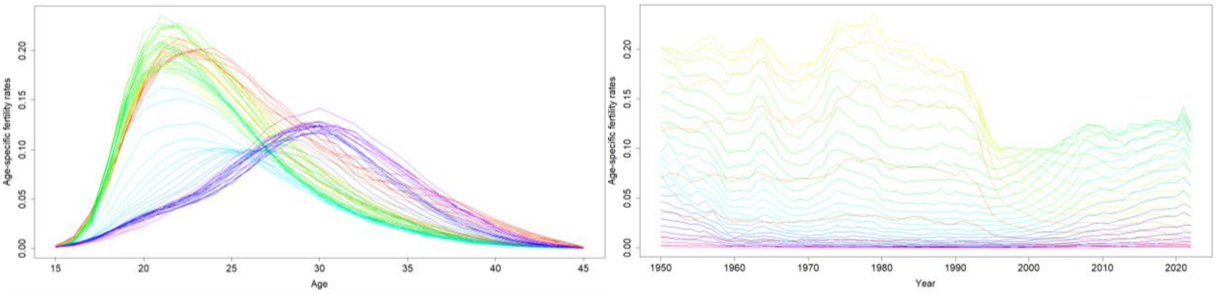


Figure 2 Empirical values of age-specific fertility rates $f_{x,t}$ in 1950–2022; changes of $f_{x,t}$ over time (left); changes of $f_{x,t}$ within age groups (right). Source: CZSO, own calculations and illustration based on Hyndman's (2022) package `demography` in R.

2.2 Model

The age-specific fertility rates $f_{x,t}$ at age x and time t create (in our case) 31 × 73 dimensional matrix **F**. The standard Lee-Carter model is applied on the $f_{x,t}$ logarithms. Consequently, certain modifications are done to the model and the results are compared.

Lee-Carter model (Lee and Carter, 1992) is one of the most used models describing the age and year structure of the life table. Its main feature is the stochasticity of the processes. It uses the trends and main components of previous development of demographic events – usually fertility or mortality rates. The model is simple because it has a limited number of the parameters and the parameters can be computed using amenable techniques such as the singular value decomposition. (Koissi, Shapiro and Hognas, 2006). The approach is based on the principal component analysis and decompose empirical values of age-specific fertility rates (in logarithms) as (1) (see e.g. Lee and Carter (1992), Lee and Tuljapurkar (1994) or Shang and Hyndman (2010).

$$f_{x,t} = a_x + b_x k_t + \varepsilon_{x,t} \tag{1}$$

where x are 15–45 years of life, t is time period t = 1,2, ..., T (T = 2022). Vectors of parameters $\mathbf{a}_x$ and $\mathbf{b}_x$ are time independent. Age specific component $\mathbf{a}_x$ is calculated as simple arithmetic mean of age-specific fertility rates at each age (2).

$$a_{x,t} = \frac{\sum_{t=1}^{T} f_{x,t}}{T} \tag{2}$$

There are 4 models with different $\mathbf{a}_x$ parameter compared in the paper: a) standard $\mathbf{a}_x$ parameter of the Lee-Carter model b) median of age-specific fertility rates, c) $\mathbf{a}_x$ calculated as standard arithmetic mean, but on data 1999–2022 (t = 50, … 73) and d) $\mathbf{a}_x$ calculated as standard arithmetic mean, but on data 2008–2022 (t = 59, … 73). The modifications are done in order to improve the abilities of the model. Shorter time-series might be better, because the fertility pattern has changed in the past significantly. Therefore, only the fertility rates from 1999 were included into the calculation of arithmetic mean, because there was the lowest total fertility rate in this year and started to increase since that. Similarly, there was a local maximum of fertility in 2008 and the rates are stabilized since that. Besides, also median instead of arithmetic mean was also calculated and included into the mode.

Vector of parameter $\mathbf{b}_x$ is the additional age-specific components that determine how much each age group changes when parameter $\mathbf{k}_t$ changes. Finally, $\mathbf{k}_t$ are the time-varying indexes – the total fertility indices. $\varepsilon_{x,t}$ is the error term with characteristics of white noise. In the standard LC model (which uses SVD), the error terms are assumed Gaussian with constant variance, e.g. a homoscedasticity is assumed, while it can be violated in many applications of the model. (Koissi, Shapiro and Hognas, 2006).

The estimation of $\mathbf{b}_x$ and $\mathbf{k}_t$ is based on Singular Value Decomposition (SVD) of matrix of age-specific fertility rates (Bell and Monsell, 1991 or Hyndman and Booth, 2008). The model in (1) is overparametrized, so the identification of the model is ensured by (3).

$$\sum_{x=15}^{45} b_x = 1 \quad \text{and} \quad \sum_{t=1}^{73} k_t = 0 \tag{3}$$

For predicting the future age-specific fertility rates it is necessary to forecast the values of $\mathbf{k}_t$ index. However, for the back-projection, it is not necessary. The main aim is to compare the MSE of the four models on the back-projected data. Therefore, the fertility rates are calculated by four approaches and the results are compared by MSE (4):

$$\frac{\sum_{x=15}^{45}\sum_{t=1}^{73}\left(f_{x,t}-f_{x,t}^{'}\right)^2}{45\times 73} \tag{4}$$

where $f_{x,t}'$ are fitted values of fertility rates projected backwards by four different approaches.

Hyndman `demography` package in RStudio software is utilized for the estimation of the standard Lee-Carter model. This and other standard packages in RStudio are utilized for creation of the figures.

Results

First, we estimated standard parameters $\mathbf{a}_x$ (age-specific fertility profiles independent of time), $\mathbf{b}_x$ (additional age-specific components that determines how much each age group changes when $\mathbf{k}_t$ changes) and $\mathbf{k}_t$ (time-varying fertility index) for standard Lee-Carter model based on data matrix from 1950 to 2022 using the SVD method. The estimation was done by command `lca()` in the R package `demography` (Hyndman, 2022). Parameter $\mathbf{a}_x$ will be discussed later. Hence, we start with presentation of the parameter $\mathbf{b}_x$ that captures the fluctuation of the fertility rates in particular age. It is displayed at Figure 3 on the left. It decreases up to age of 19 years and increases in the direction to higher ages. The non-typical break in ages 44 and 45 is due to the fact that the fertility rates in those high ages are in general low and if a significantly different value appears in any year, a bias may occur. Fertility index $\mathbf{k}_t$ is displayed also at Figure 3, but on the right. It somehow copies the development of the total fertility index, but its lowest value is 10 years earlier. It is more volatile in its decreasing part than in increasing part. The development of both can be compared e. g. to Šimpach (2015): parameter $\mathbf{b}_x$ is decreasing again in this study and $\mathbf{k}_t$ index is increasing in earlier years. This is due to different time series (since 1998) and the highest age category (49 years) used in this study.

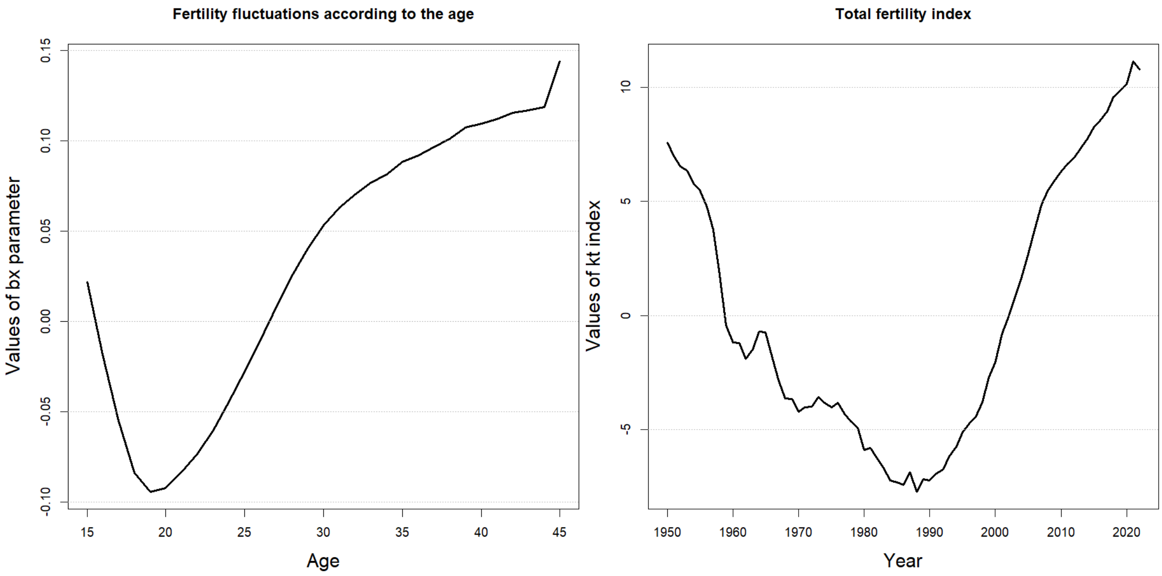


Figure 3 Values of estimated parameter $\mathbf{b}_x$ (left) and index $\mathbf{k}_t$ (right) of Lee-Carter model. Source: own calculations and illustration based on CZSO (2024) data

Let us focus on parameter $\mathbf{a}_x$. Its concave development is comparable to Šimpach (2015). All 4 variants (in natural logarithms) are displayed at Figure 4. It can be seen that the maximum of standard $\mathbf{a}_x$ is in lower ages than in case when shorter time series with fertility postponed to higher ages are used. Hence, those parameters calculated from shorter time-series (newer data) reflect the current situation in the Czech Republic better. This can be an advantage for future projection. Let us see whether the models with those parameters $\mathbf{a}_x$ are also better in backward projections.

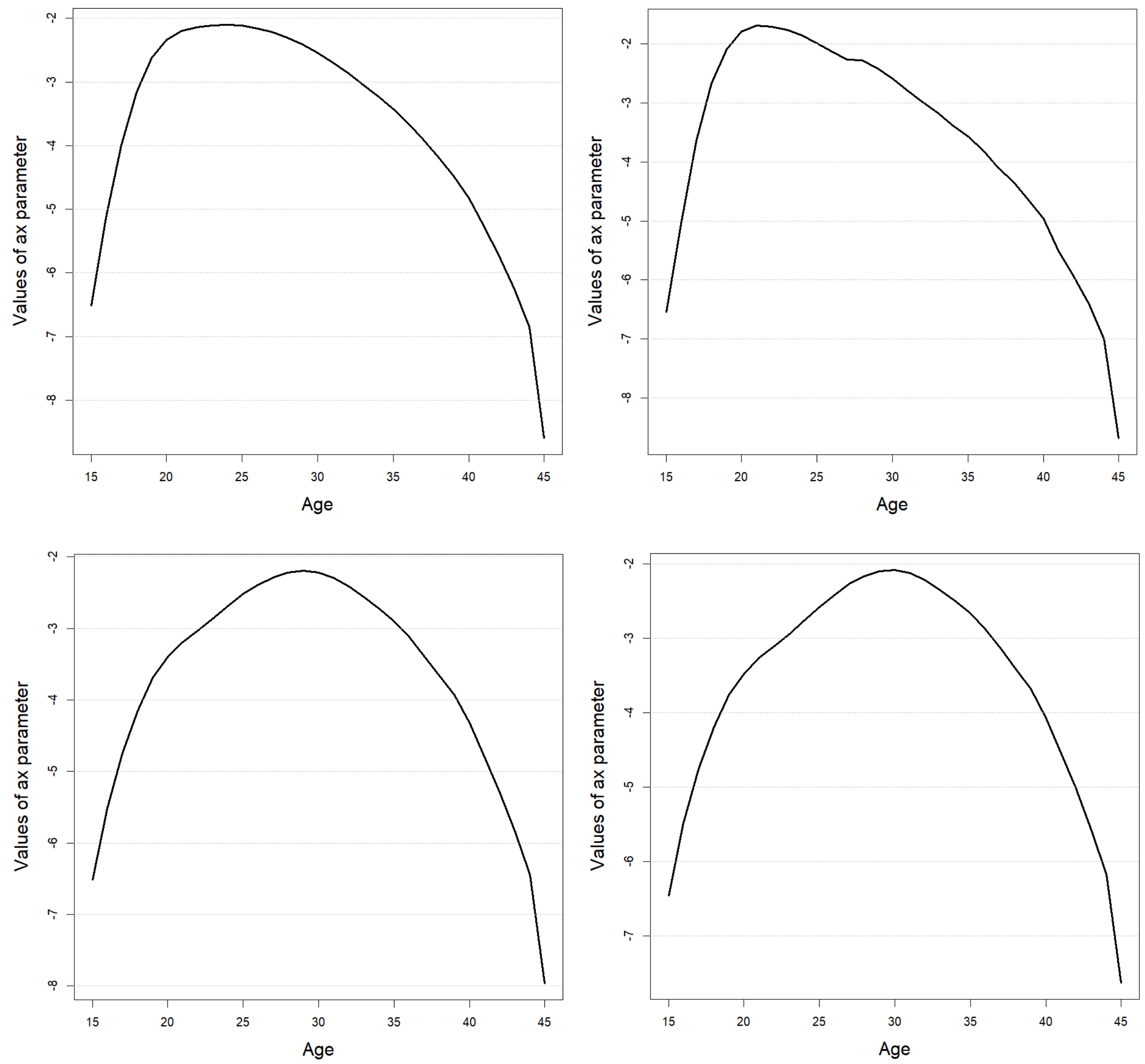


Figure 4 Values of estimated parameter $\mathbf{a}_x$ of Lee-Carter model in logs. – standard $\mathbf{a}_x$ (upper left), $\mathbf{a}_x$ calculated as median (upper right), $\mathbf{a}_x$ calculated on data 1998–2022 (lower left), $\mathbf{a}_x$ calculated on data 2008–2022 (lower right). Source: own calculations and illustration based on CZSO (2024) data

All four parameters $\mathbf{a}_x$ were included in the calculation of four models of fertility rates. The empirical fertility rates and the results of the models are displayed in Figure 5. At first glance, it can be seen that while the differences between the various parameters $\mathbf{a}_x$ are relatively small, the differences in the predicted birth rates are huge. The projections are diametrically opposed and at first glance unusable.

Using the median of specific fertility rates has proven to be inappropriate because it predicts highly skewed values. Especially in the 1970s, fertility rates are unrealistically high. In these years, due to political measures, the birth rate was high, but not nearly as much as this model predicts. For this reason, it cannot be used.

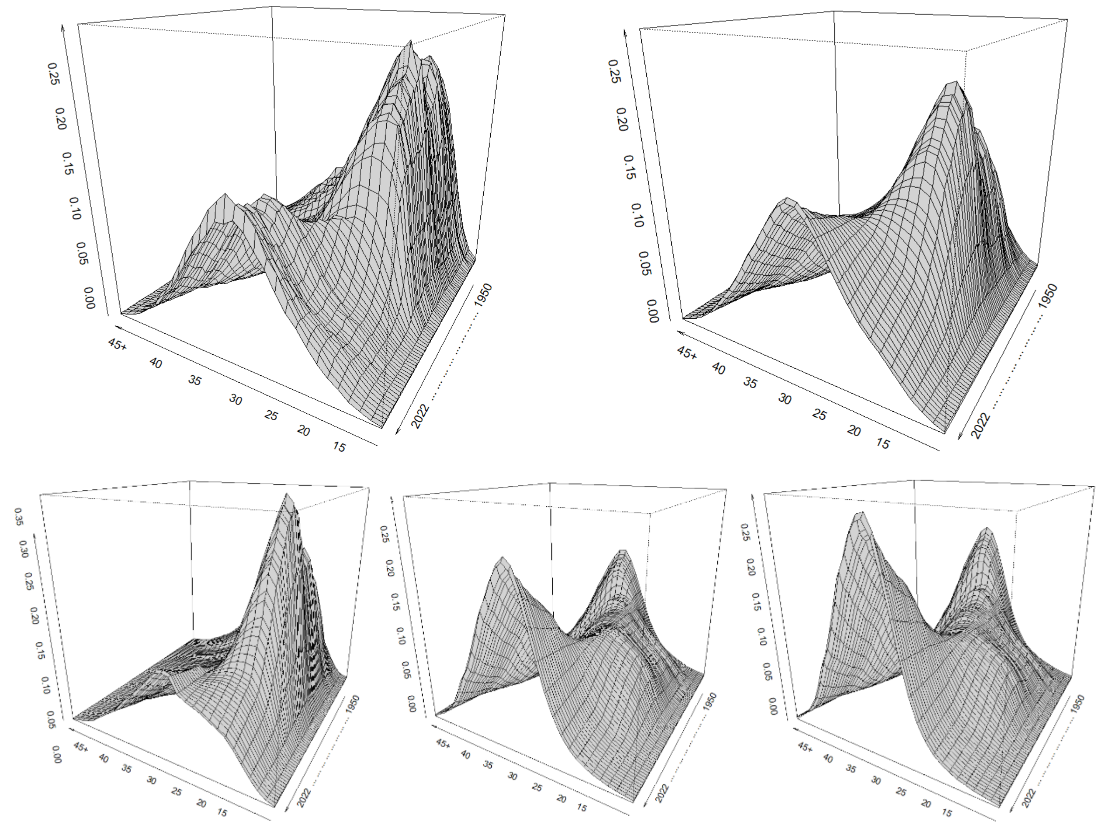


Figure 5 Age specific fertility rates: empirical $f_{x,t}$ (upper left), $f_{x,t}$ fitted by standard LC model (upper right), $f_{x,t}$ with $\mathbf{a}_x$ as median (lower left), $f_{x,t}$ with $\mathbf{a}_x$ calculated on data from 1999 (lower middle), $f_{x,t}$ with $\mathbf{a}_x$ calculated on data from 2008 (lower right). Source: own calculations and illustration based on CZSO (2024) data

When we compare the MSE of the models among each other, we can clearly see that standard calculation of the parameters $\mathbf{a}_x$ as an arithmetic means is the best because the error is the lowest by 10 times.

$\mathbf{a}_x$	**MSE**
Standard	0.000686
Median	0.001436
Period 1999–2022	0.002392
Period 2008–2022	0.002986

Table 1 Mean Squared Error (MSE) of fertility rates estimated by different approaches. Source: own calculations and illustration based on CZSO (2024) data

Conclusion

The aim of the paper was to find ways to improve fertility rates forecasting. Particularly we focused on the stochastic Lee-Carter model and compared four different parameters $\mathbf{a}_x$ (age-specific fertility profiles independent of time). There are 4 models with different $\mathbf{a}_x$ parameters compared: a) standard parameter of the Lee-Carter model b) median of age-specific fertility rates, c) $\mathbf{a}_x$ calculated on data 1999–2022 and d) $\mathbf{a}_x$ calculated on data 2008–2022. However, no approach was found to be better than the original calculated as simple arithmetic mean of fertility rates in specific age. This approach ensured the lowest mean squared error, i.e. the lowest average squared differences between empirical and estimated fertility rates.

The maximum of standard $\mathbf{a}_x$ parameter was in lower age than in case when shorter time series with fertility postponed to higher ages are used. Hence, those parameters calculated from shorter time-series (newer data) reflect the current situation in the Czech Republic better. This feature can be utilized for fertility rates projections into the future. The challenge for the consequent research is to utilize different $\mathbf{a}_x$ parameters also for the future projections of the fertility rates and compare the models.

Acknowledgements

This paper was supported by long term institutional support of research activities by Faculty of Informatics and Statistics, University of Economics, Prague.

References

Bell W R, Monsell B (1991) Using principal components in time series modelling and forecasting of age-specific mortality rates. Proceedings of the American Statistical Association, Social Statistics Section, pp. 154–159

CZSO (2023) Czech Demographic Handbook – 2022. https://www.czso.cz/csu/czso/czech-demographic-handbook-2022

CZSO (2024) Population change - year 2023. Population of the Czech Republic exceeded 10.9 million. https://www.czso.cz/csu/czso/ari/population-change-year-2023

Hyndman R J, Booth H (2008) Stochastic population forecasts using functional data models for mortality, fertility and migration. International Journal of Forecasting 24:323–342

Hyndman R J (2022) demography: Forecasting mortality, fertility, migration and population data. R package v. 2.0, https://cran.r-project.org/web/packages/demography/demography.pdf

Koissi M-C, Shapiro A F, Hognas G (2006) Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval. Insurance: Mathematics and Economics 38(1):1–20

Langhamrová J, Šimpach O (2013) Základy demografie: (materiály ke cvičením). Vysoká škola ekonomická v Praze. Praha: Oeconomica.

Lee R D, Carter L R (1992) Modeling and forecasting U.S. mortality. Journal of the American Statistical Association 87:659–675.

Lee R D, Tuljapurkar S (1994) Stochastic population forecasts for the United States: beyond high, medium, and low. Journal of the American Statistical Association 89:1175–1189

Shang H L, Hyndman, R J (2010) Exploratory graphics for functional data. Working paper of the Department of Econometrics and Business Statistics, Monash University, Clayton, Australia, Aug. 3, 2010, pp. 1–9.

Šimpach O, Langhamrová J (2014) Stochastic Modelling of Age-specific Mortality Rates for Demographic Projections: Two Different Approaches. In: Mathematical Methods in Economics 2014. Conference Proceedings. Palacký University in Olomouc, Olomouc, pp. 890–895

Šimpach O (2015) Application of the Modern Approach to Age-specific Fertility Rates Stochastic Modelling in the Czech Republic. In: Mathematical Methods in Economics 2015. Conference Proceedings. University of Western Bohemia in Plzeň, Plzeň, pp.799–804

Exploring Fuzzy Functional Dependencies in Assessing Relationships between Criteria in EU ETS Impacts

Markéta Šindlerová[1], Miroslav Hudec[2], [3]

Abstract. Emissions Trading Systems (ETS) have emerged as a supporting policy tool for tackling climate change by putting a price on carbon emissions. However, understanding the complex relations between environmental and economic criteria remains a challenge. This study explores the possibilities of fuzzy functional dependencies as a method to model the relationships between various criteria involved in assessing the effectiveness and efficiency of emissions trading schemes and the economic impacts of emissions trading within the EU ETS framework. By employing fuzzy functional dependencies, the analysis captures the inherent uncertainty and vagueness in the relationships between environmental and economic variables and explains directions of dependencies between attributes within the EU ETS context. Finally, the article raises topics for future work.

Keywords: Conformance Measure, Fuzzy Functional Dependencies, EU ETS (Emissions Trading System), Interactions between Economic and Environmental Criteria

JEL Classification: C44, Q52
AMS Classification: 62A86

1 Introduction

Emissions trading systems (ETS) have emerged as a significant policy tool for tackling climate change by putting a price on carbon emissions. The European Union Emissions Trading System (EU ETS), came into operation 2005 [8], is one of the most significant and longest-running schemes aimed at reducing greenhouse gas emissions by setting a cap on the total amount of certain greenhouse gases that can be emitted by installations covered by the system. Companies receive or buy emission allowances, which they can trade with one another as needed. Understanding the complex relationships between environmental and economic criteria within the EU ETS remains a significant challenge. Traditional methods often fail to capture the nuances and uncertainties inherent in these relationships. This article aims to extend traditional correlation analysis by introducing fuzzy functional dependencies into multi-criteria analysis, which evaluates companies within the EU from both environmental and economic perspectives.

Fuzzy logic deals with the reasoning that is approximate rather than fixed and exact. It provides a powerful tool for modeling the uncertainty and vagueness present in real-world data. By employing fuzzy functional dependencies, this study explores and models the complex relationships between various criteria involved in assessing the effectiveness and efficiency of emissions trading schemes. This approach will help uncover underlying patterns and dependencies that may be hidden by traditional statistical methods.

Moreover, fuzzy functional dependencies not only allow us to determine the strength of the relationships but also their direction [12]. Traditional correlation analysis can reveal the strength of dependency between criteria, but with fuzzy extensions, we can discern the direction of the dependency. This distinction is crucial because the influence of one criterion on another does not necessarily imply that the reverse influence will be of the same magnitude. For example, if environmental performance influences economic outcomes, it does not automatically mean that economic performance will equally influence environmental outcomes.

Our focus is on the paper industry, a sector significantly affected by emissions trading due to its high energy consumption and substantial carbon footprint [2, 7]. By examining the environmental and economic impacts of the EU ETS on companies in the paper industry, we aim to provide a more comprehensive understanding of how emissions trading affects this specific sector. This study thus contributes to the broader field of environmental economics by demonstrating the applicability of fuzzy functional dependencies in multi-criteria decision analysis (MCDA). The

[1] VŠB-Technical university of Ostrava, Faculty of Economics, Department of Systems Engineering, 17. listopadu 15, Ostrava, 70800, Czech Republic, marketa.sindlerova@vsb.cz
[2] University of Economics in Bratislava, Faculty of Economic Informatics, Department of Applied Informatics, Dolnozemská cesta 1, Bratislava, 85235, Slovakia, miroslav.hudec@vsb.cz
[3] VŠB-Technical university of Ostrava, Faculty of Economics, Department of Systems Engineering, 17. listopadu 15, Ostrava, 70800, Czech Republic, miroslav.hudec@vsb.cz

findings will be valuable for policy-makers, business leaders, and researchers seeking to balance environmental responsibilities with economic performance, thereby supporting the development of more sustainable and efficient emissions trading policies.

Based on the identified relationships, it is possible to adjust the weights of criteria in future multi-criteria decision-making processes [1]. This ensures that decision models remain dynamic and responsive to the evolving interplay between environmental and economic factors. By re-calibrating these weights, decision-makers can enhance the accuracy and relevance of their assessments, ultimately leading to more informed and effective policy and business strategies.

The reminder of this contribution is organized as follows. Section 2 introduces fuzzy functional dependencies. Section 3 is devoted to the EU ETS data and analysis, whereas Section 4 discussed the results and perspectives for adjusting weights. Finally, Section 5 concludes this work.

2 Methodology

In this section, we describe the methodology adopted to analyze the relationships between criteria in the paper industry within the EU ETS framework.

2.1 Theoretical Background of Fuzzy Functional Dependencies

Fuzzy Functional Dependencies (FFDs) have been recognized as an approach for examining flexible dependencies between two or more attributes [13]. They extend classical functional dependencies to handle cases when two entities (e.g., companies) share similar values of attribute A and attribute B, e.g, production and pollution (see e.g., [6, 11, 12]).

If entities e_1 and e_2, share a similar value within the domain of the attribute A (or criterion), they might also share similar value within the domain of the attribute B. FFD is denoted as $A \xrightarrow{\theta} B$ where θ represents the degree of strength of functional dependence. This degree can range from complete dependence to no dependence at all. FFDs are particularly useful in scenarios where data are uncertain, imprecise or incomplete, such as fuzzy numbers or belonging to the multiple neighboring values from the discrete domains (i.e.{*little, medium*}) in a case of Likert scale surveys or experts evaluation of tacit knowledge.

In this paper, we firstly focus on the conformance measure [11] adopted to the similarity between attributes, in our case economic and environmental criteria in the paper industry. Conformance measure computes the degree of similarity between values of attribute A considering two entities as follows

$$C(A[e_i, e_j] = \min \left\{ \min_{x \in d_i}\{\max_{y \in d_j}\{s(x,y)\}\}, \min_{x \in d_j}\{\max_{y \in d_i}\{s(x,y)\}\} \right\} \tag{1}$$

where d_i represents the values (subset of attribute domain) for entity e_i, d_j represents the values of A for entity e_j, where $d_i \in D_i$ and $d_j \in D_j$ (e.g., linguistic terms from linguistic variables created over respective domains), and $s(x, y)$ denotes a proximity relation between the values x and y within the respective domains D [5, 10].

The next step is recognising in which proportions of all pairs of entities e_i and e_j the fuzzy functional dependence is relevant. We calculate for which pairs it holds as [12]

$$C(A[e_i, e_j]) \geq \min(\theta, C(B[e_i, e_j]) \tag{2}$$

and finally, we compute the proportion of pairs of criteria where fuzzy functional dependencies are satisfied. Details can be found in e.g., [12].

2.2 Defining Linguistic Variables and Fuzzy Sets

Each criterion (Profit, EBITDA, Firm Size, Allocated Allowances, Pollution, and ROCE) was categorized (discretized) into linguistic variable [14] consisting of linguistic terms formalized by fuzzy sets. For example, pollution levels are categorized into fuzzy sets *Low*, *Medium*, and *High*. Similarly, the other attributes are categorized. It is possible for one value to partially belongs to two neighboring fuzzy sets. It holds, for the uncertainty area, i.e., areas where neighboring values overlap (marked as α in Fig. 1). In this way, inherent vagueness and non clear boundaries between terms are included in model. Contrary, by the traditional intervals very similar values might fall into the different intervals and therefore treated differently, even though the difference is very small, or it is caused by the measurement error.

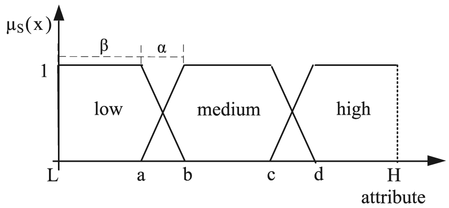


Figure 1 Linguistics variable for an attribute

In Table 1, we see an example of created fuzzy sets for the ROCE criterion. The values within each cell denote the corresponding ROCE values associated with each linguistic term. If a particular variant falls within the range between two fuzzy sets, it belongs simultaneously to both sets. In Table 1, L, a, b, c, d, and H represent the boundaries of the fuzzy sets cores (clear belonging to a fuzzy set, marked as β) and uncertain area (marked as α) for the ROCE criterion.

ROCE	**L**	**a**	**b**	**c**	**d**	**H**
Low (S)	-100	-25	-15			
Medium (M)		-25	-15	15	25	
High (L)				15	25	100

Table 1 Parameters for fuzzy sets of criterion ROCE, where L, a, b, c, d, H corresponds with Fig 1

.

For example, if the ROCE is 30, it falls clearly into the *High* category, representing a case with no overlap. Additionally, a value can belong to two linguistic terms simultaneously if it cannot be assigned to just one core with a membership degree of 1. Thus, if the ROCE of a company is -20, it would partially fall into the *Low* category and partially into the *Medium* category. Our approach can handle this overlap, and it does not hinder subsequent calculations if a criterion belongs to multiple fuzzy sets simultaneously. This method allows us to handle imprecision and uncertainty in the data, making the analysis more robust and realistic.

2.3 Creation of Proximity Relation

The next step involves creating values of proximity relation between linguistic terms. In a general case (irrelevant whether we have ordinal or nominal linguistic terms), this measure is symmetric and reflexive [10]. But for the full consistency of ordinal linguistic terms this measure should meet the properties of monotonicity and min-min transitivity [5].

For example, for *high* and *low* categories, the resulting similarity strength is weak, such as a value of 0.2. Conversely, for *high* and *medium* the proximity is higher, as is in Table 2 This relation serves as the foundation for computing conformance measure and FFD. Because of the reflexivity property, the table is diagonally symmetrical. Thus, the area below the diagonal might remain empty.

proximity measure	Low	Medium	High
Low	1	0.6	0.2
Medium		1	0.4
High			1

Table 2 Proximity Relation

3 Data and Analysis

In this section, we introduce data preparation and the subsequent calculations performed to analyze the influence between criteria in emissions trading on firms in the paper industry.

When selecting criteria, we drew from the concept of a popular Corporate Social Responsibility (CSR) model described in [9]. This model is based on groups of social, economic, and environmental criteria, similar to those presented in models referenced in the review paper [3].

3.1 Data Preparation

For our analysis, we gathered data on six key criteria for firms in the EU paper industry: Profit, EBITDA, Firm Size, Allocated Allowances, Pollution, and ROCE. These criteria were chosen because they encompass both economic and environmental dimensions, which are essential for assessing the impact of emissions trading.

Data are from the year 2021, which provided a snapshot of the current state of this industry within the EU under the emissions trading system (EU ETS). We aimed to include a comprehensive sample of firms. So, we collected data from various sources including financial reports, environmental performance reports, and the EU ETS registry. However, we faced challenges in data availability, which led to the exclusion of some firms from our initial list. Despite this, the final dataset consists of 10 major firms, ensuring that our analysis is still robust and representative of the industry, but also to introduce the new concept of evaluation dependencies in this field.

3.2 Experiments

At the beginning, each criterion was categorized into three fuzzy sets (see Section 2.2: *Low*, *Medium*, *High*. This categorization allows handling uncertainty and imprecision in the data. With the data prepared and categorized, we performed the consequent calculations to understand the relationships between the criteria.

Calculation of Dependency Strengths

For each pair of criteria, we calculated the strength of dependency for each variant (company). This has been realized by comparing the fuzzy sets of each criterion belonged to each firm. For instance, if the pollution level of a company was categorized as *high* and the profit was *low*, we assigned a similarity strength of 0.2 (see Table 2. If a company's values for two criteria fell into the same fuzzy set, the similarity strength was assigned value 1. Variants that belong to multiple fuzzy sets simultaneously had their similarity calculated using equation 1.

Aggregation of Dependencies

After calculating the dependency strengths for each company, we aggregated these results to determine the overall dependency strength for each pair of criteria. This involved summing the similarity scores above threshold value θ (2) and normalizing them by the number of firms. Dependencies were calculated in both directions for each pair of criteria. For example, we calculated how strongly profit depended on pollution and vice versa. This included evaluating the dependencies in both directions for all criteria to provide the full picture of the relationships.

This thorough calculation method allowed us to capture the nuances of the relationships between different criteria, considering both the strength and direction of dependencies. By employing fuzzy functional dependencies, we could shed more light to complex relationships than with traditional correlation analysis.

Calculated Dependencies

Table 3 displays the revealed dependencies between pairs of criteria in both directions by FFD. This brings a significant new information when compared to the traditional correlation analysis. It is important to note that if criterion *C1* is influenced by *C2*, it does not necessarily mean that *C2* is influenced by *C1* with the same strength.

	CO2 Production	Profit	EBITDA	Firm Size	ROCE	Allowance Allocation
CO2 Production	1	0.4	0.4	0.4	0.5	0.2
Profit	0.7	1	1	0.6	0.6	0.8
EBITDA	0.6	0.7	1	0.5	0.5	0.5
Firm Size	0.7	0.4	0.6	1	0.9	0.3
ROCE	0.7	0.3	0.5	0.8	1	0.3
Allowance Allocation	0.7	0.8	0.8	0.4	0.6	1

Table 3 Dependency Matrix of Criteria calculated by FFD

Next, Table 4 shows the differences in dependencies between pairs of criteria based on the direction of influence. Value 0 indicates that the dependency between the given pair of criteria is symmetrical, meaning that the first criterion influences the second one with the same strength as the second criterion influences the first one. The greater the value, the greater the difference in dependency between the criteria depending on the direction of analysis. This approach allows us to understand the nuanced relationships between criteria that are not captured by traditional correlation analysis.

	CO2 Production	Profit	EBITDA	Firm Size	ROCE	Allowance Allocation
CO2 Production	0					
Profit	0.3	0				
EBITDA	0.2	0.3	0			
Firm Size	0.3	0	0.1	0		
ROCE	0.2	0.3	0	0.1	0	
Allowance Allocation	0.5	0	0.3	0.1	0.3	0

Table 4 Dependency Difference Matrix of Criteria

Based on the provided tables, we can interpret the dependencies among different criteria. For example, the Profit criterion has a strong influence on the EBITDA criterion with a coefficient of 1. In traditional correlation analysis, such a strong influence might suggest redundancy between these two criteria, leading to the exclusion of one of them. However, when considering the reverse influence, where EBITDA affects Profit, the coefficient is lower (0.7), indicating a less dominant relationship in that direction.

Similarly, other criteria exhibit varying degrees of dependency on each other. For instance, CO2 Production, Profit, and EBITDA seem to be relatively closely interrelated, with coefficients ranging from 0.6 to 0.7. On the other hand, Firm Size and ROCE show relatively stronger dependencies with each other (0.9).

Analyzing these dependencies provides valuable insights for decision-making processes, allowing stakeholders to understand not only the strength but also the direction of influence among different criteria.

4 Discussion

Our findings highlight the need for sophisticated analytical approaches, such as fuzzy functional dependency analysis, to fully understand complex systems like emissions trading. This method helps policymakers, business leaders, and researchers make informed decisions and develop effective strategies, revealing subtle patterns not seen by the traditional correlation analysis. Identifying these dependencies allows stakeholders, for instance, to adjust weights in multi-criteria decision-making processes. By re-calibrating weights based on observed dependencies, decision models can remain dynamic and responsive to the evolving relationship between environmental and economic factors, enhancing policy and business strategies. Adjusting weights as follows based on dependency strength and direction:

- Strong mutual dependency: decrease weights of both criteria,
- No strong dependency: increase weights,
- One-way strong influence: increase the weight of the influencing criterion and decrease the weight of the influenced criterion, and vice versa.

Future analysis could explore hierarchical clustering of similar criteria, assigning group weights to further refine decision-making in multi-criteria analyses. Incorporating these adjustments and potential expansions can enhance decision-making accuracy and effectiveness within the EU ETS framework and similar contexts. We emphasize, that the proposed approach augments existing methods with additional insights, rather than competing with them. The main benefit of FFD is providing the direction of dependency between criteria, handling diverse data types and mitigating imprecision in data. However, the procedure is more demanding than the correlation analysis used in multi-criteria decision making for adjusting weights and is subjective in assigning proximity values, where new mathematical proprieties might be helpful.

This approach should be further improved: (i) a higher number of companies should be evaluated (currently, in this highly pollutant industry it is problematic due to missing data) (ii) fuzzification might be improved by considering data distribution (e.g., instead of equi-length fuzzy sets (Fig. 1, we might apply equi-dept approach) (iii) experimenting with different values of proximity relation.

5 Conclusions

Our analysis highlights the importance of considering both the strength and direction of dependencies within the criteria in EU ETS framework. Traditional correlation analysis overlooks the directional aspect, which might be useful in interpretations. We employed fuzzy functional dependency analysis to cover both strength and direction of dependencies. In this context, we found that economic and environmental criteria within the EU ETS are closely interrelated. Specifically, profit, EBITDA, and CO2 production exhibit significant dependencies, but not symmetric, emphasizing the link between financial performance and environmental impact. The weaker dependency found between CO2 emissions and profits (in the direction from CO2 production to profit) does not suggest a direct influence of the EU ETS or a strong incentive for firms to transition to greener technologies. Conversely, this revealed dependency tends to support the findings in [4], which highlighted the limited effectiveness of cost pass-through mechanisms.

Nonetheless, it's reasonable to conclude that the criteria associated with the EU ETS significantly influence individual firms' market positions and are inherently tied to economic and social criteria. Firm Size and ROCE also show strong dependencies, underscoring the role of firm size in determining return on capital employed. The proposed approach, allows values from different domains to be compared by applying linguistic variables, fuzzy sets, proximity relation and conformance measure. The reveled dependencies can be used in many tasks, for instance in adjusting initial weights, i.e., for strong mutual dependency, weights of both criteria are decreased; for weak mutual dependency, weights are increased; for one-way strong influence, increase the weight of the influencing criterion and decrease the weight of the influenced criterion. Finally, we emphasize, that the proposed approach augments existing methods with additional insights, rather than competing with them.

Acknowledgements

This article has been produced with the financial support of the SGS project No. SP2024/083 of the Ministry of Education, Youth and Sports of the Czech Republic and VEGA project No. 1/0120/23 of the Ministry of Education, Research, Development and Youth of the Slovak Republic.

References

[1] Ayan, B., Abacıoğlu, S., Basilio, M.P. (2023). A Comprehensive Review of the Novel Weighting Methods for Multi-Criteria Decision-Making. *Information*, 14(5), 285. https://doi.org/10.3390/info14050285.

[2] Bernardi, M., Davis-Peccoud, J., Lichtenau, T., Straehle, O. (2023). Putting Sustainability to Work in Paper and Packaging. Bain Company. Retrieved from https://www.bain.com/insights/putting-sustainability-to-work-paper-and-packaging-report-2023/

[3] Chufama, M., Sithole, F., Utaumire, Y. (2021). A Review of corporate social responsibility theories and models. *International Journal of Economics, Commerce and Management*, 9(3), 104-119.

[4] Cludius, J., et al. (2020). Ex-post investigation of cost pass-through in the EU ETS – an analysis for six industry sectors. *Energy Economics*, 91, 104883. DOI: 10.1016/j.eneco.2020.104883.

[5] Hudec, M., Vučetić, M., Barčáková, N. (2024). An extended proximity relation and quantified aggregation for designing robust fuzzy query engine. *Knowledge-Based Systems*, id. 111574.

[6] Hudec, M., Vučetić, M., Čermáková, I. (2020). The synergy of linguistic summaries, fuzzy functional dependencies and land coverings for augmenting informativeness in smart cities. In *Proceedings of 28th Telecommunications Forum (TELFOR)*, (pp. 1-4). Belgrade, November 24-25, 2020. IEEE.

[7] International Energy Agency. (2023). Tracking Clean Energy Progress 2023. Retrieved from https://www.iea.org/reports/tracking-clean-energy-progress-2023

[8] NDC Partnership. (n.d.). Emissions Trading Schemes (ETS). Retrieved from *https://ndcpartnership.org/knowledge-portal/good-practice-database/european-union-emissions-trading-system-eu-ets*.

[9] O'Riordan, L., Fairbrass, J. (2008). Corporate social responsibility (CSR): Models and theories in stakeholder dialogue. *Journal of Business Ethics*, 83, 745-758. DOI: 10.1007/s10551-008-9662-y.

[10] Shenoi, S., Melton, A. (1999). Proximity relations in the fuzzy relational database model. *Fuzzy Sets and Systems*, 100, 51–62.

[11] Sözat, M., Yazici, A. (2001). A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations. *Fuzzy Sets and Systems*, 117(2), 161–181.

[12] Vučetić, M., Hudec, M., Vujošević, M. (2013). A new method for computing fuzzy functional dependencies in relational database systems. *Expert Systems with Applications*, 40(7), 2738–2745.

[13] Vučetić, M., Vujošević, M. (2012). A literature overview of functional dependencies in fuzzy relational database models. *Technics Technologies Education Management*, 7(4), 1593–1604.

[14] Zadeh, L.A. (1975). The Concept of a Linguistic Variable and Its Application to Approximate Reasoning: Part I. *Information Sciences*, 8, 199–249.

Application of conventional DEA and ZSG-DEA models in university budget allocation

Eva Štichhauerová[1], Miroslav Žižka[2]

Abstract. This paper investigates the feasibility of applying two Data Envelopment Analysis (DEA) models in allocating budgets among departments in a university. The paper is divided into three parts. The first part describes the existing methodology of budget allocation in a selected faculty of a given university. Then, two DEA models that have been used in alternative budget construction are characterized. The first model is based on the classical CCR model, supplemented by the super-efficiency model of Andersen and Petersen. The model redistributes resources from inefficient units towards efficient units, depending on the degree of their super-efficiency. The reallocation takes place until all units are efficient. The second model is based on the zero-sum gains (ZSG) approach, where the sum of the resources of all units remains constant. Thus, the sum of the improvement and the worsening of all units under study must remain equal to zero. The last part of the paper compares the current budget allocation with alternative options under the classical and ZSG-DEA models.

Keywords: Budget allocation, Data Envelopment Analysis, zero-sum gains DEA model, Andersen and Petersen super-efficiency model

JEL Classification: C61, H52, I22
AMS Classification: 90C08

1 Introduction

The preparation of the university budget and its distribution to individual units (faculties, institutes, and departments) is one of every university's most important economic activities, and it is done every year. The management of the university or faculty should strive to ensure that the rules according to which the budget is drawn up are both motivating and objective. Through the budget, the university or faculty management implements its priorities and, at the same time, directs the units to achieve the best possible results in educational and creative activities. Budgeting is often discussed in the Academic Senate concerning rules and objectivity. It is not different at the Faculty of Economics of the Technical University of Liberec. The faculty uses multi-criteria evaluation to allocate the budget among departments. The size of a department's budget depends on various indicators of educational and research activities or the qualification structure of academic staff. Individual indicators have weights in the calculation, which the faculty management determines. These weights are often discussed, and whether they are set fairly is addressed. Therefore, this paper aims to compare the results of budget allocation using the current methodology with an alternative approach based on Data Envelopment Analysis. One of the advantages of DEA is that it determines the weights of the inputs and outputs used to be the best for the unit under consideration.

The paper is organized as follows. First, the DEA models used for subsequent efficiency analysis of each department are presented. These are the conventional CCR DEA model with the super-efficiency model and the alternative zero-sum gains (ZSG) DEA approach. Subsequently, the data used are explained, the technical efficiency scores are calculated, and the results are commented. Finally, the departments' shares of the total budget under the current methodology and the DEA models are compared.

2 Theoretical background

Data Envelopment Analysis (DEA) is a modelling technique based on linear programming that is used to evaluate the relative technical efficiency of Decision Making Units (DMUs). This analysis can identify efficient DMUs that lie on the empirical production possibility frontier. These units represent best-practices for other inefficient units that lie outside the efficient frontier. The technical efficiency score is a multiple performance metric because it

[1] Technical University of Liberec, Faculty of Economics, Department of Business Administration and Management, Studentská 2, Liberec 1, Czechia, eva.stichhauerova@tul.cz.
[2] Technical University of Liberec, Faculty of Economics, Department of Business Administration and Management, Studentská 2, Liberec 1, Czechia, miroslav.zizka@tul.cz.

typically includes more inputs and outputs. This larger number of inputs and outputs is then reduced to one virtual input and one virtual output using weights. The weights are set to maximize the efficiency of the units under study. The ratio between the virtual output and the virtual input indicates the degree of technical efficiency of the DMU (Dlouhý et al. 2018). The basic DEA model is the CCR, named after its authors in 1978 - Charnes, Cooper, and Rhodes. This model is based on the assumption of constant returns to scale (CRS). The model was developed to measure relative efficiency, mainly where market prices are absent. The original model was used to evaluate schools in Texas. However, the range of applications of the model expanded to other areas later, including the corporate sector (Zhu 2014). In our analysis, we use an input-oriented CCR model. We maximize the criterion function (1) under conditions (2), see e.g. (Dlouhý 2014).

$$\theta_q = \sum_{i=1}^{r} u_i y_{iq} \tag{1}$$

$$\begin{gathered} \sum_{i=1}^{r} u_i y_{ik} \le \sum_{j=1}^{m} v_j x_{jk}, k = 1, 2, \dots, n, \\ \sum_{j=1}^{m} v_j x_{jq} = 1 \\ u_i \ge \varepsilon, i = 1, 2, \dots, r, \\ v_j \ge \varepsilon, j = 1, 2, \dots, m \end{gathered} \tag{2}$$

Where x_{jk} are the inputs of production unit k, y_{ik} are the outputs of the same unit, u_i are the weights of the outputs, v_j are the weights of the inputs, ε is the infinitesimal constant, and θ_q is the efficiency score of unit q.

If the number of identified efficient units is high or the units need to be ranked according to the efficiency score, it is appropriate to apply super-efficiency models. There are several super-efficiency models; we use the Andersen and Petersen model from 1993 in our paper. This model compares the unit under study with a linear combination of all other units in the population. This means that the unit under evaluation is excluded from the solution, i.e., its weight λ_q is zero. This is because an efficient DMU can proportionally change its input vector while remaining efficient. The super-efficiency measure is then greater than one and measures the distance of the DMU from the new efficient frontier (Andersen and Petersen 1993). The essence of the calculation of the super-efficiency score under this input-oriented model and the CRS assumption is to minimize the criterion function θ_q^{AP} under conditions (3), see e.g. (Dlouhý et al. 2018).

$$\begin{gathered} \sum_{i=1}^{n} x_{ij} \lambda_i + s_j^- = \theta_q^{AP} x_{qj}, \qquad j = 1, 2, \dots, m, \\ \sum_{i=1}^{n} y_{ik} \lambda_i - s_k^+ = y_{qk}, \qquad k = 1, 2, \dots, r, \\ s_k^+ \ge 0, k = 1, 2, \dots, r, \\ s_j^- \ge 0, j = 1, 2, \dots, m, \\ \lambda_i \ge 0, i = 1, 2, \dots, n, i \ne q, \\ \lambda_q = 0 \end{gathered} \tag{3}$$

Where λ_i are the weights of the evaluated units, θ_q^{AP} is the super-efficiency measure, s_k^+ a s_j^- are dual variables for the weights of inputs and outputs.

When allocating a budget among a certain number of units, it is a simple allocation model with one input or one output. The budget can be seen as a single input (a basic resource) that influences the performance of departments in educational, creative, or other areas of activity. We could also understand the budget in reverse as a single output, where, in turn, departments try to influence the criteria (inputs) for its distribution. In this case, we understand the budget as a basic resource, i.e., we work with a single-input model. For such situations (one input or one output), the zero-sum gains DEA model (ZSG-DEA) can be used. ZSG-DEA is based on the assumption that DMUs are not completely independent in deciding the amount of resources consumed and outputs produced (Gomes et al. 2008). These are situations where resources are shared (typically departmental budgets within a faculty budget or faculty budgets within a university budget) or where the higher output of one unit limits the output of other units. That is, if a less efficient unit seeks to reach the efficient frontier, the position of other DMUs will logically change so that the sum of resources remains constant.

The ZSG-DEA model was developed (Lins et al. 2003) to assess the ranking of countries according to the Olympic medals won. Each country has given (sporting) resources that it tries to evaluate regarding medals won. This model assumed a constant sum of outputs (medals awarded). The additional output that a unit needs to become efficient must be taken away from all other units. The sum of the improvement and deterioration of outputs across all DMUs must equal zero. However, the original model was non-linear. Several approaches are in the literature to simplify that model, e.g. (Bernardo et al. 2020; Bouzidis and Karagiannis 2022).

The efficiency of unit q in the input-oriented ZSG-DEA model is given by relation (4), e.g. (Dlouhý et al. 2018).

$$\theta_q^{ZSG} = \theta_q \left(1 + \frac{\sum_{i \in E} x_{ij} \left(1 - \left(\frac{\theta_q}{\theta_i} \right) \theta_q^{ZSG} \right)}{\sum_{i \in E} x_{ij}} \right), j = 1, 2, \dots, m. \quad (4)$$

Where θ_q^{ZSG} is the unit efficiency q in the ZSG-DEA model, θ_q is the unit efficiency q in the traditional CCR model, and E is the set of unit efficiency indices.

In a practical application, it is necessary to proceed by first solving a conventional CCR model with one input and several outputs. The initial CCR DEA model can be significantly simplified by dividing the input equally among the units. The efficiency scores of each unit are determined. The input is then redistributed according to $\theta_q / \sum \theta_i$.

3 Data and methodology

The Faculty of Economics consists of departments marked in Table 1 with codes K1 to K7. We used input and output data from the 2023 faculty budgeting process and recalculated them to reflect the departments' shares of the faculty's activity. The budget, which is distributed to departments, consists of three resources on the income side (INP) - funds from indicator A (volume of educational activity), indicator K (quality), and institutional support for long-term conceptual development. Each resource is reallocated according to specific indicators (OUT). For this reason, we constructed three variants for each model. The resulting budget is then composed of the sum of the revenues from these three resources.

The first model contains the following input and output set:

- INP1: income from educational activities according to indicator A,
- OUT1: performance in educational activities over the past two semesters expressed in student hours,
- OUT2: number of thesis supervised in the last two semesters.

Second model:

- INP2: income by K (quality),
- OUT3: the current weighted number of academic staff per FTE, where the weights are the qualification achieved (assistant professor = 1, associate professor = 1.5, and professor = 2.5),
- OUT4: earmarked funds received for grants and research projects in a given year,
- OUT5: numbers of newly acquired PhD students in a given academic year.

Third model:

- INP3: income from institutional support for the long-term conceptual development of the faculty (science and research),
- OUT6: stabilizing institutional support according to the scores of individual publication outputs (different types of results have different score weights),
- OUT7: incentive institutional support (set in proportion to the publication allowances of departmental staff, which are derived from the quality of published work).

The analysis procedure can be divided into the following six steps. The procedure below has been carried out for all three sets of inputs and outputs:

Step 1: Absolute inputs and outputs values were converted to relative values, expressing the department's share in the sum of the given indicator for all faculty units.

Step 2: Each department's initial technical efficiency scores according to the CCR models were calculated (see relationships 1 and 2). Next, the super-efficiency score was determined for the efficient units according to the

Andersen and Petersen (AP) model, see relation (3). It is true that the higher the super-efficiency score, the better the department is rated.

Step 3: The redistribution mechanism described in (Dlouhý et al. 2018) was applied. The mechanism is based on the reallocation of resources from inefficient units towards efficient DMUs. The resource reduction rate is determined as the product of the resource amount and the efficiency score according to the CCR model. The resources thus obtained are reallocated among the efficient DMUs according to their super-efficiency score. The calculation of the CCR model efficiency score is repeated. The procedure is repeated until all DMUs are efficient.

Step 4: Next, efficiency scores were calculated according to the ZSG-DEA model with all departments' budgets equal in input. The initial efficiency scores were determined using the CCR model with one input. The budget was then reallocated according to the relationship $\boldsymbol{\theta_q / \sum \theta_i}$**.** We determined the super-efficiency scores according to model (3).

Step 5: The total budget of each department was determined as the sum of the shares of all three resources.

Step 6: Finally, we compared the departments' current budget shares with the shares proposed under the super-efficiency and ZSG-DEA models' redistribution mechanisms.

Dep.	INP1	INP2	INP3	OUT1	OUT3	OUT4	OUT5	OUT6	OUT7	OUT8
K1	0.2057	0.0989	0.0245	0.2048	0.0000	0.0874	0.4872	0.0000	0.0206	0.0286
K2	0.1206	0.1500	0.1402	0.1145	0.1801	0.1850	0.0000	0.1667	0.1295	0.1515
K3	0.1236	0.0960	0.0921	0.1255	0.1532	0.1275	0.0000	0.2500	0.0558	0.1307
K4	0.1403	0.0943	0.2041	0.1464	0.1290	0.1262	0.0000	0.0833	0.2558	0.1492
K5	0.1240	0.2200	0.2295	0.1314	0.2392	0.1599	0.3250	0.1667	0.1985	0.2624
K6	0.2016	0.3025	0.2734	0.2077	0.2285	0.2306	0.1878	0.3333	0.3241	0.2197
K7	0.0842	0.0383	0.0362	0.0696	0.0699	0.0834	0.0000	0.0000	0.0157	0.0579

Table 1 Input data for departmental budget allocation

4 Results

Table 2 shows the results of the baseline CCR model. The results show that only one department, K5, is fully efficient in the case of the distribution of indicator A (model 1). Still, the other five departments are close to the efficient frontier. Thus, we shift the reduced resources from the six departments in favour of K5. However, such a drastic shift will subsequently cause a decrease in the efficiency score of department K5. Therefore, it is necessary to continue the iterations. After seven iterations, we reach a state where all units are efficient.

Dep.	CCR eff. scores	Sup.-ef. scores	Modified budget	Sup.-ef. scores 2	Modified budget 2	Sup.-ef. scores 3	…..	Modified budget 7	Sup.-ef. scores 7
K1	0.9390	0.9390	0.1932	1.0000	0.1932	1.0000		0.1932	1.0006
K2	0.8963	0.8963	0.1081	1.2680	0.1451	0.8134		0.1246	1.0167
K3	0.9578	0.9578	0.1184	1.0000	0.1184	1.0264		0.1221	1.0155
K4	0.9845	0.9845	0.1381	1.0000	0.1381	1.0000	…	0.1382	1.0045
K5	1.0000	1.2917	0.1806	0.7947	0.1435	1.2883		0.1577	1.0495
K6	0.9720	0.9720	0.1959	1.0000	0.1959	1.0000		0.1980	1.0062
K7	0.7798	0.7798	0.0657	1.0000	0.0657	1.0000		0.0661	1.0005

Table 2 Efficiency scores according to CCR and AP super-efficiency model - model 1 (indicator A)

In the case of indicator K, the differences in the degree of efficiency between departments are more pronounced. Only three departments are efficient. The calculation procedure is analogous to that for indicator A. The optimal budget share distribution is reached after four iterations, see Table 3.

Four departments are fully efficient according to the CCR model for institutional support, but the remaining three are close to the efficient frontier, see Table 4. We found the optimal budget allocation after four iterations, with only a slight reallocation of resources between the third and fourth iterations.

Dep.	CCR eff. scores	Sup.-ef. scores	Modified budget	Sup.-ef. scores 2	Modified budget 2	Sup.-ef. scores 3	Modified budget 4	Sup.-ef. scores 4
K1	1.0000	3.3349	0.2222	0.9282	0.2063	1.2486	0.2080	1.2501
K2	0.7333	0.7333	0.1010	1.0505	0.1352	1.0654	0.1366	1.0218
K3	1.0000	2.3424	0.1827	0.7329	0.1339	1.1776	0.1354	1.1721
K4	0.7471	0.7471	0.0705	1.0645	0.0960	0.8999	0.0865	1.0166
K5	0.6255	0.6255	0.1376	1.4255	0.1718	1.1915	0.1734	1.1898
K6	0.5902	0.5902	0.1785	1.3215	0.2102	1.1138	0.2117	1.1178
K7	1.0000	1.6264	0.0985	0.4730	0.0466	1.3085	0.0484	1.1812

Table 3 Efficiency scores according to the CCR and AP super-efficiency model - model 2 (indicator K)

Dep.	CCR eff. scores	Sup.-ef. scores	Modified budget	Sup.-ef. scores 2	Modified budget 2	Sup.-ef. scores 3	Modified budget 4	Sup.-ef. scores 4
K1	0.9974	0.9974	0.0245	1.0000	0.0245	0.9999	0.0244	1.0000
K2	1.0000	1.0002	0.1402	0.9998	0.1402	1.0000	0.1402	1.0000
K3	1.0000	1.0004	0.0922	1.0006	0.0922	1.0007	0.0922	1.0007
K4	1.0000	1.0578	0.2041	1.0573	0.2041	1.0573	0.2041	1.0573
K5	0.9999	0.9999	0.2295	1.0000	0.2295	1.0001	0.2295	1.0001
K6	0.9997	0.9997	0.2734	1.0002	0.2734	1.0002	0.2734	1.0002
K7	1.0000	1.1278	0.0362	1.1270	0.0362	1.1269	0.0362	1.1269

Table 4 Efficiency scores according to the CCR and AP super-efficiency model - model 3 (indicator IS), note: the difference in budget shares is up to the fifth decimal place.

We calculated the efficiency scores in the next stage according to the ZSG-DEA models. The inputs for all three models will be the same (equal) departmental shares of the faculty budget, and the outputs are the same as for the previous models. Proposed budget allocations and associated efficiency or super-efficiency scores for indicators A and K are shown in Table 5 and for IP in Table 6.

	Indicator A				Indicator K		
Dep.	Equal budget	CCR eff. scores	Modified budget	ZSG-DEA efficiency	CCR eff. scores	Modified budget	ZSG-DEA efficiency
K1	0.1429	0.9859	0.1809	1.0000	1.0000	0.1837	1.4753
K2	0.1429	0.7613	0.1397	1.0000	0.8021	0.1473	1.0000
K3	0.1429	0.6631	0.1217	1.0000	0.7500	0.1377	1.0000
K4	0.1429	0.7048	0.1293	1.0000	0.5471	0.1005	1.0000
K5	0.1429	1.0000	0.1835	1.0113	0.9840	0.1807	1.0000
K6	0.1429	1.0000	0.1835	1.0431	1.0000	0.1837	1.3320
K7	0.1429	0.3352	0.0615	1.0000	0.3617	0.0664	1.0000

Table 5 Results of ZSG-DEA models 1 + 2 - indicators A and K

Table 7 then compares the current departmental shares of each budget resource with the options determined by the iterative procedure based on the super-efficiency models and the ZSG-DEA model.

Department	Equal budget	CCR eff. scores	Modified budget	ZSG-DEA efficiency
K1	0.1429	0.1089	0.0259	1.0000
K2	0.1429	0.5929	0.1408	1.0000
K3	0.1429	0.4982	0.1183	1.0000
K4	0.1429	0.7894	0.1875	1.0000
K5	0.1429	1.0000	0.2375	1.0069
K6	0.1429	1.0000	0.2375	1.0780
K7	0.1429	0.2206	0.0524	1.0000

Table 6 Results of ZSG-DEA model 3 - IS

Dep.	Proposed change compared to current budget - application of CCR + AP models			Proposed change compared to current budget - application of ZSG-DEA model		
	A	K	IS	A	K	IS
K1	-0.0126	0.1091	-0.0006	-0.0248	0.0848	0.0013
K2	0.0041	-0.0133	0.0001	0.0191	-0.0027	0.0007
K3	-0.0015	0.0394	0.0004	-0.0020	0.0417	0.0262
K4	-0.0021	-0.0079	0.0004	-0.0110	0.0062	-0.0165
K5	0.0338	-0.0465	-0.0001	0.0595	-0.0392	0.0081
K6	-0.0035	-0.0907	-0.0007	-0.0181	-0.1188	-0.0359
K7	-0.0182	0.0100	0.0004	-0.0227	-0.0281	0.0162

Table 7 Comparison of departmental budget shares by method

5 Conclusions

The results of the analyses show that the most minor differences in departmental shares compared to the current budget are in the institutional support field. Most departments were also at or near the efficient frontier for this indicator. Significant differences between departments were also not found in the case of super-efficiency. There were also relatively small differences in the distribution of funds from indicator A. A more significant increase in resources would only occur in the case of one department (K5). However, more significant differences from the current budget were found when allocating resources by indicator K. For four departments, the reallocation of resources was relatively high, two cases approaching 10%. The results imply that faculty management should reflect on the construction of the K indicator, including the weights of the sub-indicators. The K indicator should reflect the quality of activities of individual departments. It is questionable whether the current composition of the indicator stimulates departments to increase quality. If we compare the results of the budget allocation by the super-efficiency and ZSG-DEA models, we can conclude that the proposed alternative departmental budgets are similar. The most considerable difference for indicator A is for department K5 (2.6 percentage points). For indicator K, the difference is 2.8 percentage points for department K6. The most significant difference for institutional support is 3.6 percentage points for department K6. The differences are since the first model considers the super-efficiency of the assessed departments. On the other hand, the ZSG-DEA method is more straightforward to apply. In conclusion, applying DEA models can help objectify discussions regarding departmental performance within the faculty and consequently guide departmental budgeting.

Acknowledgements

The contribution was supported by institutional funds for long-term conceptual development of the Technical University of Liberec research organization.

References

Andersen P, Petersen NC (1993) A Procedure for Ranking Efficient Units in Data Envelopment Analysis. Manag Sci 39:1261–1264. https://doi.org/10.1287/mnsc.39.10.1261
Bernardo M, De Souza MAM, Moreira Lopes RS, Rodrigues LF (2020) University library performance management: Applying zero-sum gains DEA models to resource allocation. Socioecon Plann Sci 100808. https://doi.org/10.1016/j.seps.2020.100808
Bouzidis T, Karagiannis G (2022) A note on the zero-sum gains data envelopment analysis model. Oper Res 22:1737–1758. https://doi.org/10.1007/s12351-021-00631-8
Dlouhý M (2014) Models of Subsidy Allocation among City Districts. Prague Econ Pap 23:108–120. https://doi.org/10.18267/j.pep.475
Dlouhý M, Jablonský J, Zýková P (2018) Analýza obalu dat [Data Envelopment Analysis], Professional Publishing, Průhonice
Gomes EG, Mello JCCBSD, Meza LA (2008) Large discreet resource allocation: a hybrid approach based on dea efficiency measurement. Pesqui Oper 28:597–608. https://doi.org/10.1590/S0101-74382008000300011
Lins MPE, Gomes EG, Soares De Mello JCCB, Soares De Mello AJR (2003) Olympic ranking based on a zero sum gains DEA model. Eur J Oper Res 148:312–322. https://doi.org/10.1016/S0377-2217(02)00687-2
Zhu J (2014) Quantitative Models for Performance Evaluation and Benchmarking. Springer International Publishing, Cham

Generation and Analysis of Intralogistics Solution Scenarios

Dominika Tejnicka Bordacsova[1], Jana Hanclova[2], Lucie Chytilova[3]

Abstract. This contribution focuses on the identification and generation of intralogistics solution combinations that meet specific customer requirements. The classical framework and processes are innovated based on expert experiences and the future development of the company in a competitive environment. The proposed system utilizes historical data and expert knowledge while respecting customer requirements when establishing individual scenarios.

For each scenario, an intralogistics solution that minimizes the final cost while adhering to all constraints is found through optimization. For the most probable scenario, a distribution including risk will be sought, and changes in the solution's performance will be monitored. Performance values will be compared with the remaining scenarios to model the time intensity of picking within the designed intralogistics solutions. The results obtained enable the identification of the impacts of time intensity on performance within the examined scenarios, allowing for a more realistic and precise simulation of intralogistics processes.

A critical part of our method is the evaluation of various scenarios, which enables us to quantify the impacts of uncertainty in picking speed on the performance of automated systems and to identify the optimal solution for each unique customer situation. This work offers a new perspective on planning and managing intralogistics by combining stochastic modeling with practical market needs. Our analysis results demonstrate how efficient and adaptive intralogistic systems can be when properly designed and calibrated considering the inherent uncertainties of operations. Our findings provide a valuable foundation for further research and development in the field of automated intralogistics and offer practical guidelines for implementation in real logistic operations.

Keywords: intralogistics, optimization, scenario analysis

JEL Classification: C61, L90, O14,
AMS Classification: 90B06, 90C05

1 Introduction

1.1 Intralogistics and Industry 4.0

The industrial world is witnessing exponential change and continuous evolution. What industry was a century ago is now unrecognizable due to the rapid advancements and transformations. These changes occur as reactions to tough competition and difficulties to ensure sustainability and profit. One of the latest and most significant transformations would be the transition to Industry 4.0. Also known as the fourth industrial revolution, Industry 4.0 is a term used to describe the growing trend of automation and data exchange in manufacturing technologies. (Jamwal et al., 2021) (Jaskó et al., 2020) (Zheng et al., 2021)

Industry 4.0 creates what has been called a 'smart factory'. The term smart factory is used when machines are augmented with wireless connectivity and sensors and can make decisions without human involvement. These technologies can assist in the revolution of the traditional supply chain to what is now deemed intralogistics. Intralogistics covers a wide range of solutions for the optimization of workflows. (Pivoto et al., 2021) (Ahmed et al., 2021) (Javaid et al., 2023)

[1] VSB – Technical University of Ostrava, Faculty of Economics, Department of Systems Engineering and Informatics, 17. listopadu 2172/15, 708 00 Ostrava-Poruba, Czech Republic, dominika.bordacsova.st@vsb.cz
[2] VSB – Technical University of Ostrava, Faculty of Economics, Department of Systems Engineering and Informatics, 17. listopadu 2172/15, 708 00 Ostrava-Poruba, Czech Republic, jana.hanclova@vsb.cz
[3] VSB – Technical University of Ostrava, Faculty of Economics, Department of Systems Engineering and Informatics, 17. listopadu 2172/15, 708 00 Ostrava-Poruba, Czech Republic, lucie.chytilova@vsb.cz

The field of intralogistics is evolving through significant technological advancements that emphasize automation, data integration, and human-machine cooperation. These developments not only optimize current systems, but also set the stage for future innovations that could further transform the landscape of internal logistics operations. In today's rapidly evolving industrial environment, efficiency and optimization are key terms that influence every aspect of business, particularly in areas such as storage and intralogistics.

1.2 Decision support system

In the realm of decision-making exclusively, decision-making systems fall into two broad categories. The first is the independent and somewhat intuitive decisions of an individual or group, usually based on accumulated experience or rules of thumb. The second is the systematic analysis of the data. This may lead to the best decision with the available information through quantitative analysis, decision analysis, or optimization. The latter method of decision-making is combined with problem-solving in a single system to find a solution. (Burton et al., 2020) (Ouyang et al., 2022)

For a company focused on warehouse automation solutions, continual improvement of internal logistics processes is essential not only to maintain competitiveness, but also to ensure maximum efficiency and minimize operational costs. To find the most suitable solution that meets customer needs, it's necessary to analyze various parameters such as customer space, expected performance, and other operational requirements to suggest the most suitable machinery configuration.

Problem solving can be simplified into three steps. First, identify and define the problem. Second, find different possible solutions, and finally, rank the alternatives and select the best possible solution. The objective of decision automation systems is to provide the best solution quickly and cost-effectively, aligned with warehouse automation systems. A solution-finding system for a warehouse problem can be simplified into three steps and contrasted with current methods. (Cantini et al., 2024)

For this reason, a new decision support system was created. The optimization system designed in this article focuses on the efficient design and optimization of intralogistics solutions in warehouses using Integer Linear Programming (ILP). This system, which works with an extensive database that contains information about available machines and technologies, is designed to provide optimised solutions for specific client needs based on input parameters such as speed, performance, and other operational constraints, using advanced algorithms.

The designed optimization system offers a practical solution for enhancing logistics processes within warehouses. Its implementation can significantly increase efficiency and reduce costs for companies specializing in automated storage systems and intralogistics. The primary objective of the system is to identify the optimal combination of automated storage machines that meet the specific customer requirements while minimizing costs. Additionally, the main goal of this article is to compare the most likely scenario (MLS) solution, its performance over different timeframes, and to compare this analysis with the original scenarios.

This article is structured into five sections. The first section introduces intralogistics and the proposed optimization system. The second section details the methodology and input data. The third section is devoted to the analysis and comparison of scenarios. The final two sections cover a discussion and a conclusion.

2 Methodology and data

The system is designed to integrate customer-specific constraints and requirements, subsequently offering the optimal combination of machinery to meet those needs. The overall solution's performance is calculated using a measure known as picking time (t). A deviation in picking time (o), denoted as $\pm o$ from the average picking time, is considered to account for the human factor. Based on these two parameters, three scenarios are made ($t+o$, t, $t-o$). For each scenario, the most suitable solution is identified. Our focus will shift to the solution devised for the scenario characterized by the average picking time. In analyzing the solution for the most likely scenario, we aim to explore the impact of varying picking times ($t+o$, t, $t-o$) on its performance.

2.1 Input data

The system utilizes a database containing information about various types of storage machines, such as their dimensions, prices, installation costs, and other parameters. Additional inputs include specific customer requirements, such as the height of the stored material, the the maximum usable height and width of the warehouse, the required storage capacity, and the the performance of the system. All of them can be described using the name and [unit]: *Mean Material Height* [mm], *Max Available Space Height* [mm], *Max Available Space Width* [mm], *Required Storage Floor Capacity* [m^2], *Required System Performance* [picks/hour], *Estimated Mean Pick Time* [s/pick], *Pick Time Deviation* [s], *Estimated or Calculated Hitrate* [picks/tray presentation].

The optimization model includes an objective function (1) that aims to minimize the solution's total cost (including machine prices, installation costs, price for shelves) and a set of constraints (2 – 6) including limits on the maximum height and width of the space, the minimum required storage capacity, and the minimum required system performance, that ensure the design will meet all specific requirements:

$$Z_{MIN} = \sum_{i=1}^{n}(c_i \cdot x_i + installation_i \cdot x_i + shelf_i \cdot x_i + freight\_cost_i), \quad (1)$$

$$height_i \cdot x_i \le max\_height, \quad (2)$$

$$\sum_{i=1}^{n} width_i \cdot x_i \le max\_width, \quad (3)$$

$$\sum_{i=1}^{n} storage\_capacity_i \cdot x_i \ge required\_storage\_capacity, \quad (4)$$

$$\sum_{i=1}^{n} system\_performance_i \cdot x_i \ge required\_system\_performance, \quad (5)$$

$$\sum_i x_i \in \mathbb{Z}, \quad (6)$$

where x_i is decision variable indicating the quantity of unit *i* represented in the solution, c_i [eur] are costs associated with unit *i*, *installation* [eur] are installation costs of the solution in euros, *shelf* [eur] represents cost for the calculated number of shelves in euros, *freight_cost* [eur] are transportation costs, *width* [mm] is width of the machine in mm, *height* [mm] means height of the machine in mm, *max_height* [mm] is maximum height of the space available for use, *max_width* [mm] is maximum allowable total width of the solution, *storage_capacity* [m^2] is storage area in m^2, *required_storage_capacity* [m^2] is required storage area in m^2, *system_performance* [picks/hr] represents performance of the solution and *required_system_performance* [picks/hr] is required performance of the solution.

After entering the input data, the system dynamically loads the information from the database and generates the available solutions using the mathematical model. The output includes optimal combinations of storage machines that meet all the specified conditions at minimal cost. The results are presented in three scenarios (Best Case, Most Likely and Worst Case), allowing for comparison of different solution variants.

The system also includes options for additional conditions:

1. maximum utilization of the available warehouse height;
2. same width of all machines within the given solution.

These conditions can be applied independently or simultaneously, affecting the selection of the optimal solution. Below, you can see other additional conditions for system functionality (7 – 15):

$$shelf\_size = self\_height \cdot self\_depth, \quad (7)$$

$$total\_height = self\_height \cdot columns - 1200, \quad (8)$$

$$shelf = total\_height/(material\_height + 100), \quad (9)$$

$$vertical\_time = height_i/1000 \cdot 1{,}1, \quad (10)$$

$$floor\_kapacity = (shelf \cdot shelf\_width \cdot 820)/1000000, \quad (11)$$

$$volume\_capacity = (floor\_capacity \cdot material\_height)/1000, \quad (12)$$

$$pickrate_T = 3600/\max(picking\_ime + 14 \,||\, vertical\_time + horizontal\_time\_T), \quad (13)$$

$$pickrate_L = 3600/(picking\_time + vertical\_time + horizontal\_time\_L), \quad (14)$$

$$final\_pickrate = pickrate \ \cdot hitrate, \quad (15)$$

where *tray_size* is used to calculate the dimensions of the shelf where *self_height* and *self_depth* are parameters of the machine. *Total_height* represents the total storage height inside the given machine, *self_height* is the height of the machine, and *columns* is a constant representing the number of columns of shelves in the machine (columns=2) with a constant *1200* for a single dispensing opening in the machine. The *shelf* is used to calculate the number of shelves, where we work with the calculated *total_height*, the *material_height*, and a constant *100* for the additional height needed for each shelf. *Vertical_time* represents the time of the vertical movement of the shelf in the machine, *horizontal_time_L* and *horizontal_time_T* are constants for the time of horizontal movement in the machines. *Floor_capacity* calculates the storage area using the number of shelves *shelf*, *shelf_width*, the constant *820* representing the length of the shelf, and the constant *1,000,000* for conversion to square meters. The *volume_capacity* is calculated using *floor_capacity, material_height* and *1000* constant for conversion to square meters.

2.2 Scenario generation

The optimization system generates three different scenarios, allowing users to see how various input data and conditions affect the final solutions. These scenarios are designed to reflect different possible real-world situations. The scenarios differ on the basis of picking time and picking deviation, which are used to calculate the performance of each solution.

Best Case Scenario (BCS): This scenario represents the most optimistic outcome, where all variables are set to minimize costs and maximize efficiency. Compared to the other scenarios, this one provides the lowest price and the highest efficiency. It is suitable for situations where conditions are ideal, and variables are most favourable.

Most Likely Scenario (MLS): This scenario represents the most likely outcome based on average values of the picking time and realistic conditions. This scenario provides a realistic view of possible solutions and is suitable for most practical applications where typical operating conditions are expected.

Worst Case Scenario (WCS): Based on the slowest picking time, this scenario represents the most pessimistic outcome, leading to higher costs and lower efficiency. This scenario provides the worst possible outcome and is useful for risk analysis and preparedness for the worst-case situations. It helps users understand what costs and performance to expect under the worst conditions.

If we consider comparing scenarios in terms of cost, performance, and realism, BCS has usually the lowest costs, MLS has moderate costs, and WCS has the highest costs. Sometimes the costs of the MLS and WCS can be the same. BCS usually doesn't offer the best performance, because it's the least expensive solution. MLS offers average performance, and WCS has the highest performance under any circumstances. MLS is the most realistic scenario, considering average conditions and inputs, while BCS and WCS represent extremes.

These scenarios allow users to compare different solutions and choose the most suitable one according to their specific needs and expectations. They also help to plan and prepare for various possible situations in intralogistics and automated warehouses.

To further analyze the scenarios, we will take the Most Likely Scenario and examine its performance under the fastest, slowest, and average picking times. This analysis will then be compared with the original three scenarios. Each of the original scenarios has a different machine configuration. Therefore, we will study the performance of one specific configuration (from the MLS) under all three picking times (picking time and picking deviation).

3 Scenario analysis

Each of the original scenarios has a different machine configuration. For a comprehensive analysis, we will take one specific configuration from the Most Likely Scenario and examine its performance under the following conditions:

- *Fastest Picking Time*: Evaluating the system's performance with the minimum possible picking time.
- *Average Picking Time*: Evaluating the system's performance with the average picking time.
- *Slowest Picking Time*: Evaluating the system's performance with the maximum possible picking time.

This analysis will provide information on how the chosen configuration performs under varying conditions and will be compared to the original three scenarios. This approach helps to understand the robustness and flexibility of the proposed solutions. By conducting this detailed analysis, we aim to identify the most efficient configuration that can adapt to different operational conditions, ensuring optimal performance and cost-effectiveness in intralogistics and automated warehouse solutions.

3.1 Performance analysis of the Most Likely Scenario

Table 1 shows a specific example of the original Most Likely Scenario for a fictional customer given by the optimization decision support system after entering the input data. Each scenario includes the *machine type*, *quantity*, *price*, *height*, *width*, *number of shelves*, *floor capacity*, *volume capacity*, and *pick rate* for the given type of machine within the solution. In addition, the sum of all these variables and the overall performance of the proposed solution is calculated, adhering to the given requirements.

machine type	quantity	price	height	width	shelf	floor cap.	vol. cap.	pick rate
A	2	80678	7976	2440	42	84,03	21,01	116,47
B	2	70980	7976	2440	42	84,03	21,01	84,66
Total	4	303316	7976	9760	168	336,12	84,04	402,26

Table 1 Most likely scenario

The analysis continues with the MLS, to compare its performance behavior for all 3 picking times as shown in *Table 2* below. The performance (*Pick Rate*) for each machine was calculated using the MLS and (7 – 15).

machine / pick time	fastest time (15s)	average time (20s)	slowest time (25s)
A	120,83	116,47	101,54
B	94,80	84,66	76,49
MLS pick rate	431,25	**402,26**	356,05

Table 2 Machine performance

As is evident from both logical reasoning and the *Table 2*, the fastest time represents the highest performance, while the slowest time corresponds to the lowest performance. This table allows the customer to evaluate the behavior of the given solution at different times, considering the estimated picking time and the probable deviation.

3.2 Comparison of MLS performance with the original scenarios

Below there are the tables for the original Best Case (*Table 3*) and Worst Case (*Table 4*) scenarios. These are used for the subsequent analysis. The scenarios include the same variables as the previous MLS in *Table 1.*

machine type	quantity	price	height	width	shelf	floor cap.	vol. cap.	pick rate
A	1	80678	7976	2440	42	84,03	21,01	120,83
B	3	70980	7976	2440	42	84,03	21,01	94,8
Total	4	293618	7976	9760	168	336,12	84,04	405,23

Table 3 Best case scenario

machine type	quantity	price	height	width	shelf	floor cap.	vol. cap.	pick rate
A	4	80678	7976	2440	42	84,03	21,01	101,54
Total	4	322712	7976	9760	168	336,12	84,04	406,16

Table 4 Worst case scenario

Since the performance of a machine is also dependent on its height, and the heights of the machines are the same cross all solutions in this case, their performances are also equal. Using the calculated performance of the machines (*Table 2*), the behavior of the original scenarios for the three considered times was calculated in *Table 5*.

scenario/pick time	fastest time	average time	slowest time	price/solution [€]	price diff. [€]
Best Case	**405,23**	370,46	330,99	293 618	-9 698
Most Likely	431,25	**402,26**	356,05	**303 316**	0
Worst Case	483,32	465,88	**406,16**	322 713	19 396

Table 5 Pick rate and price comparison

Subsequently, the prices of the individual solutions were compared to the MLS. When comparing the price and performance of our considered MLS with the BCS, the performance of the BCS is slightly higher and the price is nearly €10,000 lower. The performance of the WCS is slightly higher than that of the BCS, but the price is more than €19,000 higher. This comparison of the price and performance of individual solutions can be very useful for the customer when making a decision.

4 Discussion

The comparison of the Most Likely, Best Case, and Worst Case scenarios provides a comprehensive view of potential outcomes based on different picking times and associated costs. The Most Likely scenario represents the most balanced and realistic outcome, offering a practical perspective for most real-world applications.

The BCS, while offering the lowest costs, assumes ideal conditions that may not always be achievable in practice. It serves as a benchmark for the best possible efficiency, but it may not account for unforeseen operational challenges. On the contrary, the WCS provides a valuable risk assessment tool, highlighting the highest possible costs and ensuring preparedness for less favourable conditions.

When the performance is analyzed under varying picking times, the robustness of the solutions can be evaluated. The Most Likely scenario's configuration shows resilience across different conditions, indicating its suitability for dynamic operational environments. This analysis underscores the importance of considering multiple scenarios to ensure that the chosen solution is not only cost-effective, but also adaptable to changes in operational efficiency.

Furthermore, the system's ability to dynamically integrate new data and adjust recommendations accordingly demonstrates its flexibility and relevance in an ever-changing industrial landscape. This adaptability is crucial to maintaining efficiency and competitiveness in intralogistics operations.

5 Conclusion

The optimization system developed for intralogistics solutions offers a powerful tool for designing efficient and cost-effective storage systems in warehouses. By utilizing integer linear programming and an extensive database of machine parameters, the system provides customised solutions that meet specific customer requirements.

The analysis of the Most Likely, Best Case, and Worst Case scenarios demonstrates the system's capability to handle varying operational conditions and highlights the trade-offs between cost and performance. The MLS, with its balanced approach, emerges as the most practical solution for typical operating conditions, while the BCS and WCS offer insights into the extremes of potential outcomes.

This comprehensive approach to decision-making in warehouse automation ensures that businesses can optimize their logistics processes, reduce costs, and maintain competitiveness in a rapidly evolving industrial landscape. The system allows decisions to be made based on rigorous data analysis, ensuring that all factors are considered. Additionally, it can integrate the expertise and practical experience of industry experts, enhancing the reliability and applicability of the solutions provided.

Combining analytical precision with expert knowledge, the system provides a robust framework to select the most suitable intralogistics solutions. This dual approach not only ensures optimal performance and cost-effectiveness but also adapts to real-world challenges, making it a valuable asset for any business looking to enhance its internal logistics operations.

Furthermore, future research could explore the integration of real-time data analytics and machine learning techniques to further improve decision-making accuracy and efficiency. Such advancements could lead to even more adaptive and intelligent intralogistics systems, capable of proactively responding to changing conditions and optimizing performance continuously.

Acknowledgements

This article has been produced with the financial support of the European Union under the REFRESH – Research Excellence for Region Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition and by the SGS project (SP2024/083) of the Faculty of Economics, VŠB - Technical University of Ostrava.

References

Ahmed R S, Ahmed E S A, and Saeed R A (2021) Machine learning in cyber-physical systems in industry 4.0. In Artificial intelligence paradigms for smart cyber-physical systems (pp. 20-41). IGI global. https://www.igi-global.com/chapter/machine-learning-in-cyber-physical-systems-in-industry-40/266131

Burton J W, Stein M K, Jensen T B (2020) A systematic review of algorithm aversion in augmented decision making. Journal of behavioral decision making – Wiley Online Library. https://doi.org/10.1002/bdm.2155

Cantini A, Peron M, De Carlo F and Sgarbossa F (2024) A decision support systém for configuring spare parts supply chains considering different manufacturing technologies, International Journal of Production Research, 62:8, 3023-3043. https://doi.org/10.1080/00207543.2022.2041757

Jamwal A, Agrawal R, Sharma M, and Giallanza A (2021). Industry 4.0 technologies for manufacturing sustainability: A systematic review and future research directions. Applied Sciences. *11*(12), 5725. https://doi.org/10.3390/app11125725

Jaskó S, Skrop A, Holczinger T, Chován T and Abonyi J (2020). Development of manufacturing execution systems in accordance with Industry 4.0 requirements: A review of standard-and ontology-based methodologies and tools. Computers in industry, 123, p.103300. https://doi.org/10.1016/j.compind.2020.103300

Javaid M, Haleem A, Singh R P and Suman R (2023). An integrated outlook of Cyber–Physical Systems for Industry 4.0: Topical practices, architecture, and applications. Green Technologies and Sustainability, 1 (1), p.100001. https://doi.org/10.1016/j.grets.2022.100001

Ouyang F, Zheng L, Jiao P (2022) Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. Education and Information Technologies, Springer. https://doi.org/10.1007/s10639-022-10925-9

Pivoto D G S, De Almeida L F F, Da Rosa Righi R, Rodrigues J J P C, Lugli A B and Alberti A M (2021) Cyber-physical systems architectures for industrial internet of things applications in Industry 4.0: A literature review. Journal of manufacturing systems, 58, pp.176-192. https://doi.org/10.1016/j.jmsy.2020.11.017

Zheng T, Ardolino M, Bacchetti A and Perona M (2021). The applications of Industry 4.0 technologies in manufacturing context: a systematic literature review. International Journal of Production Research, 59(6), pp.1922-1954. https://doi.org/10.1080/00207543.2020.1824085

A Coalition Formation as a Multicriteria Voting Game

Michaela Tichá[1], Martin Dlouhý[2]

Abstract. In the multicriteria voting game, we assume a set of political parties and a set of political programs with multiple dimensions of public policy. The coalition program is formulated as a weighted average of the individual political programs. The objective of each political party is to minimize the maximum distance between the coalition program and its own program, to maximize its own share of power in the winning coalition, and to maximize the stability of the winning coalition, which is measured as the maximum distance between the coalition program and the individual programs of all political parties in the coalition. The multicriteria voting game model is applied to the Chamber of Deputies of the Parliament of the Czech Republic. Deputies were surveyed using a questionnaire that included 16 key questions. For each question, deputies selected from five scaled responses. The data obtained was used to calculate the optimal winning coalition.

Keywords: game theory, multicriteria voting game, coalition stability, political power dynamics

JEL Classification: C71
AMS Classification: 91A12

1 Introduction

The realm of game theory offers a robust framework for analyzing situations where conflicts and cooperative decision-making intersect among multiple agents (Binmore, 1992). Within this domain, voting games, a specialized subset, focus explicitly on the decision-making processes that occur within political bodies or other organizational groups (Dlouhý et al., 2015). This paper extends our previous work (Dlouhý et al., 2016), applying the multicriteria voting game model to contemporary data from the 2021 Czech Republic parliamentary elections, thereby demonstrating the model's ongoing relevance and utility in understanding political dynamics.

In our enhanced multicriteria voting game model, we consider a collection of political parties and a corresponding set of political programs that encompass a broad spectrum of public policy dimensions. By constructing a coalitional program as a weighted average of individual party programs, we introduce a methodological approach that quantifies the extent of compromise and consensus among coalition members. The objectives for each political party within this model are threefold: minimizing the maximum discrepancy between the coalition program and its own program, maximizing its relative power within the governing coalition, and maximizing the coalition stability. This stability is quantitatively assessed by evaluating the maximum divergence among all parties' programs within the coalition.

The application of this model to the Chamber of Deputies, the lower house of the Czech Parliament, serves not only to validate the model but also to illustrate its practical applications. By integrating responses from a detailed questionnaire completed by the deputies, which solicited their positions on a range of policy issues, we were able to construct a detailed landscape of political alignment and potential coalition configurations.

2 Model Formulation

Let $N=\{1,2,\dots,n\}$ be the set of political parties, where a_i represents the number of deputies from political party i. The ideology, or political program, is traditionally modeled geometrically. The simplest form is the one-dimensional left-right model. More sophisticated models position political parties within a multi-dimensional ideological space (Dlouhý et al., 2016). In our model, political parties are characterized by their values across m public policy dimensions, quantified within the [0, 1] interval. The vector of values for each political party in these m dimensions is denoted by

$$k^i = (k_1^i, k_2^i, \dots, k_m^i), \tag{1}$$

[1]Jan Evangelista Purkyně University, Dpt. of Mathematics, Pasteurova 3544/1, Ústí nad Labem, e-mail: michaela.ticha@ujep.cz.
[2] Prague University of Economics and Business, Dpt. of Econometrics, W. Churchill Sq. 4, Prague 3, e-mail: dlouhy@vse.cz.

where k_j^i is the value for political party i in the j-th public policy dimension. The coalition program of the winning coalition S is denoted by

$$K^S = (K_1^S, \dots, K_m^S), \tag{2}$$

where Kj^S is the value for the j-th public policy dimension. If political parties form a winning coalition, the coalition program represents a compromise of individual political programs. The program of the winning coalition S in policy dimension j is calculated as the weighted average of the programs based on the power of individual political parties:

$$K_j^S = \sum_{i \in S} \frac{a_i}{\sum_{i \in S} a_i} \; k_j^i. \tag{3}$$

With the coalition program K^S, we can now evaluate the expected benefits for a political party from being a member of this winning coalition. We assume that each political party has three objectives: (1) to implement its political program, (2) to maximize its relative power within the coalition, and (3) to be part of a politically stable coalition, see (Straffin, 1994). Quantitatively, these objectives are formulated as follows:

1. Each political party i in the winning coalition S is interested in minimizing the maximal difference between its individual program and the coalition program:

$$f_1(\mathrm{i}, \mathrm{S}) = \max_{j \in 1,2,\dots,m} \left| k_j^i - K_j^S \right|. \tag{4}$$

2. Each political party i in the winning coalition S aims to maximize its power, measured by the number of deputies:

$$f_2(i, S) = \frac{a_i}{\sum_{i \in S} a_i}. \tag{5}$$

3. Each political party seeks to minimize the maximal distance between all individual programs and the coalition program, enhancing the stability of the winning coalition as a whole:

$$f_3(\mathrm{S}) = \max_{i \in S} \left(\max_{j \in 1,2,\dots,m} \left| k_j^i - K_j^S \right| \right). \tag{6}$$

In this multicriteria model, each political party is motivated to minimize f_1, maximize f_2 and minimize f_3. These objective functions are defined within the interval [0, 1]. The preferences of each political party are represented by non-negative weights, subject to following conditions:

$$\sum_{j=1}^{3} v_j^i = 1 \;\text{ and }\; v_j^i \geq 0 \qquad \forall i \in \{1,2,\dots,n\}. \tag{7}$$

The utility derived by each political party from being part of the winning coalition is determined by maximizing the global objective function:

$$f(i, S) = \max_{S \in \Omega} (v_1^i \big(1 - f_1(i, S)\big) + v_2^i f_2(i, S) + v_3^i (1 - f_3(S)), \tag{8}$$

where Ω represents the set of all winning coalitions. Note that utility is non-transferable among political parties.

In the process of refining potential winning coalitions, we first identify those coalitions that are non-dominated. A coalition is classified as non-dominated if no member party could achieve a higher utility in any other feasible coalition where every other member also experiences at least the same or higher utility. This criterion ensures that the coalitions considered are stable and genuinely represent the interests of all included parties without favoring any subgroups.

If only one non-dominated coalition emerges from this analysis, it is automatically selected as the winning coalition. This reflects its unique position as the most stable and satisfactory arrangement among the possible configurations.

However, if multiple non-dominated coalitions are identified, we proceed to evaluate them further by calculating the average utility U(S) for each coalition. This average is defined as:

$$U(S) = \frac{\sum_{i \in S} f(i, S)}{|S|}. \tag{9}$$

Among these non-dominated coalitions, the one with the highest average utility U(S) is chosen as the optimal coalition. This method ensures that the final selected coalition maximizes the overall satisfaction and utility of its members, thereby promoting a stable and effective governance structure. It is important to highlight that this approach prioritizes coalition stability and individual satisfaction over simply achieving the minimum size required for a majority.

3 Application

The multicriteria voting game model was applied to the Chamber of Deputies of the Czech Parliament, which is the lower house of the bicameral legislature. This Chamber comprises 200 members, each serving a term of four years, with the government primarily accountable to this body. In the 2021 elections, deputies from seven political parties, organized into two individual parties and two coalitions, were elected (refer to Table 1). The SPOLU-coalition included the political parties TOP 09, ODS, and KDU-ČSL.

Political party	**Number of Deputies**
SPOLU	71
ANO	72
SPD	20
Piráti+STAN	37

Table 1 Chamber of Deputies, October 2021.

All 200 deputies were contacted via email and requested to complete a questionnaire comprising 16 questions focused on pivotal policy issues. The questions are detailed in Table 2.

Questions for deputies
1. What would you choose as the appropriate level of state participation in the economy and society?
2. What do you think is the appropriate role of government in social policy?
3. What is your opinion on the pension insurance model?
4. How should the government support entrepreneurship and economic growth?
5. What is your position on tax policy?
6. What is your preferred model for financing health care?
7. How should the government protect consumers?
8. How should the country approach international agreements and cooperation?
9. How should a country approach military issues?
10. How should a country approach migration?
11. How should a country address the issue of climate change?
12. What is your position on the adoption of the euro?
13. How should the government address the issue of teacher salaries and the keeping of young teachers?
14. What is your position on promoting gender equality?
15. How should the country approach the development and use of artificial intelligence?
16. Do you support marriage and adoption for same-sex couples?

Table 2 Questionnaire

Each question had a five-point answer. A total of 15 deputies from different parties responded to the questionnaire. The responses from each party and coalition were averaged. The resulting values (Table 3) were taken as the opinion of the political party/coalition on the topic.

Political Party/Coalition	**SPOLU**	**ANO**	**SPD**	**Piráti+STAN**
Question 1	0.35	0.31	0.50	0.33

Question 2	0.45	0.31	0.38	0.50
Question 3	0.40	0.38	0.38	0.42
Question 4	0.30	0.31	0.25	0.42
Question 5	0.45	0.25	0.14	0.42
Question 6	0.40	0.56	0.88	0.58
Question 7	0.45	0.56	0.50	0.50
Question 8	0.80	0.56	0.00	0.67
Question 9	0.85	0.69	0.50	0.75
Question 10	0.45	0.25	0.13	0.50
Question 11	0.45	0.44	0.00	0.50
Question 12	0.85	0.25	0.00	0.92
Question 13	0.40	0.38	0.50	0.67
Question 14	0.45	0.25	0.00	0.33
Question 15	0.75	0.56	0.75	0.58
Question 16	0.80	0.25	0.00	0.83

Table 3 Views on 16 political topics

In the subsequent phase of our analysis, we constructed all feasible winning coalitions within the Chamber of Deputies. We identified eight coalitions that each included 101 or more deputies. Lacking specific information on the weights v_j^i, we followed previous research and assigned them as (0.5, 0.3, 0.2) across all political parties for this experiment. These weights reflect the presumption that implementing the political program is the paramount objective for each party, whereas the stability of the coalition is considered the least critical goal. For each potential winning coalition *S*, we calculated the vector of the coalition program K^S, and for every party within these coalitions, we computed the values of $f_1(i,S)$, $f_2(i,S)$, $f_3(S)$ and $f(i,S)$. In this phase of our analysis, we assessed which coalitions within the Chamber of Deputies were dominated by others based on their utility scores, as shown in Table 4. A coalition is considered dominated if there exists an alternative coalition in which at least one member party has the option to achieve higher utility, denoted as $f(S)$, while all other parties from the original coalition also experience increased utility in this alternative arrangement. This concept underscores the strategic dynamics where political parties continuously evaluate their positions within various potential coalitions, aiming to maximize their respective gains while ensuring that their move does not adversely affect the utility of their usual coalition partners.

Winning Coalition	SPOLU	ANO	SPD	Piráti+STAN
SPOLU, ANO	0.64	0.64		
SPOLU, Piráti+STAN	0.82			0.68
ANO, Piráti+STAN		0.70		0.49
SPOLU, ANO, SPD	0.53	0.60	0.32	
SPOLU, ANO, Piráti+STAN	0.63	0.56		0.54
SPOLU, SPD, Piráti+STAN	0.64		0.23	0.55
ANO, SPD, Piráti+STAN		0.69	0.39	0.43
SPOLU, ANO, SPD, Piráti+STAN	0.54	0.53	0.30	0.46

Table 4 Utility of individual members of Winning Coalition

In our analysis to identify dominated coalitions, we assess whether any party within a coalition can achieve a higher utility in another feasible coalition where all other members also maintain or improve their utilities. This approach determines whether a coalition is dominated and thus less preferable.

For example, consider the coalition SPOLU, ANO, SPD, where SPOLU has a utility of 0.53, ANO has 0.60, and SPD has 0.32. This coalition is dominated by the simpler coalition SPOLU, ANO, where SPOLU has a higher utility of 0.64 and ANO also has 0.64. Both SPOLU and ANO achieve higher utilities in the SPOLU, ANO coalition compared to the SPOLU, ANO, SPD coalition, making the latter dominated as both can improve their position by shifting to a smaller coalition without SPD.

Once all feasible coalitions are evaluated for dominance, we identify those that are non-dominated. If only one non-dominated coalition exists, it automatically becomes the winning coalition. If multiple non-dominated coalitions are found, we then compute the average utility for each, *U(S),* to determine the optimal coalition. These calculations are detailed in Table 5 for clarity and completeness.

Winning Coalition	SPOLU	ANO	SPD	Piráti+STAN	*U(S)*
SPOLU, ANO	0.64	0.64			0.64
SPOLU, Piráti+STAN	0.82			0.68	0.75
ANO, Piráti+STAN		0.70		0.49	0.60
SPOLU, ANO, SPD	0.53	0.60	0.32		0.48
SPOLU, ANO, Piráti+STAN	0.63	0.56		0.54	0.58
SPOLU, SPD, Piráti+STAN	0.64		0.23	0.55	0.47
ANO, SPD, Piráti+STAN		0.69	0.39	0.43	0.50
SPOLU, ANO, SPD, Piráti+STAN	0.54	0.53	0.30	0.46	0.46

Table 5 Utility of Winning Coalition *U(S)*

Our findings indicate that the coalition of SPOLU and Piráti+STAN, which aligns with the actual governing coalition in the Czech Republic, emerges as the only non-dominated coalition. To confirm the robustness of this result, we computed *U(S)* for each coalition. The calculations confirm that this coalition not only avoids dominance but also achieves the highest average utility, thus being optimal within our model framework.

4 Conclusion

In this paper, we applied the multicriteria voting game model to the Czech Republic's Chamber of Deputies, analyzing the dynamics of political objectives within coalition formation. This model offers a structured approach to assess how individual political party programs align with a coalition program, the influence each party holds within the coalition, and the overall stability of the coalition. Our empirical application, based on responses to a 16-question survey spanning various policy issues, underscores the model's practical utility and relevance in a real-world parliamentary setting.

The findings indicate that the optimal coalition, as predicted by the model, not only maximizes the collective utility of its members but also corresponds with the actual governing coalition. This congruence between theoretical predictions and real-world outcomes highlights the model's potential as a predictive tool for political analysts and strategists.

The practical applications of such a model extend beyond academic interest, providing insights that could guide political negotiations and strategy development. By quantifying trade-offs between program alignment, power maximization, and stability, our model equips political actors with a nuanced understanding of the dynamics in coalition politics.

Future research could explore refinements to the model, such as incorporating dynamic changes in party programs or examining external influences like public opinion and economic conditions on coalition stability. Applying the model to other parliamentary systems could also offer comparative insights and enhance its generalizability.

In conclusion, the multicriteria voting game model is a valuable addition to the toolkit of political science, offering theoretical insights and practical applications that deepen our understanding of coalition dynamics in parliamentary democracies.

References

Binmore K (1992) Fun and Games: A Text on Game Theory. D. C. Heath and Company, Lexington

Dlouhý M, and Fiala P (2015) Teorie ekonomických a politických her. Oeconomica, Prague

Dlouhý M, Tichá M (2016) Multicriteria Voting Game and its Application. In: 34th International Conference on Mathematical Methods in Economics 2016. Conference Proceedings. TU Liberec, Liberec, pp 160-164

Straffin P D (1994) Chapter 32 - Power and Stability in Politics. In: Handbook of Game Theory with Economic Applications - Volume 2 (Aumann, R., and Hart, S., eds.). North-Holland, Amsterdam

Comparison of Python Metaheuristic Packages

Ing. Vojtěch Vávra[1]

Abstract. Metaheuristic approaches are utilized to find sub-optimal solutions within a reasonable timeframe. This is crucial for NP-hard problems such as the traveling salesman problem, the vehicle routing problem and the knapsack problem. However, identifying the appropriate software to execute such algorithms is not straightforward. This paper presents a comprehensive study aimed at identifying a suitable package in the programming language Python. Python is one of the most widely used programming languages worldwide and is employed daily by numerous companies. The sheer number of packages available can be overwhelming, making it challenging to select the right tool for a given problem.

The objective of this paper is to locate and compare such packages, preselect suitable ones and determine the best package or packages. The criteria for decision-making include: first, the number of algorithms implemented; second, flexibility, customization and tuning capabilities; third, performance; fourth, quality of documentation; fifth, the learning curve in relation to the knowledge of AI; sixth, maintenance; seventh, community support and finally, the overall health of the package.

Keywords: Python, metaheuristics, DEAP, MEALPY, NiaPy, Opytimizer, PyGmo

JEL Classification: C61
AMS Classification: 90-08

1 Introduction

In the paper are presented metaheuristic packages for Python. Since this topic is not that widely popular, there are no such preferences for using one of them among the community, such as libraries numpy or pandas. The aim of this paper is to present such libraries, explore them and test them using optimization functions to determine which is the best of them. To scale problems, one can increase the number of dimensions, making it more challenging to find an optimal or suboptimal solution.

To test the AI's ability to utilize the packages, the following tools were employed: OpenAI ChatGPT, Codeium (Python copilot), and Microsoft Bing AI ('Copilot'). Documentation, code samples, and examples were extensively studied to optimize the calculation process for the proposed problems.

The next chapter presents the packages along with their user experience, the metaheuristics used for comparison and the problems computed in the third chapter. In the third chapter, experimental calculations are provided and in the final chapter, the paper concludes.

2 Methodology

2.1 Metaheuristics packages

Below is an unordered list of the studied packages for Python considered in this study. In brackets are cited the documentation or publications of individual packages.

- DEAP – (Fortin et al., 2012),
- MEALPY – (Thieu et al., 2022),
- NiaPy – (Vrbančič et al., 2018),
- Opytimizer – (De Rosa et al., 2019),
- PyGmo – (Biscani et al., 2020),

The first view on packages is provided by (Snyk Ltd., 2024) in Table 1, where all the mentioned packages and their overall health situation are covered. Packages **DEAP** and **NiaPy** have the best score and are mostly used by the community.

[1] Prague University of Economics and Business, Department of Econometrics, square W. Churchill 1938, 130 67 Prague 3, vavra.vojtech@vse.cz.

Package	Health	Security	Popularity	Service	Community
DEAP	77	Required	Popular	Healthy	Active
MEALPY	61	Required	Small	Sustainable	Sustainable
NiaPy	78	Alright	Limited	Healthy	Active
Opytimizer	50	Required	Small	Inactive	Sustainable
PyGmo	62	Alright	Small	Inactive	Sustainable

Table 1 Overall health of packages (observed in January 2024).

DEAP

The DEAP package allows the creation of necessary data types instead of having strictly defined inputs. It is designed for the user to customize everything according to their needs rather than conforming to system requirements. Key operations are the creator and toolbox. Implementation and experimentation with evolutionary algorithms are among the strengths of this package. It enables the creation of genetic algorithms (GA), genetic programming (GP), multi-objective optimization, parallel computations in evolutionary processes and promises easy integration with other tools. The downside of this package is that the number of metaheuristics is limited. It is also possible to use implemented functions to create algorithms on the user's behalf.

AI had huge difficulties generating functional scripts, however, at least the genetic algorithm was successfully generated. At first, it was also confused about the availability of metaheuristics.

MEALPY

The MEALPY package is among the latest and contains the largest number of metaheuristics, according to its authors. It has a well-organized database of metaheuristics, which are categorized into the following folders:

- bio_based – nature-inspired,
- evolutionary_based – evolution-based,
- human_based – human-inspired,
- math_based – mathematics-based,
- music_based – music-inspired,
- physics_based – inspired by physical laws,
- swarm_based – swarm-based,
- system_based – system-based (e.g., water circulation system).

The authors promise to implement every metaheuristic that will be proposed to them. The best practice is to combine calculations with a tool that allows users to suppress print statements, since every iteration of the algorithm returns output.

AI had difficulties generating functional scripts, especially before the update in autumn 2023. However, after correcting file locations and a deeper examination of the package, all tests were successfully performed.

NiaPy

NiaPy is a package that is used to implement algorithms that are inspired by nature. Although these algorithms are usually difficult to implement, this package makes them simple enough to use. It also offers a straightforward way to compare the outcomes of various algorithms that influenced by nature.

The documentation includes illustrative examples that help users understand how to work with the library. Although only two sample examples and a guide for the first use of the library are available, they provide a useful foundation for beginners, allowing them to quickly start working with the package.

Only two but informative examples are included in the documentation. These are considered sufficient for beginners to understand the basics and enable them to work with the package.

AI was able to generate functional scripts with reservations.

Opytimizer

The Opytimizer package promises easy usage of algorithms, including the use of agents and search spaces. The main goal of the package, in the words of the authors, is to "optimize things." It assists in creating custom optimization algorithms, creating or using pre-built tasks for solutions and combining various strategies to find solutions. However, the package specializes in minimization, which may require users to convert their problems.

Additionally, the documentation appears clear and user-friendly for first-time library users. The main components of the library are core, functions, math, optimizers, spaces, utils and visualization.

Instead of providing sample examples, the documentation describes every function that has been developed from a programming standpoint. As a result, there is a point of reference in case a user experiences problems with a certain function. However, the examples folder on GitHub should be consulted if a sample example is required.

AI-generated scripts ran properly for the first time.

PyGmo

The PyGmo package is based on the idea of parallel optimization and provides a unified interface for optimization algorithms. The library aims to facilitate the implementation of metaheuristics and leverage the mentioned parallel optimization. The package can be used to solve constrained, unconstrained, single-objective, multi-objective, continuous and integer optimization problems, both stochastic and deterministic, as stated by the authors in the documentation. However, the package is not available for installation on Windows using the pip command. It is only available on Linux. Therefore, installation requires the Anaconda manager for Windows users. Additionally, the package can be integrated into R and MATLAB environments, as it is written in C++ under the name PaGmo.

AI managed to create working scripts, however, there were several problems with identifying the right approach for the genetic algorithm and lying about the implementation of simulated annealing that was implemented in another instance.

2.2 Proposed Metaheuristics and test functions

Proposed and tested algorithms by each library are listed below. From three different classes, basic versions of metaheuristics were selected that are widely used and well known.

- Class of evolutionary algorithms: **Genetic Algorithm (GA)**, (Holland, 1962)
- Class of physics algorithms: **Simulated Annealing (SA, SIAM)**, (Kirkpatrick, 1983)
- Class of nature-inspired algorithms: **Particle Swarm Optimization (PSO)**, (Kennedy et al., 1995)

For experimental calculation, the following problems are solved: the Rastrigin function, the Ackley function and the Rosenbrock function. Functions are described as minimization problems with a known optimal solution and given boundaries. Proposed problems are defined by the number of given dimensions, which are denoted as $(i = 1,2, \dots, n)$.

Rastrigin function (1) is defined as follows:

$$f(x) = A \cdot n + \sum_{i=1}^{n} [x_i^2 - A \cdot \cos(2\pi x_i)], \quad (1)$$

Where A is usually set as 10. To find the optimal global solution, it is required to increase exploration over exploitation. For two dimensions, the function is visualized in **Figure 1**.

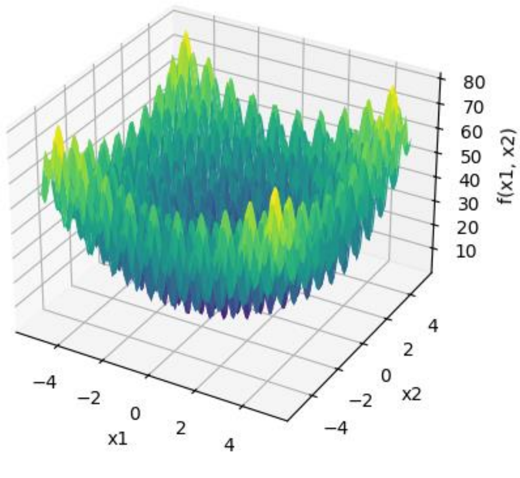


Figure 1 Rastrigin function for 2 dimensions.

Ackley function (2) is defined as follows:

$$f(x) = -a \cdot \exp\left[-b\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right] - \exp\left[\frac{1}{n}\sqrt{\sum_{i=1}^{n}\cos(cx_i)}\right] + a + \exp[1], \tag{2}$$

Where a is usually defined as 20, b as 0.2 and c as 2π. To find an Ackley global optimal solution, it is required to find a balance between exploitation and exploration. For two dimensions, the function is visualized in **Figure 2**.

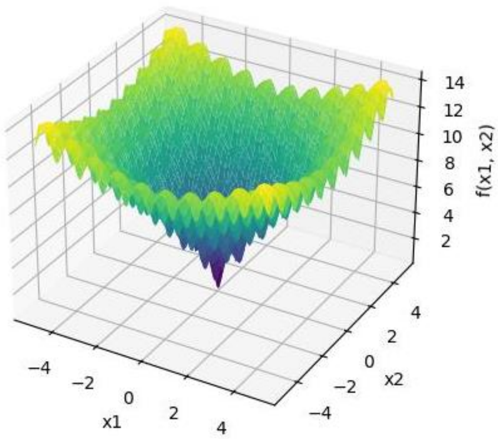


Figure 2 Ackley function for 2 dimensions.

Rosenbrock function (3) is defined as follows:

$$f(x) = \sum_{i=1}^{n-1}[100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2]. \tag{3}$$

To find a global solution, exploitation is the key parameter to be used. For two dimensions, the function is visualized in **Figure 3**.

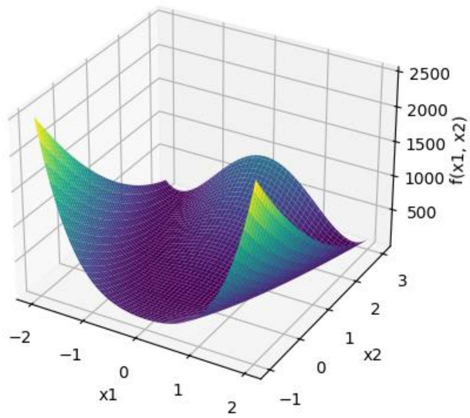


Figure 3 Rosenbrock function for 2 dimensions.

To compare functions, **Table 2** contains information about the allowed number of dimensions, recommended boundaries and value of the global minimum $f(x^*)$ with a vector of decision variables x^*.

Function	Dimensions	Boundaries	Global minimum
Rastrigin	$n \in \mathbb{Z}$	$\forall i \in \langle -5.12, 5.12 \rangle$	$f(x^*) = 0$, at $x^* = (0, \ldots, 0)$
Ackley	$n \in \mathbb{Z}$	$\forall i \in \langle 32.768, 32.768 \rangle$	$f(x^*) = 0$, at $x^* = (0, \ldots, 0)$
Rosenbrock	$n \in \mathbb{Z}$	$\forall i \in \langle -5, 10 \rangle$	$f(x^*) = 0$, at $x^* = (1, \ldots, 1)$

Table 2 Overview of functions with restrictions and global best objective.

3 Computational experiments

In the previous chapter, methods to quantitatively measure the effectiveness of packages were introduced. Only five packages have the desired algorithms: **DEAP**, **MEALPY**, **NiaPy**, **Opytimizer**, and **PyGmo**. There is one important notification. The **PyGmo** library does not contain the SA algorithm in its basic form. However, Corana's version of the SA algorithm is implemented. These numbers are presented, but they are considered deprecated (grey color).

Each package was examined to deliver objective functions for 10 repetitions of calculation. Here, $\bar{Z}$ denotes the mean of all objective function values and Z^* represents the best obtained value after 10 calculations.

The trend of time calculation between optimization functions is from fastest to slowest, as follows: Rastrigin, Ackley and Rosenbrock.

The time to find the final solution for each package significantly differs. Average time over 10 repetitions is considered to present the following order: **NiaPy** (0.02), **PyGmo** (0.07), **DEAP** (2.58), **MEALPY** (4.16) and **Opytimizer** (8.56). The numbers in brackets show the number of seconds needed to compute PSO for 20 dimensions. The calculation time is only considered for an optimization step of the calculation. Since this trend is observed in all cases, only this informative list is proposed.

For an overview of Rastring function results, see **Table 3**. The best results are obtained by packages **PyGmo**, **Opytimizer** and **MEALPY**.

Dim.	**Met.**	**DEAP**		**MEALPY**		**NiaPy**		**Opytimizer**		**PyGmo**	
		$\bar{Z}$	Z^*	$\bar{Z}$	Z^*	$\bar{Z}$	Z^*	$\bar{Z}$	Z^*	$\bar{Z}$	Z^*
5	GA	0.21	0.00	0.11	0.03	28.78	9.37	0.00	0.00	0.00	0.00
	SA	11.71	8.11	3.27	1.57	19.89	12.15	12.53	8.21	1.29	0.00
	PSO	7.42	4.96	5.87	0.99	17.54	12.84	7.76	1.99	3.03	1.76
10	GA	15.78	0.99	0.16	0.05	100.09	79.43	0.00	0.00	0.00	0.00
	SA	62.73	54.81	33.09	25.08	73.81	59.65	58.78	40.85	1.79	1.00
	PSO	42.68	26.54	24.58	14.92	72.41	50.20	24.21	10.99	22.60	15.05
20	GA	132.91	119.95	0.41	0.22	228.27	206.02	1.12	0.19	0.03	0.02
	SA	194.37	178.32	127.07	96.13	214.05	185.50	188.85	171.63	5.56	3.98
	PSO	143.69	122.44	96.06	72.21	177.35	142.15	68.60	30.18	93.01	80.83

Table 3 Experiments – Rastrigin function.

For an overview of Ackley function results, see **Table 4**. In this example, the best calculations are observed for package **MEALPY**. Overall, the worst package is NiaPy.

Dim.	**Met.**	**DEAP**		**MEALPY**		**NiaPy**		**Opytimizer**		**PyGmo**	
		$\bar{Z}$	Z^*	$\bar{Z}$	Z^*	$\bar{Z}$	$\bar{Z}$	$\bar{Z}$	Z^*	$\bar{Z}$	Z^*
5	GA	0.00	0.00	0.04	0.02	15.51	13.04	0.00	0.00	0.09	0.05
	SA	9.39	6.74	0.45	0.27	18.88	16.11	15.27	11.87	0.04	0.02
	PSO	1.19	0.38	0.33	0.00	8.31	4.22	0.40	0.00	0.74	0.00
10	GA	0.00	0.00	0.04	0.00	18.93	17.87	0.01	0.00	0.13	0.09
	SA	16.97	16.10	1.65	1.38	19.89	19.00	18.71	16.97	0.08	0.02
	PSO	3.10	2.54	1.36	0.02	13.45	11.38	4.07	1.81	3.75	2.40
20	GA	19.54	19.93	2.86	0.04	20.08	19.57	0.05	0.04	0.18	0.11
	SA	19.18	18.55	0.06	2.71	20.39	20.01	19.71	19.30	0.13	0.08
	PSO	4.29	3.82	3.92	2.32	17.01	15.09	8.61	7.02	6.27	5.30

Table 4 Experiments Ackley function.

For an overview of Rosenbrock function results, see **Table 5**. This function was challenging for most of the packages and metaheuristics, especially for PSO and 20 dimensions. Also, the NiaPy package was confused for genetic algorithm. The best results are again split between **PyGmo**, **Opytimizer** and **MEALPY**.

Dim.	Met.	DEAP		MEALPY		NiaPy		Opytimizer		PyGmo	
		$\bar{Z}$	Z^*	$\bar{Z}$	Z^*	$\bar{Z}$	$\bar{Z}$	$\bar{Z}$	Z^*	$\bar{Z}$	Z^*
5	GA	0.69	0.00	1.33	0.20	1768.97	362.67	1.39	0.27	1.61	0.23
	SA	0.13	0.02	4.11	1.34	13.02	1.06	0.31	0.15	2.66	0.41
	PSO	7.54	4.17	1.23	0.00	228.63	25.31	20.40	0.06	2.35	0.00
10	GA	4.28	0.02	23.06	2.31	54234.80	19299.55	1.07	0.27	8.71	3.33
	SA	3.54	1.99	57.55	14.97	266.66	109.32	5.26	4.20	3.53	0.22
	PSO	223.99	166.43	54.27	6.90	5179.49	2002.55	183.97	4.39	81.96	16.72
20	GA	50.92	2.27	52.02	10.40	331755.77	154841.96	16.78	2.62	52.26	12.66
	SA	59.02	45.15	298.68	140.73	997.70	604.98	29.43	26.71	31.00	0.90
	PSO	2633.4	1428.4	785.1	365.0	59878.5	38236.4	3109.6	1236.1	3617.7	1749.5

Table 5 Experiments Rosenbrock function.

4 Conclusion

All the presented packages were tested with the same tuned parameters and different results were observed for all of them. Package **NiaPy** was the fastest but achieved the worst results. Package **DEAP** was very successful for smaller problems. However, when the number of dimensions was increased, calculation performance was worse in comparison to other better packages. Package **Opytimizer** has decent objective values. However, the time to compute algorithms is too high and is overrun by other packages.

The two best packages are MEALPY and PyGmo. Package **MEALPY** has an extensive number of implemented algorithms and decent health care. However, its computational time is slower than that of its competitor, PyGmo. The **PyGmo** package excels in time and computational results, with exceptions. The downside is limited service.

Acknowledgements

The research is supported by the Internal Grant Agency, Faculty of Informatics and Statistics, Prague University of Economics and Business – project no. F4/10/2024.

References

Biscani, F., Izzo, D. (2020) A parallel global multiobjective framework for optimization: pagmo. Journal of Open Source Software, 5(53), 2338. https://doi.org/10.21105/joss.02338

Fortin F.-A., De Rainville F.-M., Gardner M.-A., Parizeau M., Gagné, C. (2012) DEAP: Evolutionary Algorithms Made Easy. Journal of Machine Learning Research, 13, 2171–2175. https://www.jmlr.org/papers/volume13/fortin12a/fortin12a.pdf

Holland J. H. (1962). Outline for a Logical Theory of Adaptive Systems. Journal of the ACM, 3(4), 297-314. https://doi.org/10.1145/321127.321128

Kennedy J., Eberhart R. (1995). Particle Swarm Optimization. In Proceedings of IEEE International Conference on Neural Networks (ICNN), Perth, Australia, Vol. 4, pp. 1942-1948.https://doi.org/10.1109/ICNN.1995.488968

Kirkpatrick S., Gelatt Jr. C. D., Vecchi M. P. (1983). Optimization by Simulated Annealing. Science, 220(4598), 671-680. https://doi.org/10.1126/science.220.4598.671

De Rosa G. H., Rodrigues D., Papa, J. P. (2019) Opytimizer: A Nature-Inspired Python Optimizer. Available at arXiv:1912.13002. https://doi.org/10.48550/arXiv.1912.13002

Snyk Ltd. (2024) Snyk Open Source Adviser. 2024-01. https://snyk.io/advisor/python.

Thieu, N. V., Mirjalili, S. (2022) MEALPY: a Framework of The State-of-The-Art Meta-Heuristic Algorithms in Python. Zenodo, v2.5.0. https://doi.org/10.5281/zenodo.6684223

Vrbančič G., Brezočnik L., Mlakar U, Fister D., Fister Jr. I. (2018) NiaPy: Python microframework for building nature-inspired algorithms. Journal of Open Source Software, 3(23). https://doi.org/10.21105/joss.00613

Refining Fourier Approach with Constrained Parameter Estimation and Penalizing Seasonal Distortions

Lukáš Veverka[1]

Abstract. This study addresses the challenge of determining initial parameters in a constrained space for Fourier transformations applied to decompose time series data of media investments, focusing on identifying the underlying seasonal components at minimal levels. Recognizing that observed peaks of any business key performance indicators (KPIs) can be attributed to various external factors such as special events and media activities. The research modifies the Sum of Squared Residuals (SSR) methodology, which aims to penalize overestimations in the form of negative residuals. Thus, the distribution of residuals becomes highly skewed. The maximum likelihood method is used to get likelihood, and the Akaike information criterion (AIC) is used to evaluate the appropriate order of Fourier transformation. Based on the outputs, it is possible to provide suitable initial parameters for more complex regression models based on non-linear optimization.

Keywords: Fourier transformation, Constrained parameter estimation, Seasonal decomposition, Data-driven marketing

JEL Classification: C22, M31
AMS Classification: 91B84

1 Introduction

Being able to estimate the initial parameters for Fourier transformation and its correct order is crucial for more advanced modelling business KPIs such as webpage visits, transactions or sales. This paper focuses on the initial parameter estimation within the Fourier transformation framework, which is used to identify seasonality in the data. The motivation for this research comes from its application in more complex models that incorporate various factors influencing KPIs, including highly non-linear behaviour of media investments and special events. In such models, the search space for finding an optimal solution significantly increases. Fourier transformation helps isolate the baseline seasonal patterns in these models, ensuring that variations beyond this baseline are attributed to other influencing factors. Knowing at least the initial parameters for the seasonality decreases the complexity of the more complex models. Given the complexity of these models, all Fourier parameters are constrained within a 0-1 range, and the target variable is also scaled to this range.

This study extends the work of Veverka, 2022 where the research was focused on optimizing tuning parameters for nested algorithms, particularly non-linear models where traditional variable selection methods in linear regression fall short. The binary version of a genetic algorithm was utilized to address the challenges posed by the varying ranges of estimated parameters for a single variable. The research highlighted the sensitivity of both genetic algorithms and non-linear optimization to parameter settings, emphasizing the need for optimal configurations to achieve the best outcomes. The optimal settings were determined through repeated cross-validation, considering different combinations for various time categories.

Fourier transformation is widely recognized for its capacity to decompose time series data into frequency components. It is valuable for analyzing non-stationary and nonlinear data, such as company sales or web page visits. The Fourier Decomposition Method (FDM) offers an adaptive approach, decomposing data into Fourier Intrinsic Band Functions (FIBFs) to reveal the data's intrinsic time-frequency-energy distribution, thus providing insights into its underlying structure which was shown by Singh et al, 2017. In practical applications, Fourier analysis has been employed for demand forecasting in industries with complex demand patterns, such as the fashion sector. This methodology leverages the Fourier method for a more effective forecasting approach, demonstrated through a comparison with traditional techniques like moving averages and exponential smoothing, using historical sales data from a fashion company done by Fumi et al, 2013.

Building on this, this paper explores the methodology and advancements related to the Fourier transformations in business KPI modelling. The Fourier transformation is very useful for estimating seasonality, which can serve as the baseline for further estimating additional factors affecting KPIs, such as media investment and special events.

[1] University of Economics, Prague, Department of Econometrics, Winston Churchill Square 4, CZ13067 Prague, Czech Republic, lukas.veverka@vse.cz

2 Analyzed Data Sample

The anonymised dataset contains information about a company in the CEE region selling clothes. The data are from 27.10.2018 to 27.12.2022 with a daily frequency (resulting in the total number of observations: $T = 1523$). The target variable is the number of transactions, which is obtained from Google Analytics. Moreover, data derived from the date, such as the day of the year, trend or national holidays, are used.

Furthermore, the data contains information about the daily media investments in TV, online, print and radio. Those data are sourced from Nielsen Ad Monitoring. All media investments incorporated into the models are measured in gross prices (CZK). In addition, there are COVID-related data and economic indicators. *Covid positive* variable represents the total amount of people with a positive test based on the data from the National Health Statistics Institute. *Lockdown* is a boolean variable indicating days when the national government has ordered a lockdown, and people were unable to go shopping physically, but they were still able to buy goods online. Data regarding economic indicators are obtained from the Czech Statistical Office. Specifically, Balances of Confidence Indicators are presented as *Consumer confidence indicator (CCI)*, Average consumer fuel prices in CR as *Energy crisis indicator* and Consumer price index according to ECOICOP as *Inflation*. However, in this study, just the target variable (Transactions) and the day of the year (derived from the date) are used.

Table 1 Sample data displaying five rows with selected columns. Note that the full dataset contains additional columns not shown here.

Transactions	Day of year	Consumer confidence indicator (CCI)	Energy crisis indicator	Inflation	TV	Online	Print	Radio	Covid positive	Lockdown
87	300	38.75	33.49	106	0	0	0	0	0	FALSE
134	301	38.75	33.49	106	0	0	0	0	0	FALSE
164	302	38.75	33.54	106	0	0	0	25,016	0	FALSE
168	303	38.75	33.54	106	0	0	0	26,888	0	FALSE
188	304	38.75	33.54	106	0	0	0	24,968	0	FALSE

3 Methodology

3.1 Maximum Likelihood Estimation with Biased Sum of Squared Residuals (SSR)

This section outlines the methodological framework for identifying the baseline seasonality in business KPIs using Fourier transformation and a modified regression approach.

Let us begin with the standard linear regression model, represented as:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \cdots + \beta_j x_{tj} + u_t, \tag{1}$$

where y_t is the dependent variable (e.g. any KPI), x_{tj} are the independent variables, β_j are the coefficients to be estimated, and u_t represents the error term. Under typical assumptions, the error term u_t is normally distributed, i.e., $u_t \sim N(0, \sigma^2)$.

To identify the baseline (seasonality) of the KPI, the fitted values should ideally represent the lower boundary of the series, capturing the minimal seasonal pattern. Hence, the residuals ($\hat{u}_t = y_t - \hat{y}_t$) should predominantly lie in the non-negative part of the distribution, ensuring the baseline is at the bottom of the series.

To achieve this, the SSR function is modified to penalize negative residuals more heavily:

$$SSR = \sum_{t=1}^{T} \begin{cases} \hat{u}_t^2 & \text{if } \hat{u}_t < 0 \\ \hat{u}_t^4 & \text{if } \hat{u}_t \geq 0 \end{cases} \tag{2}$$

Given that the residuals ($\hat{u}_t$) and the target variable are on a 0-1 scale, this modified SSR function penalises the positive residuals with a higher power (4 in this case, though this can be adjusted as a tuning parameter) and therefore reduces the impact of positive residuals while still striving to keep the fitted values close to the observed values.

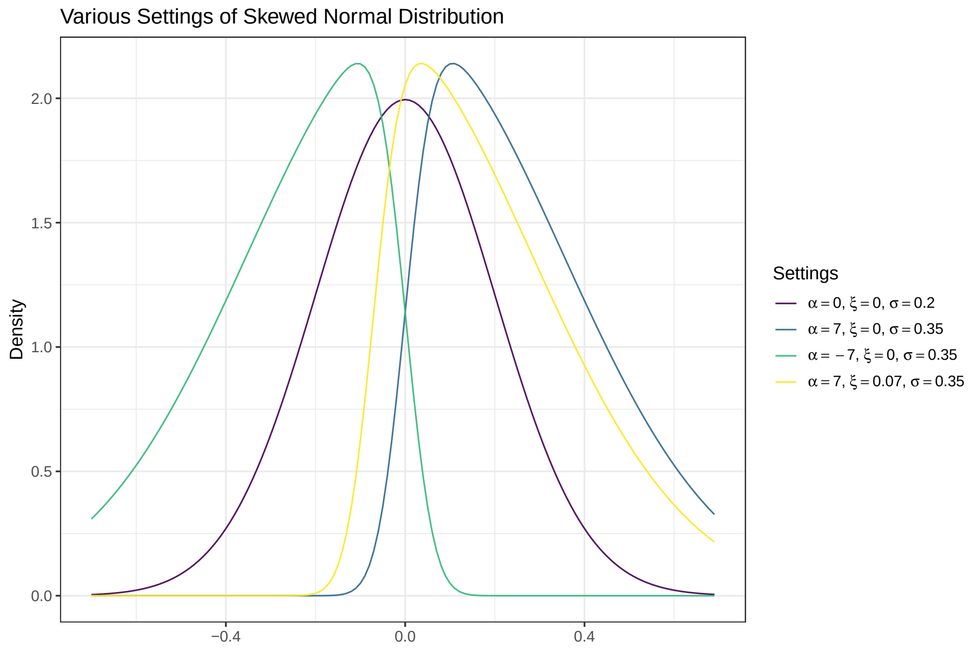


Figure 1 Probability density function under different settings of parameters for skewed normal distribution.

Introducing this penalization skews the distribution of residuals, deviating from the normality assumption. Therefore, it is assumed the residuals follow a skewed normal distribution. The probability density function of a skewed normal distribution is given by:

$$f(\hat{u}_t) = \frac{2}{\sigma}\phi\left(\frac{\hat{u}_t - \xi}{\sigma}\right)\Phi\left(\alpha\left(\frac{\hat{u}_t - \xi}{\sigma}\right)\right), \tag{3}$$

where ϕ is the probability density function of standard normal distribution, and Φ is the cumulative distribution function of standard normal distribution. The parameter α is known as the skewness parameter. It controls the asymmetry of the distribution. When $\alpha = 0$, the distribution reduces to the standard normal distribution, which is symmetric. Positive values of α introduce positive skewness (right skew), while negative values introduce negative skewness (left skew). The parameter ξ represents the location parameter, similar to the mean in a normal distribution. It shifts the entire distribution left or right along the x-axis. The parameter σ is the scale parameter analogous to the standard deviation in a normal distribution. It stretches or compresses the distribution, affecting the spread of the data. A larger σ results in a wider distribution, while a smaller σ results in a narrower distribution. In Figure 1, it is possible to see the distribution behaviour under different settings of each parameter.

This modified assumption allows us to better model the skewed nature of the residuals introduced by the penalization. The maximum likelihood method is employed to estimate the parameters in this model. The likelihood function for the model containing skewed normal distributed residuals can be expressed as:

$$L(\boldsymbol{\beta}, \alpha, \xi, \sigma) = \prod_{t=1}^{T} \frac{2}{\sigma}\phi\left(\frac{\hat{u}_t^* - \xi}{\sigma}\right)\Phi\left(\alpha\left(\frac{\hat{u}_t^* - \xi}{\sigma}\right)\right), \tag{4}$$

where α, ξ, σ are the parameters (described earlier in the text) for skewed normal distribution to be estimated, $\boldsymbol{\beta}$ is a vector of β_j parameters to be estimated and its relation to the likelihood function is through the modified residuals $\hat{u}_t^*$ which follows this function:

$$\hat{u}_t^* = \begin{cases} (y_t - \mathbf{x}_t\boldsymbol{\beta})^2 & \text{if } (y_t - \mathbf{x}_t\boldsymbol{\beta}) < 0 \\ (y_t - \mathbf{x}_t\boldsymbol{\beta})^4 & \text{if } (y_t - \mathbf{x}_t\boldsymbol{\beta}) \geq 0 \end{cases}, \tag{5}$$

where y_t is the target variable and therefore $(y_t - \mathbf{x}_t\boldsymbol{\beta})$ corresponds to the common residuals and based on their negativity/positivity different penalization parameter is applied.

To facilitate the optimization, the likelihood function is transformed into the negative log-likelihood function:

$$-\log L(\boldsymbol{\beta}, \alpha, \xi, \sigma) = -\sum_{t=1}^{T} \log \frac{2}{\sigma} \phi \left(\frac{\hat{u}_t^* - \xi}{\sigma} \right) \Phi \left(\alpha \left(\frac{\hat{u}_t^* - \xi}{\sigma} \right) \right) \tag{6}$$

The negative log-likelihood function is minimized using standard numerical non-linear optimization algorithms, ensuring the parameters are estimated to capture the KPI's underlying seasonality best. This methodological approach provides a more accurate baseline for KPIs and simplifies the estimation of subsequent complex models that incorporate various factors influencing business metrics.

3.2 Fourier Transformation for Time Series Decomposition

Fourier transformation is a mathematical technique that is commonly used in econometrics to decompose time series data into its underlying frequency components. This approach is useful for identifying the presence of cyclical patterns or seasonality in the data. The basic idea behind Fourier transformation is that any time series can be represented as a combination of sine and cosine waves of varying frequencies, with each wave having its own amplitude and phase. By applying the Fourier transformation to a time series, it is possible to identify the frequency components present in the data and estimate the amplitude and phase of each component. The Fourier transformation can be written as:

$$f(t) = \beta_0 + \sum_{n} (A_n cos(n\omega t) + B_n sin(n\omega t)), \tag{7}$$

where β_0 is the intercept, A_n and B_n are the amplitude of the sine and cosine waves at frequency $n\omega$ and when plugged in the regression model (i.e. in Equation 5) both A_n and B_n are general β parameters. ω is the angular frequency ($2\pi C$), where C is the duration of one complete cycle (inverse of the frequency) and t is the time variable. Finally, n is an integer representing the Fourier transformation's order.

Determining the optimal Fourier order is crucial for accurately capturing the seasonality in the time series data without overfitting. The Akaike Information Criterion is utilized to achieve this. The AIC is a widely used measure for model selection that balances model fit and complexity. It is defined as:

$$\text{AIC} = 2k - 2\ln(L), \tag{8}$$

where k is the number of parameters in the model, and L is the likelihood function as stated in the Equation 4. To find the optimal Fourier order, the AIC is calculated for models with different orders of Fourier components and select the model with the lowest AIC value.

Fourier transformation helps us identify the dominant periodicities in the data by decomposing the time series into its frequency components. These periodic components can then be analyzed individually to understand the underlying seasonal patterns. In the context of business KPIs, this decomposition allows us to isolate the baseline seasonality, which is crucial for accurate modelling.

4 Results

This section presents the results of the analysis based on the methodology described previously. The Maximum Likelihood Estimation was applied using a skewed normal distribution and incorporated the Fourier transformation to capture seasonality in the time series data.

The Fourier transformation orders were tested, ranging from 1 to 30, to determine the optimal order based on the Akaike Information Criterion. The results indicated that the optimal Fourier order was 18, as it minimized the AIC value. The parameters for the skewed normal distribution resulted in $\alpha = 5.39$, $\xi = 0.034$, $\sigma = 0.08$.

Figure 2 illustrates the distribution of residuals obtained from the normal and skewed estimations. The normal estimation assumes residuals are normally distributed with a mean of zero, resulting in approximately centred around zero residuals. In contrast, the skewed estimation, which incorporates bias and skewed normal distribution, shows predominantly non-negative residuals, effectively placing the fitted values at the lower boundary of the series. This demonstrates that the skewed estimation better captures the baseline seasonality.

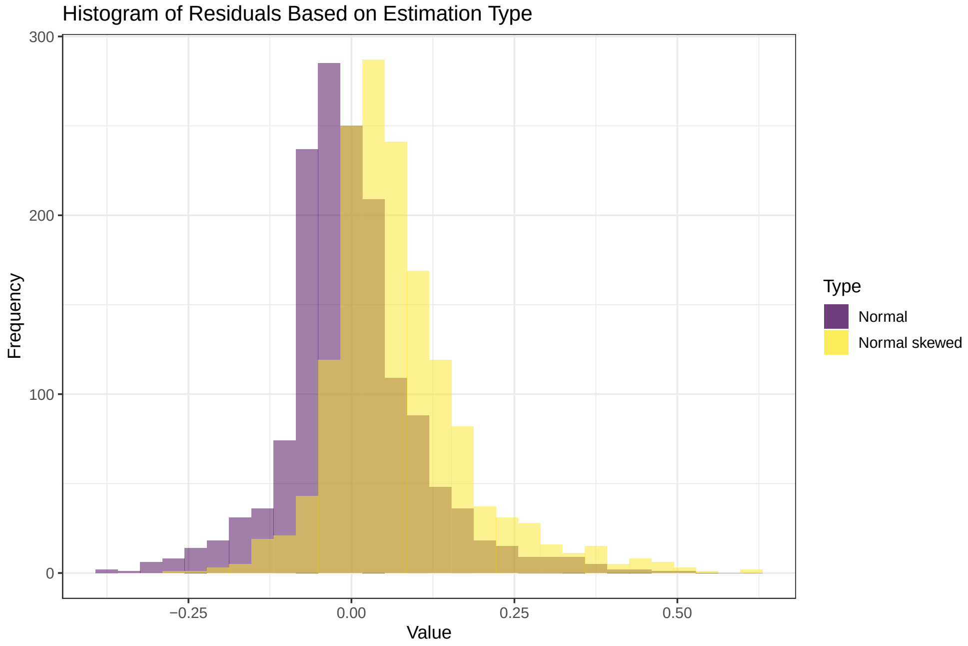


Figure 2 The distribution of residuals with and without penalization for negativeness.

Figure 3 presents the time series data over three years, with daily observations. The normal estimation, shown in the figure, places the fitted values approximately in the middle of the series, consistent with the assumption that the residuals have a zero mean. However, the skewed estimation positions the fitted values near the bottom of the series. Although not perfect (since there are still some observations below the fitted values) this positioning can be adjusted by tuning the penalization parameter for residuals (e.g., adjusting the power from 4 in Equation 2 to a different value). For the purposes of this research, the current results are satisfactory.

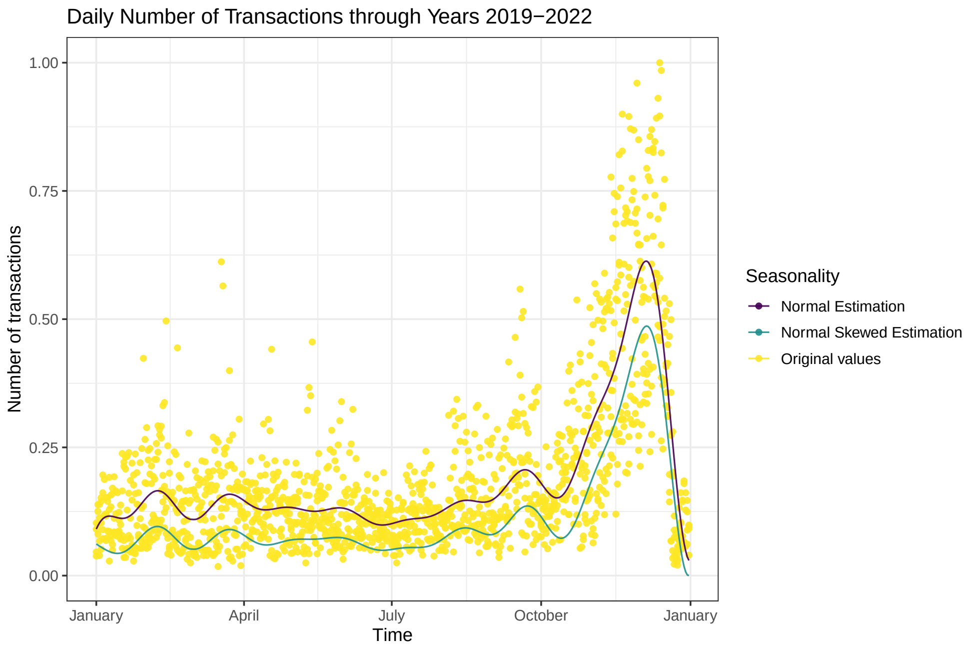


Figure 3 Transactions scaled to 0-1 and the fitted values based on normal and skewed estimation.

The results indicate that using a skewed normal distribution and optimizing the Fourier order based on AIC provides a more accurate baseline for seasonality in business KPIs. This approach effectively captures the underlying patterns in the data, providing better initial parameters for more advanced models and, therefore, their faster estimation.

5 Conclusion

This study estimated initial parameters for Fourier transformations to model business KPIs such as webpage visits, transactions or sales. The baseline patterns were isolated by identifying the underlying seasonality within a constrained 0-1 range, while the remaining variations were attributed to factors like media investments or special events. The SSR was modified to penalize negative residuals and align fitted values closer to minimal KPI levels. Using Maximum Likelihood Estimation with a skewed normal distribution and the Akaike Information Criterion, the Fourier order of 18 was found to capture the seasonality optimally. The results showed that the skewed estimation effectively positioned fitted values at the lower boundary of the series, providing a more accurate baseline than traditional normal estimation. This refined method enhances the accuracy and reliability of complex regression models, with the potential for further tuning to improve forecasting in various business contexts.

Acknowledgements

This research was supported by the Internal Grant Agency of the Prague University of Economics and Business under grant F4/22/2023.

References

Fumi A, Pepe A, Scarabotti L, et al (2013) Fourier Analysis for Demand Forecasting in a Fashion Company. International Journal of Engineering Business Management 5:30. https://doi.org/10.5772/56839, http://journals.sagepub.com/doi/10.5772/56839

Singh P, Joshi SD, Patney RK, et al (2017) The Fourier decomposition method for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 473(2199):20160871. https://doi.org/10.1098/rspa.2016.0871, https://royalsocietypublishing.org/doi/10.1098/rspa.2016.0871

Veverka L (2022) The Optimal Settings of a Genetic Algorithm for Variable Selection in a Non-linear Time Series Model. In: Vojackova H (ed) 40TH INTERNATIONAL CONFERENCE MATHEMATICAL METHODS IN ECONOMICS 2022. Coll Polytechnics Jihlava, Jihlava, pp 392–397, https://www.webofscience.com/wos/woscc/full-record/WOS:000936355000062, num Pages: 6 Web of Science ID: WOS:000936355000062

Influence of Weights in Central Weight Vector on Additional Information in SMAA Method

Jan Volny[1]

Abstract.
Decision-making processes often rely on stochastic models due to the inherent uncertainty in evaluating alternatives. This study investigates how the weights of criteria within the central weight vector, derived via the Stochastic Multi-criteria Acceptability Analysis (SMAA) method, affect the valuation of additional information. Applying SMAA across various decision-making scenarios, we examined the correlation between criteria weights and the variability of acceptability and confidence indices after integrating supplementary data. Our analysis reveals a direct relationship: higher weights in the central weight vector significantly increase the impact of additional information on alternative evaluations, as evidenced by increased variance in both indices. These findings offer a strategic framework for decision-makers to efficiently allocate resources towards obtaining additional information for the most influential criteria and alternatives, thereby optimizing the decision-making process.

Keywords: Stochastic Multi-criteria Acceptability Analysis (SMAA), Decision-Making Process, Central Weight Vector, Additional Information, Decision Optimization.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

In decision-making processes, stochastic models are crucial for handling inherent uncertainties when evaluating various alternatives. These models enhance the understanding and quantification of uncertainties in complex decision-making environments. This study explores how the weights of criteria in the central weight vector, defined by the Stochastic Multi-criteria Acceptability Analysis (SMAA) method, affect the valuation of additional information. By applying SMAA in different decision-making scenarios, we examined the correlation between criteria weights and the variability of acceptability and confidence indices after incorporating supplementary data.

Our analysis reveals a direct relationship: higher weights in the central weight vector significantly influence the impact of additional information on evaluating alternatives, as evidenced by increased variance in both indices. These findings provide a strategic framework for decision-makers to allocate resources efficiently towards obtaining additional information for the most influential criteria and alternatives, optimizing the decision-making process. Adopting this approach helps decision-makers make better, more informed decisions that align closely with the complex and dynamic conditions they often face.

2 Theoretical background

According to Taherdoost and Madanchian (2023), multi-criteria decision making (MCDM) is a discipline within operations research aimed at supporting decision-making among alternatives defined by various criteria. The standard MCDM problem involves a set of alternatives **V** with m options, a set of criteria **K** with n criteria, and an evaluation matrix **X**, where element $x_{i,j}$ represents the evaluation of the i-th alternative according to the j-th criterion. Typically, the problem also includes a weight vector w that assigns relative importance to each criterion.

MCDM can be categorized into two main classes: deterministic and stochastic. In deterministic problems, the values of the weights and the evaluations of the alternatives according to the criteria are fixed and known. Conversely, stochastic approaches allow for the existence of randomness in the values of the weights or the evaluations of the alternatives, reflecting uncertainty in real-world decision-making scenarios. In this context, stochastic

[1] VSB - Technical university of Ostrava, Faculty of Economics, 17. listopadu 15, 70800, Ostrava, Czech Republic, jan.volny@vsb.cz.

variables can take discrete values, where the frequencies of the occurrence of individual values are known, or they can be continuous, meaning they are characterized by a probability distribution, such as the normal distribution.

2.1 Stochastic Multi-criteria Acceptability Analysis

The SMAA (Stochastic Multiobjective Acceptability Analysis) method operates on the principle of determining the percentage of weights for which a given variant is the best. This metric is referred to as the acceptability index, and it also identifies the weight vector that is the centroid of the hypersurface of all weights for which the given variant is the best. In cases where the evaluations of the variants are stochastic, a confidence factor is also obtained, indicating the likelihood that the weight vector, which is the centroid of the hypersurface of weights where the variant was the best, actually emerges as the best for the given variant.

According to Lahdelma (1997), the acceptability index for the *i*-th variant in a deterministic setting is calculated as the ratio of the volume of all weight vectors for which the *i*-th variant is the best to the volume of all weight vectors. Here, **W** represents the set of all weight non-random vectors that meet the user's or problem's criteria, and $\mathbf{W_i}$ is a subset of **W** for which the *i*-th variant is the best. The function *vol* represents the "volume" and measures how much of the vector space corresponds to the given set.

$$a_i = \frac{vol(\boldsymbol{W_i})}{vol(\boldsymbol{W})} \tag{1}$$

In the case of stochastic evaluations, the expected value of the volume of the weight vector is used in the ratio.

$$a_i = \frac{E\left(vol\big(\boldsymbol{W_i}(\gamma)\big)\right)}{vol(\boldsymbol{W})} \tag{2}$$

The central weight vector for alternative *i* is defined as the expected center of gravity and can be calculated as follows for deterministic (3) and stochastic (4) cases:

$$w_i^c = \frac{\int_{W_i} w\, dw}{\int_{W_i} dw} \tag{3}$$

$$w_i^c = \int_{\gamma} f(\gamma)\left(\frac{\int_{W_i(\gamma)} w\, dw}{\int_{W_i} dw}\right) d\gamma, \tag{4}$$

where $f(\gamma)$ is joint probability distribution for the criterion values. The confidence factor is obtained as the area of the probability distribution function for which the utility of variant *i* is greater than the utility of other variants.

$$p_i^c = \int_{\gamma: u_i(\gamma, w_i^c) > u_k(\gamma, w_i^c)} f(\gamma) d\gamma \tag{5}$$

where w_i^c is the central weight vector in which variant *i* is optimal.

3 Data and analysis

In this paper, we utilized data from two distinct research studies addressing critical decision-making problems. In both studies were used continuous uniform distribution for stochastic values. The first study, originally conducted by Hokkanen (1997), focuses on municipal planning, optimizing the development of urban areas. The second, explored by Hokkanen (2000), pertains to decisions about soil cleaning, involving environmental remediation strategies. We analyzed each problem by calculating central weight vectors, acceptability, and confidence indices using the Monte Carlo method, noted for its robustness in handling stochastic variability and uncertainty. We selected suitable alternatives for our research, simulated obtaining additional information by reducing the ranges of stochastic values, and then calculated the variance of acceptability and confidence indices.

3.1 Municipal Planning Analysis

First problem is decision about municipal general plan applications with six alternatives and four criteria. This problem was introduced by Hokkanen (1997) and later was used by Lahdelma (1998), whose values we will use. Alternatives are areas that could be developed. The alternatives are areas that could be developed, each evaluated against the following criteria:

- C1 – Environment – Values range from 1 to 5, with higher values indicating that the plan will have a lower impact on the environment.
- C2 – Costs – Costs of development in area.
- C3 – Land ownership – Percentage of land that city owns in area.
- C4 – Community structures – Values are 1-5 and higher values are preferred.

Alternative	C1	C2	C3	C4
I	5	-1.5	15	4
II	2.5	-3.8	25	2.5
III	3	-2.8	10	2.8
IV	3	-3.2	16	3.2
V	2	-6.7	0	1
VI	4	-3.4	30	3.5
δ	±0.5	±10 %	±10	±0.5

Table 1 Values of criterion in Municipal decision-making problem

By running Monte Carlo simulations, we derived the central weight vectors and computed the acceptability and confidence indices for each alternative, as presented in Table 2. For our research, we will focus on Alternative VI, as other alternatives either have low acceptability indices or high confidence indices, which would hinder our ability to observe changes effectively.

Alternative	a	p	w_1	w_2	w_3	w_4
I	74	99	28	29	17	26
II	1	13	8	15	66	11
III	0	0	3	47	41	7
IV	0	1	8	18	52	21
V	0					
VI	24	72	17	15	47	21

Table 2 Values of indices in Municipal decision-making problem

Following the initial assessment of the municipal planning problem, we proceed to evaluate the indices under varying ranges of criterion values to understand the impact of additional information. For each criterion we will be observing changes when we will reduce range to 40 % of its original span at eight different intervals, from the lowest values (e.g., for C1: <3.5, 3.9>) to the highest values (<4.1, 4.5>). We can see variance for each criterion are shown in table 3.

Criterion	a - variance	p - variance	Weight
C1	4.49	2.79	17
C2	0.58	0.50	15
C3	34.10	221.41	47
C4	7.90	5.57	21

Table 3 Variances in Municipal decision-making problem

3.2 Soil Cleaning Analysis

The second issue pertains to the selection of a company for soil cleaning in Finland. This decision, made by Hokkanen (2000), involved choosing among nine alternatives (companies) based on five criteria, which are detailed in Table 4. The criteria were as follows:

- C1 – Costs – Cost of soil cleaning.

- C2 – Environment – scale from 1 to 3, where higher values are better.
- C3 - Innovation – scale from 1 to 4, where higher values are better.
- C4 – Credibility - scale from 1 to 3, where higher values are better.
- C5 – Project management - scale from 1 to 3, where higher values are better.

Alternative	C1	C2	C3	C4	C5
I	-19±5.9	2	3	1	1
II	-2.3	1	1.8	2	2
III	-6	2	1.9	2	1.1
IV	-3.6	2.3	2	1.5	1.2
V	-8.4	2.2	2.7	2	2
VI	-13.6±2.85	2.8	2.8	2	2
VII	-5.2	1	1	2.8	3
VIII	-5.6	2	2	2.7	2.9
IX	-6	2	1	2.5	3
δ	±10 %	±0.5	±0.3	±0.3	±0.3

Table 4 Values of criterion in soil cleaning problem

By employing Monte Carlo simulations, we derived the central weight vectors and computed the acceptability and confidence indices for each alternative, as displayed in Table 5. For our research, we will concentrate on Alternatives V, VI, and IX, as the other alternatives either have low acceptability indices or high confidence indices, impeding our ability to effectively observe changes.

Alternative	a	p	w_1	w_2	w_3	w_4	w_5
I	0	1	6	11	74	5	4
II	3	84	66	4	14	9	6
III	0	4	33	40	10	15	0
IV	4	62	45	31	14	6	2
V	8	35	19	19	40	12	10
VI	22	81	9	35	31	13	11
VII	7	38	23	5	6	37	30
VIII	44	95	20	15	16	25	23
IX	9	30	19	23	5	19	34

Table 5 Values of indices in soil cleaning problem

Following the initial assessment of the municipal planning problem, we proceed to evaluate the indices under varying ranges of criterion values to understand the impact of additional information. For each criterion, we will observe changes when we reduce the range to 40% of its original span at eight different intervals. The variance for each criterion is shown in Table 6.

	V			VI			IX		
Criterion	**a - variance**	**p - variance**	**Weight**	**a - variance**	**p - variance**	**Weight**	**a - variance**	**p - variance**	**Weight**
C1	0.38	12.98	19	2.89	0.41	9	0.32	1.71	19
C2	9.97	4.41	19	20.55	120.13	35	12.7	164	23
C3	13.36	289.98	40	13.87	28.13	31	0.22	0.27	5

C4	2.13	10.29	12	4.11	4.86	13	4.75	6.57	19
C5	1.03	4.57	10	2.19	2.27	11	7.09	88.55	34

Table 6 Variances in soil cleaning problem

4 Statistical analysis of Criteria Weights and Decision Index Variances

In the previous chapter, we calculated the variances of indices for the selected alternatives in our case studies. In this chapter, we will demonstrate that higher weights in the central weight vector correlate with higher variances in indices. We will first examine the correlation between weights and the calculated variance of indices. Then, we will categorize our results into three groups and employ an ANOVA test to determine if there is a significant difference between the groups based on weight.

4.1 Correlation Analysis between Criteria Weights and Variances

Examining the acceptability index, we observe that, generally, the index increases with weight. This trend is illustrated in Figure 1.

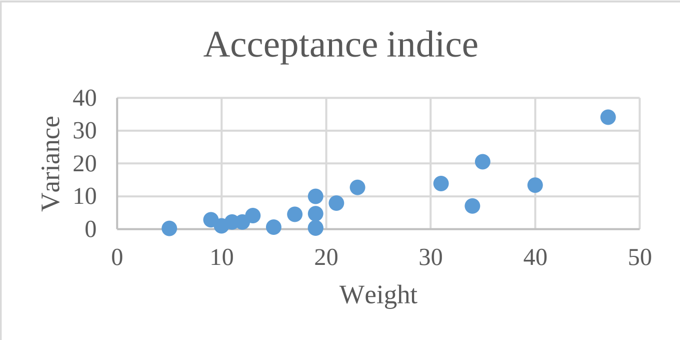


Figure 1 Acceptance indices Variance based on weights

Additionally, we can calculate the correlation coefficient, which is 0.85, indicating a strong correlation.

Similarly, for the confidence index, we observe a trend of increasing index with increasing weight, as shown in Figure 2.

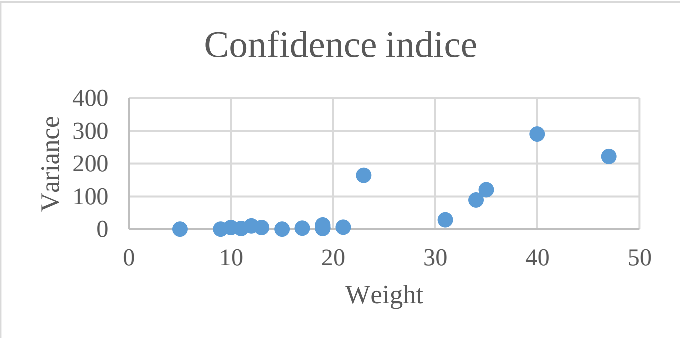


Figure 2 Confidence indices Variance based on weights

The correlation for the confidence index is also strong, with a coefficient of 0.81.

4.2 Analysis of Variance (ANOVA)

According to Ostertagova (2013) is ANOVA very useful technique in revealing important information in interpreting experimental outcomes and in determining the influence of some factors on other processing parameters.

In our case we will study, whether weights have influence on acceptability and confidence indices. For this reason we divide our results into three groups based on weight and perform an ANOVA test.

First, we will calculate expected weight as can be seen in the following equation (6).

$$w_e = \frac{1}{number\ of\ criteria} \quad (6)$$

Second, we categorize the weights into three groups: low, medium, and high. These groups are defined as follows:

- **Low Weight Group:** Weights lower than 0.75 x w_e.
- **Medium Weight Group:** Weights between 0.75 x w_e and 1.25 x w_e.
- **High Weight Group:** Weights higher than 1.25 x w_e.

The ANOVA test results are summarized in Table 7.

Group	Mean Variance (Acceptability)	Mean Variance (Confidence)
Low Weight	1.88	3.31
Medium Weight	7.72	31.37
High Weight	18.78	180.01
p-value	0.001889	0.007726

Table 7 ANOVA results summarized

As we can see, p-value for both acceptability and confidence indices is lower than 0.05, so we can assume, that there is difference between groups and by looking at mean variance values, we can assume, that groups with higher weights, will have higher values.

5 Conclusion

The study demonstrates a significant correlation between the weights of criteria in the central weight vector and the impact of additional information on decision-making outcomes, evidenced by changes in acceptability and confidence indices. By analyzing data from municipal planning and soil cleaning scenarios, it was observed that higher weights associated with specific criteria consistently led to greater variances in both indices upon the integration of supplementary data. This finding underscores the importance of strategic information gathering, particularly in areas deemed crucial by the weight vectors. Decision-makers are advised to prioritize resource allocation towards obtaining detailed information for the most influential criteria, thereby enhancing the precision and effectiveness of their decisions under uncertainty. The analysis also supports the use of ANOVA to validate the differences in variances, providing a robust statistical foundation for interpreting the influence of criteria weighting on decision-making processes.

Acknowledgements

This article has been produced with the financial support of SGS project No. SPS2024/078. This support is gratefully acknowledged.

References

[1] Hokkanen, J., Lahdelma, R., Miettinen, K., & Salminen, P. (1997). Determining the implementation order of a general plan by using a multicriteria method. Journal of MultiCriteria Decision Analysis, 7(5), 273-284. https://doi.org/10.1002/(SICI)1099-1360(199809)7:5<273::AID-MCDA198>3.0.CO;2-1

[2] Hokkanen, J., Lahdelma, R., & Salminen, P. (2000). Multicriteria decision support in technology competition for cleaning polluted soil in Helsinki. Journal of Environmental Management, 60(4), 339-348. https://doi.org/10.1006/jema.2000.0389.

[3] Lahdelma, R., Hokkanen, J., Salminen, P. (1998). SMAA - Stochastic multiobjective acceptability analysis. European Journal of Operational Research, 106(1), 137-143. https://doi.org/10.1016/S0377-2217(97)00163-X

[4] Ostertagova, E., Ostertag, O. (2013). Methodology and Application of Oneway ANOVA. American Journal of Mechanical Engineering, 1(7), 256-261.

[5] Taherdoost, H., Madanchian, M. (2023). Multi-Criteria Decision Making (MCDM) Methods and Concepts. Encyclopedia, 3(1) , 77-87. https://doi.org/10.3390/encyclopedia3010006

Exploring Market Attention's Impact on Portfolio Optimization

Jialei Xiong[1]

Abstract. This study examines whether market attention has a positive impact on portfolio optimization. Using a rolling-window approach, we extract Google Trends time series data for specific companies as a proxy for market attention. This market attention indicator is then incorporated into a suitable Google Trends model. By applying this model, the market attention of a stock is effectively translated into appropriate weights for portfolio optimization. The findings of this study suggest that incorporating market interest into portfolio optimization can effectively enhance portfolio performance and diversify risk. This study contributes to further validate the constructive impact of market attention on portfolio optimization.

Keywords: market attention, Google Trends, time series analysis, portfolio optimization, rolling window

JEL Classification: G11, G41
AMS Classification: 90A09, 91B28, 91G10

1 Introduction

The application of investor attention has attracted a great deal of attention in recent years of research. Information occupies the receiver's attention (Simon & Greenberger, 1971); hence, a glut of information leads to a loss of attention, and people must concentrate on the information they require. According to Kahnemann (1973), attention is a limited cognitive resource, and investors will pay attention to assets with more information. Whereas online search queries can be an excellent mapping of an investor's current desire for information, they have also been used to measure investor attention (de Castro et al., 2023).

Online search queries enable market participants to have online access and collect and analyze financial information in real-time. The efficient way to obtain stock market information through the (mobile) internet can reduce search costs, trading costs, and information dissemination costs (Johri et al., 2023). With the rapid development of the internet, online search queries have become a significant means for individual/retail investors to get stock market information in a timely and easy way. As the most important tool for online search queries, the search intensity in Google Trends is defined as the online search attention, and the attention of individual/retail investors can be well captured by the Google search intensity (Ginsberg et al., 2009; BenRephael et al., 2017; de Castro et al., 2023).

In recent years, scientists have explored the relationship between investor attention and the price of assets, mostly using the search intensity of Google Trends (Da et al., 2011), demonstrating that investors' active search for stocks can have an impact on both stock prices and returns. In addition to investigating the relationship between investor attention and asset price returns, Vlastakis & Markellos (2012) and Vozlyublennaia (2014) analyze the in-sample properties of the relationship between investor attention and volatility using Google search intensity. Hamid and Heiden (2015) discovered that the accuracy of volatility forecasts can be positively correlated with investor attention during highly volatile market phases. Although considerable research has focused on investigating the relationship between investor attention and asset prices and volatilities, less attention has been given to apply investor attention for portfolio optimization. Little study on how investor attention might be used to optimize portfolios. Our research will help fill this gap.

The goal of the study is to verify the applicability of the market attention, proxied by the Google Trends data, in the portfolio optimization. In this study, we use data from Google Trends search engine data to analyze the relationship between investor attention and portfolio optimization. We apply the diversification strategy proposed by Kristoufek (2013) to explore whether incorporating Google Trends search intensity data in portfolio optimization diversifies portfolio risk and results in better performance. Furthermore, in order to examine how the Google Trends intensity of each chosen stock fluctuates over time and contributes to the evaluation of its stability, we

[1] VŠB – Technical University of Ostrava, Department of Finance, 17. listopadu 2172/15, 708 00, Ostrava-Poruba, jialei.xiong@vsb.cz.

apply a rolling window approach to rebalance the Google Trends search intensity obtained for each stock applied for one window. We apply the chosen stock data from the Hong Kong Stock Exchange. To the best of our knowledge, there is no paper that examines the relationship between search intensity and stock trading behavior in the Hong Kong market, this study will fill this gap.

The study is structured as follows. The first chapter is an introduction. The second chapter briefly describes the methodology used in this project. The third chapter presents the applicable data. The fourth chapter presents the obtained results. The final chapter is the conclusion.

2 Methodology

This study investigates the effectiveness of risk diversification using Kristoufek's (2013) Google Trends model. To go deeper in identifying changes or outliers in the data over time in a lengthy time series, the rolling-window method is applied to obtain the Google Trends of selected stocks for each time period. Subsequently, these data are used in the Google Trends model to derive different weights for each selected stock at different levels of diversification.

Google Trends consolidates the intensity of user searches for specific keywords in Google searches, allowing users to analyze trends across different time horizons and geographic regions. To investigate the incorporation of market attention into portfolio optimization, this study refers to the portfolio diversification strategy proposed by Kristoufek (2013), which measures the intensity of relevant searches using stock-related terms as keywords, known as the Google Trends strategy. The strategy operates on the premise that an increase in search intensity for terms is associated with an increase in stock risk. To identify the popularity of the stock, lower weights are assigned to stocks with higher search intensity, and higher weights are assigned to stocks with lower search intensity during the portfolio optimization process. Therefore, in week *t*, the weight of the stock *i* in the portfolio $w_{i,t}$ can be calculated as follows,

$$w_{i,t} = \frac{V_{i,t}^{-\beta}}{\sum_{j=1}^{n} V_{j,t}^{-\beta}}, \tag{1}$$

where $V_{i,t}$ represents the Google Trends search intensity for stock *i* in week *t*. A higher search frequency for stock *i* signifies a higher search intensity. The β is a parameter of the power law model, quantifying the discriminatory influence of frequent searches for a given stock. The β value greater than 0 suggests that stocks with higher search intensity receive lower weights. On the contrary, the β value of lower than 0 indicates that stocks with higher search intensity receive higher weights. The β value equal to 0 implies that each stock in the portfolio is assigned an equal weight.

3 Data

This study creates an 11-year data set from January 2013 to December 2023, consistent with the time frame of the selected stocks. The data set contains monthly data on stock prices and records stock price information for approximately 132 months. It is assumed that there are no transaction costs associated with portfolio creation each month. All stock prices are in Hong Kong dollars (HKD). According to Hong Kong 10-Year bond historical data - Investing.com. (2024), Hong Kong's 10-year bond yield was 3.272%. The initial wealth in January 2013 is set to HKD 1. The Hang Seng Index is applied as a benchmark to compare our results with.

To use Google Trends effectively, it is important to consider certain key characteristics. Firstly, Google Trends data do not provide the exact search volumes for selected keywords, but rather present trends in search intensity over time in a relative way. In other words, Google Trends data are normalized, and it is proportionally transformed to the total search volume relative to a given time period and geographic location. Google Trends presents a search intensity between 0 and 100, where 100 represents the highest search intensity for the keyword searched in the selected time period and geographic location. It is worth noting that when comparing different Google Trends data obtained for different keywords, the respective search intensities are normalized according to the maximum search intensity of all data. It should be considered that our study relies on data from the existing Google Trends platform, which allows the comparison of up to five keywords on a single webpage. Thus, we divided the selected 30 stocks into six groups for comparison purposes. We then normalized and rescaled the search intensities of these six groups to obtain the Google Trends search intensity dataset of all 30 stocks required for our analysis.

Google Trends allows users to define the time frame according to their preferences. Google Trends provides data on a weekly basis, from each Sunday to the following Saturday. Users can access weekly data for up to five years or monthly data if the period is longer. In this study, to obtain 11 years of weekly Google Trends data, we strategically chose two intermediate weeks in the 11-year timeframe as the key points to connect three different time periods. In this way, a uniform and standardized Google Trends search intensity dataset is ensured. This approach ensures that a uniform and standardized data set of Google Trends search intensity is obtained for the specified time periods. We take each month as a window, average the acquired weekly frequency Google Trends data on a monthly basis, and then apply it to *equation* (1) to obtain the monthly weights for each stock. It is worth noting that since some weekly frequency Google Trends search intensity data, we obtained has the value of 0, which would make the power law function of *equation* (1) meaningless, we replace these zeros with a small positive value of 0.1, which is very common in time series analysis (Hastie et al., 2009).

Considering that Google Trends as market attention can have an impact on the future stock market (Aboody et al., 2010; Ballinari et al.,2022), we assume that the market attention captured by Google Trends weights from the previous month impacts stock return in the current month. Additionally, we conduct a monthly rolling-window analysis to allow the portfolio rebalancing. When using Google Trends, our focus is to maximize the searchability of all selected stock keywords. We chose to enter the abbreviation "HKG" as well as the respective stock code for each stock. For example, when searching for Tencent Holdings Limited, we enter the search term "HKG:0700" in Google Trends. To improve the accuracy of Google Trends in describing each stock, we chose to search using the "Topics" category. However, due to insufficient search data for terms related to CHINA Longfor CO Ltd, ANTA Sports Products Ltd and Shenzhou International Group Holdings Limited, we had to remove their data from our analysis.

4 Results

In the Google Trends strategy, we specify the power law parameter β. This power law parameter β varies in increments of 0.1 over a range of -1 to 1, and incorporating it into *equation* (1) allows us to obtain the weights of the stocks under different power law parameters. To facilitate the observation of changes in stock weights under different settings of the power law parameters β, we selected five of the 21 power law parameters β to be analyzed.

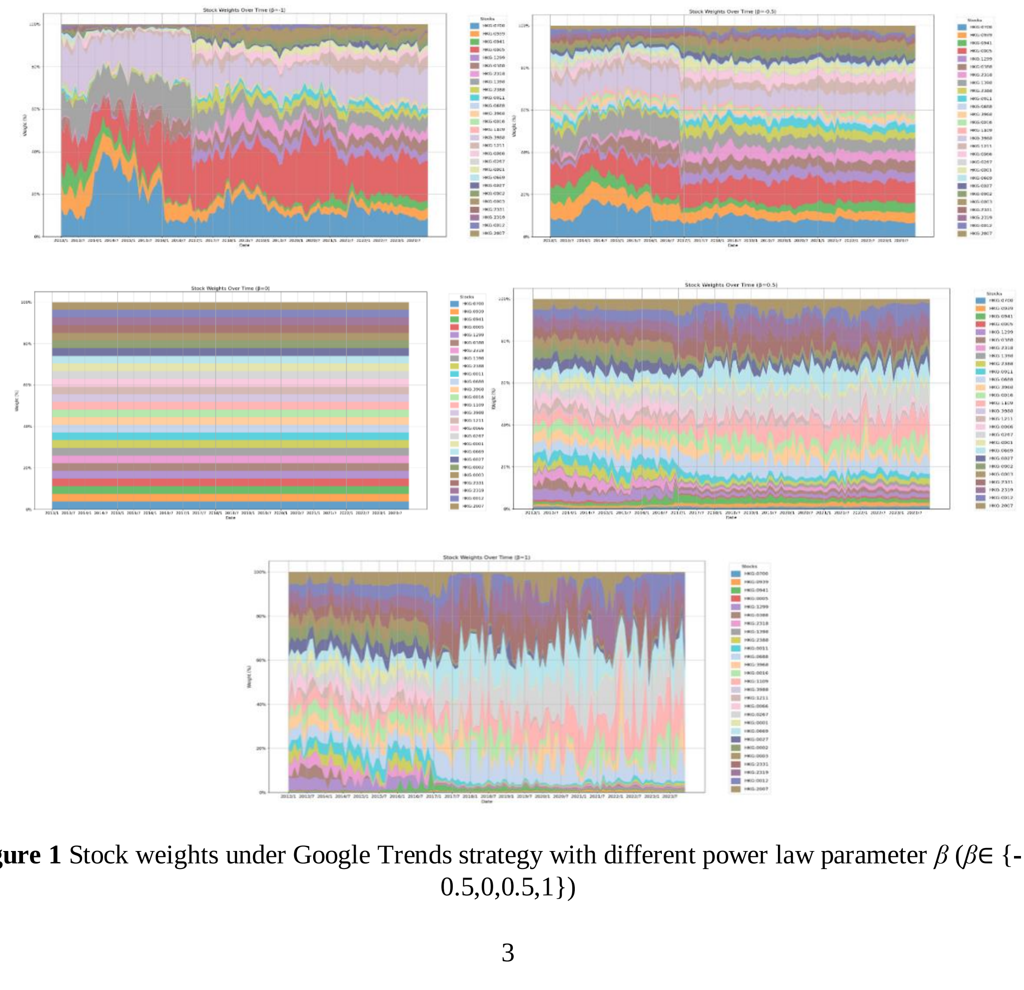


Figure 1 Stock weights under Google Trends strategy with different power law parameter β ($\beta \in \{-1,-0.5,0,0.5,1\}$)

Positive values of the power law parameter β are consistent with the expected application of the risk diversification strategy. Therefore, when β exceeds 0, the weights allocated to the stocks decrease accordingly as the search intensity increases. On the contrary, when β is less than 0, an increase in search intensity leads to an increase in the weights allocated to the corresponding stocks. A comparative analysis of the four figures representing positive and negative beta values shows that the stocks with a large share of stock weights at positive and negative values are different from each other, which is consistent with the basic principle of the applied Google Trends strategy.

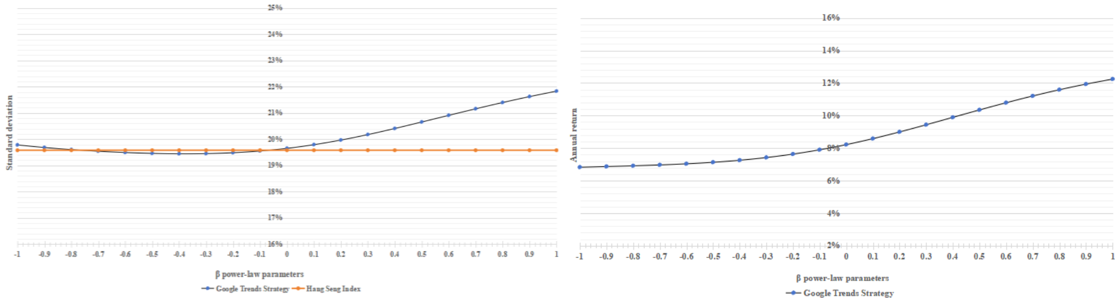


Figure 2 Comparison of standard deviation and annual returns between Google Trends Strategy and Hang Seng Index

This project uses the Hang Seng Index as the benchmark. **Figure 2** shows that for negative values of β, the standard deviation of the Google Trends strategy remains relatively consistent compared to the benchmark, and even on closer examination, the standard deviation of most negative values of β is lower than the standard deviation of the benchmark. However, as β becomes positive, the standard deviation of the Google Trends strategy shows a clear upward trend. It is observed that the corresponding higher standard deviation has higher annual returns. We also considered the performance of the Sharpe ratio and the maximum drawdown, taking into account the fact that there is a case where the higher the standard deviation, the higher the return.

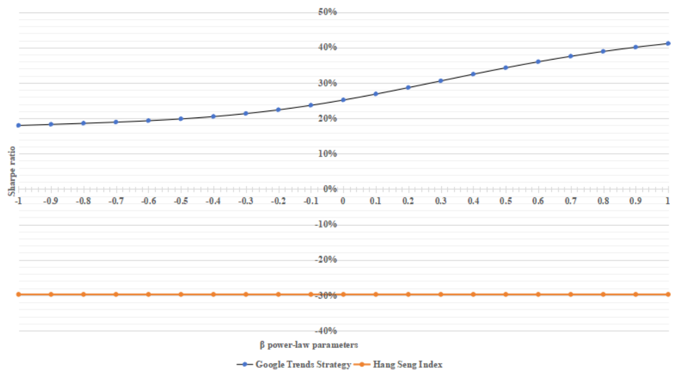


Figure 3 Comparison of Sharpe ratio between Google Trends Strategy and Hang Seng Index

Figure 3 shows the higher the value of β, the higher the Sharpe ratio. This is especially evident when β is positive, in which case the Sharpe ratio shows a clear upward trajectory. When comparing the Google Trends strategy with the benchmark, the Sharpe ratios of the Google Trends strategy for the various power law parameters β significantly outperform the negative Sharpe ratios of the benchmark. Since the correlation between larger Sharpe ratios and higher risk-return ratios is obvious, the Google Trends strategy outperforms based on the Sharpe ratio as a performance indicator.

As can be seen in **Figure 4**, the maximum drawdown (MDD) under the Google Trends strategy is much lower than the benchmark, indicating that the Google Trends strategy performs extremely well when compared to the MDD as a performance.

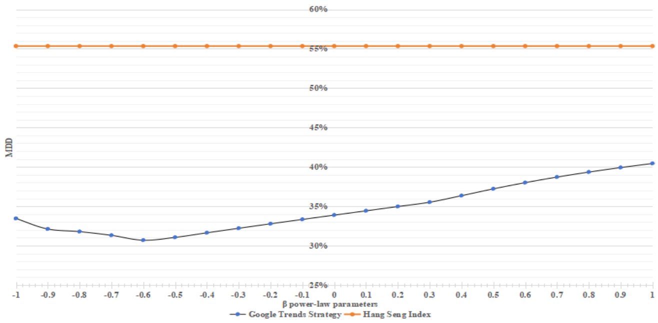


Figure 4 Comparison of the MDD between Google Trends Strategy and Hang Seng Index

In summary, comparing the performance of the benchmark with the Google Trends strategy across all performance categories indicates that the latter demonstrates significantly superior performance. Despite taking into account the smaller subset of constituents (27 for the Google Trends Strategy and 76 for the Hang Seng Index), the Google Trends strategy demonstrates superior performance, particularly its superior Sharpe ratio and MDD. Despite slightly underperforming the benchmark in terms of standard deviation, the strong performance of the Google Trends strategy highlights its effectiveness as a diversification strategy to focus market attention. As a result, we believe that market attention can contribute significantly to portfolio optimization efforts.

5 Conclusion

The purpose of this study is to assess whether market attention has a positive impact on portfolio optimization and improves overall performance, with a particular focus on risk diversification, as well as improving the Sharpe ratio and MDD.

This study allows us to get weekly market attention for selected stocks by using their respective codes as keywords to search Google Trends intensity. Using a rolling window approach, we dynamically adjust these attention levels to obtain appropriate portfolio weights for rebalancing and subsequently assess performance within a portfolio optimization framework.

This study compares the Google Trends strategy with the Hang Seng Index as a benchmark. Our analysis shows that the Google Trends strategy is effective in reducing risk to a certain degree while generating a superior Sharpe ratio and MDD. However, the Google Trends Strategy slightly underperforms the benchmark in terms of risk diversification, which is mainly attributed to the fact that its superior returns lead to relatively high risk. However, the risk level of the Google Trends strategy does not differ significantly from the benchmark.

In conclusion, our findings suggest that market attention significantly improves portfolio optimization performance. Future research could examine the comparative analysis between the Google Trends strategy and the traditional portfolio optimization method in more detail, leading to a more detailed assessment of the effectiveness of market attention strategies. Furthermore, the results observed in this study are slightly different from those of Kristoufek (2013) on the relationship between β and diversification, who concluded that a positive β leads to better diversification. However, we find that for better diversification, β should be negative. This requires further investigation to elucidate the underlying reasons.

Acknowledgements

The research was supported by the Student Grant Competition of the Faculty of Economics, VSB - Technical University of Ostrava, project number SP2024/047.

References

Aboody, D., Lehavy, R., & Trueman, B (2010). Limited attention and the earnings announcement returns of past stock market winners. Review of Accounting Studies, 15, 317-344.

https://link.springer.com/article/10.1007/s11142-009-9104-9

Ballinari, D., Audrino, F., & Sigrist, F (2022). When does attention matter? The effect of investor attention on stock market volatility around news releases. International Review of Financial Analysis, 82, 102185. https://www.sciencedirect.com/science/article/pii/S1057521922001466

Ben-Rephael, A., Da, Z., Israelsen, R.D (2017). It depends on where you search: institutional investor attention and underreaction to news. Rev. Financ. Stud. 30 (9), 3009–3047. https://academic.oup.com/rfs/article-abstract/30/9/3009/3830159?login=true

Da, Z., Engelberg, J., & Gao, P (2011). In search of attention. The journal of finance, 66(5), 1461-1499. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1540-6261.2011.01679.x

de Castro, J., & Piccoli, P (2023). Do online searches actually measure future retail investor trades?. International Review of Financial Analysis, 86, 102552. https://www.sciencedirect.com/science/article/pii/S1057521923000686

Fu, C., G. Jacoby, & Y. Wang (2015). Investor sentiment and portfolio selection. Finance Research Letters, 15, 266–273.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L (2009). Detecting influenza epidemics using search engine query data. Nature, 457(7232), 1012–1014. https://www.nature.com/articles/nature07634

Hamid, A., & Heiden, M (2015) Forecasting volatility with empirical similarity and Google Trends. Journal of economic behavior & organization, 117, 62-81. https://www.sciencedirect.com/science/article/pii/S0167268115001687

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.

Johri, A., Wasiq, M., Kaur, H., & Asif, M (2023). Assessment of users' adoption behaviour for stock market investment through online applications. Heliyon, 9(9). https://www.cell.com/heliyon/pdf/S2405-8440(23)06732-4.pdf

Kahneman, D (1973). Attention and effort,. Vol. 1063: Prentice-Hall Englewood Cliffs. NJ.

Kristoufek, L (2013). Can Google Trends search queries contribute to risk diversification? Sci Rep 3, 2713. https://www.nature.com/articles/srep02713

Nguyen, C., Hoang, L., Shim, J., & Truong, P (2020). Internet search intensity, liquidity and returns in emerging markets. Research in International Business and Finance, 52, 101166. https://www.sciencedirect.com/science/article/pii/S0275531919306130

Sharpe, W. F (1966). Mutual fund performance. The Journal of business, 39(1), 119-138. https://www.jstor.org/stable/2351741

Simon, H., & Greenberger, M (1971). Computers, communications and the public interest. Computers, communications, and the public interest. Johns Hopkins Press, Baltimore, 40-41.

Vlastakis, Nikolaos, & Raphael N. Markellos (2012). Information demand and stock market volatility, Journal of Banking & Finance 36, 1808{1821. https://www.sciencedirect.com/science/article/pii/S0378426612000507

Vozlyublennaia, N (2014). Investor attention, index performance, and return predictability. J. Bank. Finance 41C, 17–35. https://www.sciencedirect.com/science/article/pii/S0378426613004822

Website sources

Hong Kong 10-Year bond historical data - Investing.com. (2024). Investing.com. https://www.investing.com/rates-bonds/hong-kong-10-year-bond-yield-historical-data

The Problem of Optimal Delivery of Frozen and Chilled Goods with Given Priority from Multiple Warehouses

Jaromír Zahrádka[1]

Abstract. This article comes up with a specific example of the vehicle routing problem. The selling company has to deliver the ordered frozen or chilled goods from s warehouses to n customers as efficiently as possible. Each customer has ordered goods stored in a certain number of containers which need to be transported. All customer points of delivery and warehouse points are given by GPS coordinates. The objective of the solution is to select the number of vehicles and their routes between suitable warehouse and customers in such a way that the total travel distance or travel time is as short as possible. The order of customers on each route respects the priority of delivery of frozen goods over chilled ones. This means that chilled goods are unloaded from the truck only after all frozen goods have been unloaded at previous customer delivery points. Each of these delivery points is visited only once by one of the vehicles. In each warehouse, the same number of trucks ends the journey as they left. All trucks have the same pre-limited capacity of containers. In this article, the algorithm of the vehicle routing problem with multiple warehouses and priority of delivery of frozen goods was created and implemented in Matlab code.

Keywords: Chilled goods, frozen goods, Matlab, mixed integer linear programming, optimization, point of delivery, priority, vehicle routing problem

JEL Classification: C64
AMS Classification: 68W04, 90C11, 05C20

1 The Vehicle Routing Problem with Multiple Warehouses.

The vehicle routing problem with multiple warehouses (VRPMW) is inspired by a practical situation that a trading company supplying, for example, food products has several warehouses in its territory from which it distributes goods to customers. The entire range of delivered goods is available in all warehouses and there is always a sufficient number of vehicles available. The solution to the VRPMW consists in planning the routes of trucks delivering goods so that the total distance travelled by all vehicles is as small as possible.
Our presented multi-storage VRPMW solution with customer priority according to the temperature regime of the delivered goods is a continuation of the single-storage VRP solution [1], and the Matlab one published in [6]. Our solution uses the principles of integer programming, which are contained in [4, 5]. They respect the principles of operating intelligent transport systems [2].

1.1 Mathematical Formulation of VRPMW

We calculate that n customers are served from s warehouses. The customers sites (nodes) are marked by numbers $1, \dots, n$, and warehouses (depots) are marked by numbers $n+1, \dots, n+s$. The vehicle routing problem with multiple warehouses (VRPMW) can be presented as the subsequent graph problem. Let $G = (V, E)$ be a complete directed graph where $V = \{1, \dots, n, n+1, \dots, n+s\}$ is the nodes set and E is the set of all oriented arcs. Nodes $i = 1, 2, \dots, n$ correspond to the customers, each with a number q_i of demand containers, which form the row vector $\boldsymbol{q} = (q_1,\ q_2,\ \dots\ , q_n)$. Each oriented $arc(i,j)$ is associated with non-negative values d_{ij} of travel distance (in meters), and c_{ij} of travel time (in sec) from node i to node j. For easier reference we assume, $I = \{1,\ \dots\ ,\ n\}$, and $I_D = I \cup \{n+1, \dots, n+s\}$. The distance matrix $\boldsymbol{D} = (d_{ij})_{i,j \in I_D}$ and time distance matrix $\boldsymbol{C} = (c_{ij})_{i,j \in I_D}$ are non-negative and generally asymmetric.

The VRPMW consists of finding a collection of simple cycles and simple paths between warehouses, each corresponding to a vehicle route with minimal sum of the distances of the cycles and path arcs, such that:
a) each cycle starts and ends in one of the warehouses;
b) each path starts in one of the warehouses and ends in another one;
c) each vertex $j \in I$ is visited by exactly one cycle or path;

[1] University of Pardubice, Department Mathematics and Physics, Studentská 95, 53210 Pardubice, jaromir.zahradka@upce.cz.

d) the same number of cycles and paths start and end in each warehouse;
e) the sum of delivered containers during a cycle or path not exceed the vehicle capacity Q.
At the beginning and end of each cycle or path is one of s stores. All customers who have ordered frozen goods must be visited. After the delivery of all frozen goods, the customers to whom the chilled goods are delivered are visited.
We assume, that m_i is the service time associated with the unloading of goods and dealing with the customer for each customer $i \in I$. All service times are arranged in a row vector $\boldsymbol{m} = (m_1, m_2, \dots, m_n)$.

1.2 Mathematical Solution of VRPMW

The used optimization model is generally described by

$$\min_{V}(\boldsymbol{f} \cdot \boldsymbol{V}) \text{ subject to} \begin{cases} \boldsymbol{V}_{intcon} \text{ are integers} \\ \boldsymbol{A} \cdot \boldsymbol{V} \leq \boldsymbol{b} \\ \boldsymbol{A}_{eq} \cdot \boldsymbol{V} = \boldsymbol{b}_{eq} \\ \boldsymbol{l}_b \leq \boldsymbol{V} \leq \boldsymbol{u}_b \,. \end{cases} \tag{1}$$

The vector $\boldsymbol{V}$ is the column vector of all flow variables; $\boldsymbol{f} \cdot \boldsymbol{V}$ is the objective function, the coefficients of this objective function are components of the row vector $\boldsymbol{f}$; $\boldsymbol{V}_{intcon}$is the list of variable indices of the vector $\boldsymbol{V}$ that takes only the integer values; $\boldsymbol{A} \cdot \boldsymbol{V} \leq \boldsymbol{b}$ determines the system of inequality constrains; $\boldsymbol{A}_{eq} \cdot \boldsymbol{V} = \boldsymbol{b}_{eq}$ determines the system of linear equations; and inequalities $\boldsymbol{l}_b \leq \boldsymbol{V} \leq \boldsymbol{u}_b$ indicate the lower and upper bounds of flow variables.

The core of the solution of VRPMW is finding suitable configurations of paths and cycles in the graph that start and end in one of the warehouses. Each customer is on exactly one path or cycle. Another condition is that the same number of vehicles arrive at each warehouse, as left the warehouse. The optimal solution is the configuration of paths and cycles that gives the shortest total sum of the lengths of all paths and cycles. To realize the solution by the optimization method, integer flow variables x_{ij} are introduced on the set of all oriented arcs of the complete oriented graph, which can only take on binary values 0 or 1. The value $x_{ij} = 1$ means that the arc from i node to j is included in one path or cycle, and the value $x_{ij} = 0$ means that the corresponding arc is not included. For systemic reasons, variables $x_{i\,i}$ are used but are fixed $x_{i\,i} = 0$, for each $i \in I_D$. Variables x_{ij} are elements of a matrix $\boldsymbol{X} = (x_{ij})_{ij \in I_D}$, and the number of x_{ij} variables is $(n+s)^2$.

In our work we use other specific integer flow variables, y_i, for each $i \in I$, which indicate the number of containers that were unloaded to customers during the journey from the depot up to and including the node i. The variables y_i are n elements of the vector $\boldsymbol{y} = (y_1, \ y_2, \ \dots \ , y_n)$. The sum of all flow variables is $(n+s)^2 + n$.
Elements of matrix $\boldsymbol{X} = (x_{i\,j})_{i\,j \in I_D}$ and vector $\boldsymbol{y} = (y_1, \ y_2, \ \dots \ , y_n)$ are arranged in the column vector V so that,
$\boldsymbol{V}^T = (x_{1,1}, x_{1,2}, \ \dots \ , x_{1,n+s}, \ x_{2,1}, \ x_{2,2}, \ \dots \ x_{1,n+s}, \ \dots \ , \ x_{n+s,1}, \ x_{n+s,2}, \ \dots \ , x_{n+s,n+s}, \ y_1, \ y_2 \ , \dots \ , \ y_n)$.
The solution of VRPMW is the optimal solution of a mixed-integer linear programming problem:

$$\min_{(\boldsymbol{X},\boldsymbol{y})}\left\{\sum_{i,j=0}^{n+s} d_{ij} \cdot x_{ij} + \sum_{i=1}^{n} y_i\right\} \text{ subject to} \tag{2}$$

$$x_{ij}, \quad i,j \in I_D \text{ are binary, i.e.} x_{ij} \in \{0,1\} \tag{3}$$

$$y_i, \quad i \in I \text{ are integer} \tag{4}$$

$$Q\,x_{ij} + y_i - y_j \leq Q - q_j, \quad i,j \in I, \quad i \neq j \tag{5}$$

$$Q\,x_{kj} - y_j \leq Q - q_j, \quad k \in \mathrm{I}_D - I, j \in I \tag{6}$$

$$\sum_{j \in I_D} x_{ij} = 1\,, \qquad i \in I \tag{7}$$

$$\sum_{i \in I_D} x_{ij} = 1\,, \qquad j \in I \tag{8}$$

$$\sum_{j \in I} x_{kj} - \sum_{i \in I} x_{ik} = 0, \qquad k \in \mathrm{I}_D - I \tag{9}$$

$$x_{ii} = 0, \ i \in I_D \tag{10}$$

$$x_{ij} = 0,\ i,j \in I_D - I, \tag{11}$$

$$0 \le x_{ij} \le 1, \quad i,j \in I_D \tag{12}$$

$$q_j \le y_j \le Q, \quad j \in I \tag{13}$$

It means, that the linear optimization function is minimized such that

$$\boldsymbol{f} = \sum_{i,j=1}^{n+s} d_{ij} x_{ij} + \sum_{i=0}^{n} y_i. \tag{14}$$

The sum of distances (in meters) $\sum_{i,j=1}^{n+s} d_{ij} x_{ij}$ guarantees finding a collection of cycles and paths such that the sum of their lengths is minimal. The part $\sum_{i=1}^{n} y_i$ of the optimization function (13) is about four orders of magnitude smaller and does not affect the optimization according to the smallest length. This part guarantees that the solutions of flow values y_i are also as small as possible. Constraint (5) defines $n(n-1)$ conditions between flow variables y_i, y_j and x_{ij}. In the case $x_{ij} = 1$, the inequality (5) expresses the relationship $y_j \ge y_i + q_j$. Thanks to the inclusion of the variables $y_1,\ y_2,\ \dots\ ,\ y_n$ in the optimized function (13), it is ensured that the values of y_j will always be minimal, i.e. the equation $y_j = y_i + q_j$ will be applied instead of an inequality. In the case $x_{ij} = 0$, for $i \ne j$, the inequality (5) expresses the relationship $y_j \ge y_i + q_j - Q$.

Statements (7) and (8) declare $2n$ equations, which express that only one arc leads from each node $i \in I$ and only one arc leads to each node $j \in I$. Statement (9) declares that for each store nodes $n+1, \dots, n+s$, the number of arcs leaving is equal to the number of arcs entering. Statement (10) declares that there are no loops in the graph. Statement (11) declares that there are no arcs between store nodes.

The inequalities in (12) state that the lower and upper bounds of flow variables x_{ij} are 0 and 1. The inequalities in (13) declare that each flow variable y_j is greater than or equal to q_j, and each y_j can't be greater than the capacity Q of vehicle.

1.3 Mathematical Solution of VRPMW with Given Priority of Delivery

For the delivery of goods in two different temperature regimes (e.g. frozen goods -18°C and chilled goods +6°C), they can be transported with a separating partition: The partition enables the respective temperature regimes to be maintained in individual parts of the transport space. In such a case, the sequence of customers should be prepared so that frozen goods are unloaded first, followed by customers with chilled goods. It means that customers (nodes) with frozen goods have a higher priority delivery.

A row vector $p = (p_1,\ p_2,\ \dots\ , p_n)$. with component values of 1 or 2 is introduced to indicate the priority of customer nodes. The node priority condition does not allow a cycle or path with any arc leading from a node with priority 2 to a node with priority 1. This is achieved by adding a new constraint

$$x_{ij} = 0, \quad i,j \in I, \quad p_i > p_j \tag{15}$$

in the optimization model (2).

1.4 Implementation of the Solution in Matlab Code

The solution to the VRPMW problem is implemented in the Matlab programming language and is contained in the M-function *SOLVER_VRPMW_Z.m*, which is presented as an Appendix at the end of this article. In the Matlab solution, the same variables are used as in the mathematical description. However, variable identifiers are adapted to the syntax of the Matlab language.

The M-function *SOLVER_VRPMW_Za.m* is started using the special setup M-script, which is not published in this article. Using the setup M-script, all input data are introduced, all output data are taken, and the all figures shown in this article are plotted.

The output flow variables are the components of column vector V, which is obtained as the result of the command *V=intlinprog(f,intcon,A,b,Aeq,beq,lb,ub,[],options)* (App. row No. 23). A more detailed explanation of command *intlinprog* can be found in the User's Guide [3].

The matrix A of system linear inequalities and the column vector b of their right sides are created in Matlab code by statements on lines No. 4 to 10 in the Appendix. The matrix Aeq of system linear equalities and the column vector beq of their right sides are created by statements on lines No. 11 to 15 of the Appendix.

The lower and upper bounds of the flow variables with respect to the relations (10), (11), (12), (13) are created as components of vectors Lb and ub on lines No. 16, 17, 18, 19 in the Appendix. The fulfilment of condition (15) is ensured by the command in line No. 20.

The coefficients of the objective function are created by commands on line No. 22 of the Appendix. By last command on the line No. 22 all flow variables are taken as integers.

The for-cycle commands on lines No. 26 and 27 allow calculation of the total length of cycles and paths `TotLgth` and the total driving time `TotDur` of all used trucks. The total driving time `TotDur` includes service times of all customers. The M-function *VRPMW_SOLVER_Z.m* is executed by using a special startup M-script which enables the introduction of input data, processing of output data, and drawing output cycles and paths. The startup M-script is not presented in this article.

2 Demonstration Applied Task

Three warehouses in the region of Czech Republic and 27 customers were selected for the demonstration of the applied task. The GPS coordinates of customer locations and warehouses are listed in Table 1. In the table q_i means the number of ordered containers by customer i and the value p_i specifies frozen or chilled goods. The value $p_i = 1$ is for frozen goods and $p_i = 2$ means chilled ones. The customer and warehouse locations according to their GPS coordinates you can see in Figure 1. The considered cargo capacity of all vehicles is 18 containers. The service time m_i for each customer i is given by 20 minutes plus one minute for each unloaded container q_i, i.e. $m_i = (20 + q_i)\ 60$ in seconds.

	Customers									
i	1	2	3	4	5	6	7	8	9	10
E_i (°)	16.24160	16.08722	14.69055	15.27025	16.66231	16.98885	14.90985	17.11870	15.58698	14.85888
N_i (°)	50.57513	50.27706	49.41570	49.95616	49.72949	49.95986	50.42491	49.46988	49.61639	49.99393
q_i	7	1	5	4	4	3	3	6	3	3
p_i	1	2	1	2	1	2	1	2	1	2
	Customers									
i	11	12	13	14	15	16	17	18	19	20
E_i (°)	14.74052	15.63217	15.36320	15.598478	15.30779	15.04006	15.22476	16.65998	15.99651	15.90839
N_i (°)	50.15575	50.35722	50.43471	49.42432	50.73916	50.76432	49.42156	49.49256	49.84431	50.02656
q_i	7	1	1	7	1	5	1	6	3	5
p_i	1	2	1	2	1	2	2	2	1	1
	Customers							Stores		
i	21	22	23	24	25	26	27	28	29	30
E_i (°)	14.4257	16.5257	16.4257	16.1565	16.3101	17.3520	17.3535	14.6359	16.2264	16.7487
N_i (°)	49.7105	50.0569	50.8749	50.7268	49.6891	50.8844	50.4806	51.0259	49.3061	50.6544
q_i	3	5	1	1	3	5	3	-	-	-
p_i	1	2	1	2	1	2	1	-	-	-

Table 1 The customers and stores GPS coordinates E_i , N_i , numbers of delivered containers q_i, priorities p_i

By running the *VRPMW_SOLVER_Z.m* function with above-given input parameters, an optimal solution which consists of 4 cycles and 3 paths is found and the sum of travel distances of all vehicles is minimal. All cycles and paths are shown in Figure 1. This means that the delivery of 97 containers to customers can be ensured by seven trucks, and the total distance traveled by vehicles distance is 1 839,585 km the sum of driving times (hh:mm:ss) of all seven trucks, including the service times for unloading goods, is 41:32:44. The list of customers on the cycle/path, numbers of delivered containers, priorities, total lengths, and durations can be found in Table 2.

If the priority in the delivery of goods is canceled, it is possible to reach it by canceling condition 15, which is achieved in the m-function by canceling all command line No. 20. Then, the optimal solution consists of 3 cycles and 3 paths. The total distance traveled by vehicles is reduced to 1 689,264 km, and the sum of driving times of vehicles is reduced to 39:06:09.

If frozen and chilled goods will be delivered separately, in the command on line No. 18, the condition `p(i)~=p(j)` will be replaced by the condition `p(i)>p(j)`. Then the optimal solution consists of 3 cycles and 3 paths such that the total distance traveled by vehicles is increased to 2 010,762 km, and the sum of driving times of vehicles is increased to 44:04:14.

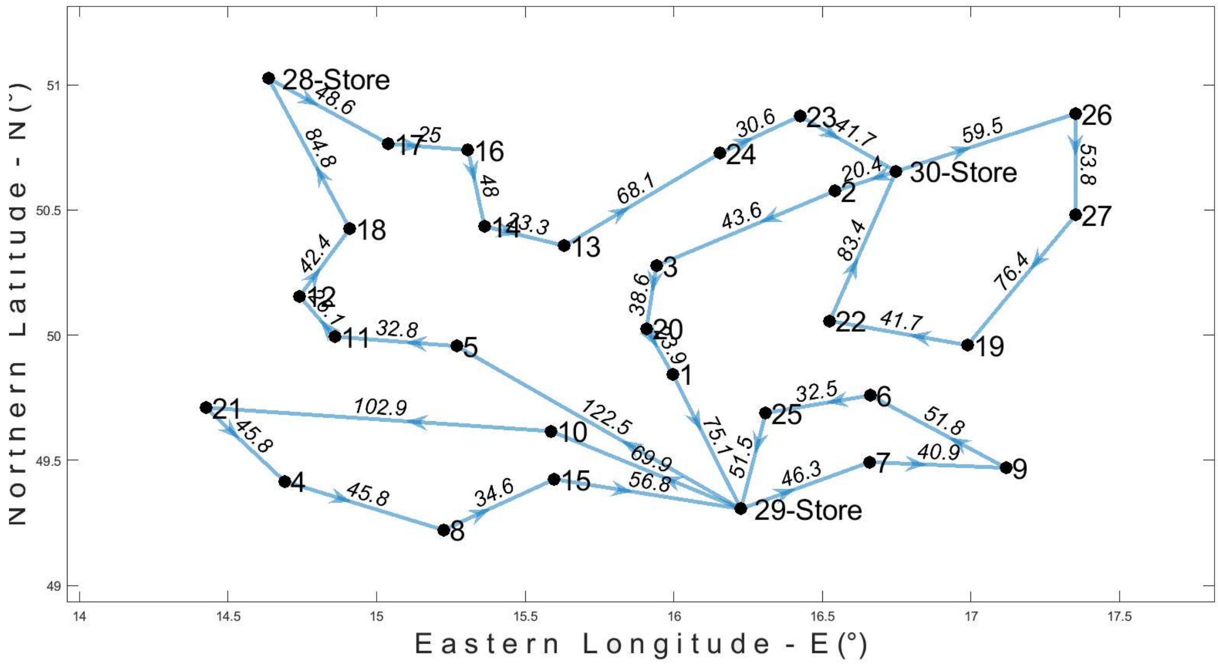


Figure 1 The optimal solution VRP for minimal traveling distance

No.		Start Store	Customers					Stop Store	Lengths (km)	Duration (hh:mm:ss)
1 Path	**i**	**28**	**17**	**18**	**12**	**4**	**8**	**29**	373.342	8:21:10
	q_i	-	1	6	1	4	6	-		
	p_i	-	2	2	2	2	2	-		
2 Cycle	**i**	**29**	**15**	**21**	**11**	**5**	**10**	**29**	374.607	8:16:03
	q_i	-	1	3	7	1	3	-		
	p_i	-	1	1	1	1	2	-		
3 Path	**i**	**29**	**25**	**1**	**20**	**2**	**-**	**30**	216.588	5:17:01
	q_i	-	3	7	5	1	-	-		
	p_i	-	1	1	1	1	-	-		
4 Cycle	**i**	**29**	**7**	**9**	**19**	**22**	**6**	**29**	309,669	7:00:25
	q_i	-	3	3	3	5	3	-		
	p_i	-	1	1	1	2	2	-		
5 Cycle	**i**	**30**	**27**	**26**	**-**	**-**	**-**	**30**	172.132	3:40:08
	q_i	-	5	5	-	-	-	-		
	p_i	-	1	2	-	-	-	-		
6 Cycle	**i**	**30**	**23**	**24**	**-**	**-**	**-**	**30**	124.304	2:46:17
	q_i	-	1	1	-	-	-	-		
	p_i	-	1	2	-	-	-	-		
7 Path	**i**	**30**	**3**	**13**	**14**	**16**	**-**	**28**	268.943	6:11:38
	q_i	-	5	1	7	5	-	-		
	p_i	-	1	1	2	2	-	-		
								Total	1 839.585	41:32:44

Table 2 The customers cycles and paths with respect priorities.

3 Conclusion

The set goal of developing an optimization model for solving the vehicle routing problem for multiple warehouses and prioritizing customers according to the temperature regime of the delivered goods was met. Optimization is realized by minimizing the sum of the distances traveled by all vehicles used to deliver goods. The completion of the created model is a software application, i.e. the creation of the M-function *VRPMW_SOLVER_Z.m*, which realizes the solution of the VRPMW problem for an arbitrary number of n customers and s warehouses. The created software was applied to an illustrative task with 3 warehouses and 27 customers. Its solution took about 15 minutes on a common laptop. The software is practically usable for up to 40 customers, but the processing time is increasing rapidly. The application of the created software in logistics practice enables the saving of vehicle distances traveled and thus the fuel consumed.

Acknowledgements

The paper was supported by the grant No. CZ.01.1.02/0.0/0.0/21_374/0027244 "Technology development for intelligent traffic flow management - 2nd part - optimization and extension" of Czech Ministry of Industry and Trade.

References

[1] Eksioglu, B., Vural, A.V. & Reisman, A. (2009). The vehicle routing problem: A taxonomic review. *Computers & Industrial Engineering,* 57, 1472-1483.

[2] Jonak, R., Smutný, Z., Simunek, M. & Dolezel, M. (2020). Rout and Travel Time Optimization for Delivery and Utility Services. *Acta Informatica Pragensia,* (2) 9, 200-209.

[3] Math Works. Inc. (2024). *Optimalization Toolbox^{TM}. User´s* Guide*.* Natick.

[4] Winston, W. L. (1994). *Operations Research. Applications and Algorithms.* Duxbury: Duxbury Press.

[5] Toth, P. & Vigo, D. (1998). *Exact algorithms for vehicle routing.* Boston: Kluwer Academic Publisher

[6] Zahrádka, J. (2023) The Exact Solution of Vehicle Routing Problem by Mixed Integer Linear Programming in Matlab. In J. Zahrádka, *Proceedings of 41st International Conference on Mathematical Methods in Economics* (pp, 433-438). Prague: Czech Society for Operations Research.

Appendix

```
function [X, y, TotLgth, TotDur] = VRPMW_SOLVER_Z(n, s, D, C, Q, q, p, m)
options = optimoptions('intlinprog', 'MaxTime', 3600, 'MaxNodes', 3000000);
A=zeros(n*n-n, (n+s)*(n+s)+n); Aeq=zeros(2*n+s,(n+s)*(n+s)+n); TotLgth=0;
k=0; for i=1:n for j=1:n
if i~=j k=k+1;A(k, (n+s)*(n+s)+i)=1;A(k, (n+s)*(n+s)+j)=-1; end; end; end
k=0; for i=1:n for j=1:n if i~=j k=k+1; A(k, (n+s)*(i-1)+j)=Q; end; end; end
for i=1:n*n-n for j=1:n if A(i,(n+s)*(n+s)+j)==-1 b(i,1)=Q-q(j); end; end; end
k=n*(n-1); for i=1:s for j=1:n
k=k+1;A(k,n*(n+s)+(i-1)*(n+s)+j)=Q;A(k,(n+s)*(n+s)+j)=-1;b(k,1)=Q-q(j);end;end
for i=1:n for j=1:n+s Aeq(i, (i-1)*(n+s)+j)=1; end; beq(i,1)=1; end
for i=1:n for j=i:n+s:(n+s)*(n+s) Aeq(n+i,j)=1; end; beq(n+i,1)=1; end
for i=1:s for j=1:n
Aeq(2*n+i, (n+s)*n+(i-1)*(n+s)+j)=1; Aeq(2*n+i,(n+s)*(j-1)+n+i)=-1;
end; beq(2*n+i,1)=0; end
lb=zeros((n+s)*(n+s),1); ub=ones((n+s)*(n+s),1);
for i=1:n ub(i+(n+s)*(i-1), 1)=0; end
for i=1:s for j=1:s ub((n-1)*(n+s)+n+i*(n+s)+j, 1)=0;  end; end
for i=1:n lb((n+s)*(n+s)+i, 1)=q(i); ub((n+s)*(n+s)+i,1)=Q; end
for i=1:n for j=1:n if p(i)>p(j); ub((n+s)*(i-1)+j)=0; end; end; end
ub=ub'; lb=lb'; DT=D';
f=DT(:);f((n+s)*(n+s)+1:(n+s)*(n+s)+n)=1; intcon=1:(n+s)^2+n;
intlinprog(f, intcon, A, b, Aeq, beq, lb, ub, [], options);
V=round(V); X=V(1:(n+s)*(n+s));
X=reshape(X, [n+s, n+s]); X=X'; y=V((n+s)*(n+s)+1: end) ; TotDur=sum(m);
for i=1:(n+s) for j=1:(n+s) if X(i, j)==1
TotLgth=TotLgth+D(i, j); TotDur=TotDur+C(i, j); end; end; end
end
```

Dependencies Between Criteria and Their Weights: A Reflective Analysis

František Zapletal[1], Miroslav Hudec[2], Miljan Vučetič [3], Marek Štěpán[4], Nina Barčáková[5]

Abstract. In multi-criteria decision-making (MCDM), the weights, expressing the importance of criteria for a decision-maker, are a crucial factor. Usually, the MCDM methods assume mutually independent criteria. On the other hand, in some situations, this assumption is too strong and could bring heavily distorted results. The literature provides some ways how to reflect such dependencies in the weights of criteria (e.g., CRITIC method or Choquet integral), but they suffer from weaknesses like the direction of dependence. This paper adopts the concept of fuzzy functional dependencies (FFDs) for criteria weights assignment. FFDs reveal not only the strength of the dependence but also its direction (causality); they can work with the uncertain input data (including qualitative variables or fuzzy numbers). The goal of this paper is to propose a model which uses FFDs to adjust the criteria weights in order to reflect the impact of dependencies. The model is verified using a real-life example.

Keywords: Fuzzy functional dependencies, weights, dependence, multi-criteria decision-making

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Decision making is a process that is absolutely elementary not only for managers, but for people and their development and survival in general. Many of the decision-making problems that people have to deal with fall into the category of multi-criteria decision-making (MCDM). That is, problems where a discrete countable set of options is evaluated based on their performance according to selected evaluation criteria.

One of the challenges of MCDM is to take into account the preferences of the decision maker when aggregating performance in terms of individual criteria. To this end, weights are typically used to quantify the decision maker's preferences. These weights can be determined in many different ways, both subjectively, e.g., Saaty's method (Saaty, 2003), and objectively, e.g., standard deviation method (Jahan et al, 2012). Most of these weighting methods, assume that the criteria are mutually independent, i.e., that the performance in one criterion does not imply the performance value of the same alternative in another criterion. However, such assumption is very difficult, if not even impossible, to meet in practical problems. The author of (Grabisch, 1995) demonstrates this fact using the student evaluation example. If a student is assessed based on the results of the test in several fields (courses), which are considered criteria there, it would be naive to assume that the results within the group of natural sciences or social sciences are absolutely independent. If a student is good at mathematics, it can be expected than his or her performance at physics will be also good (or at least better than at history, for instance).

In this research, two types of dependencies between criteria Figueira et al, 2009, are considered:

- If the performances of alternatives in two criteria are substantially positively dependent, their impact on the resulting order of the alternatives should be decreased for alternatives with good performance in both criteria (weakening) because the criteria are redundant to some extent.
- If the performances of alternatives in two criteria are substantially negatively dependent, their impact on the resulting order of the alternatives should be increased for alternatives with good performance in both criteria (strengthening) because such performance portfolio is rather rare.

An effect of the dependencies between criteria on their weights has already been explored by many researchers. In general, two different approaches can be distinguished: (a) Data-driven approach, (b) Expert opinion approach. The former is based on the calculations with values in the decision matrix (based on the performance values of the alternatives) where the dependence is given by the correlations between the performance values. The CRITIC

[1] VŠB-TU Ostrava, Faculty of Economics, Dept. of Systems Engineering and Informatics, Ostrava, Czech Republic, frantisek.zapletal@vsb.cz
[2] VŠB-TU Ostrava, Faculty of Economics, Dept. of Systems Engineering and Informatics, Ostrava, Czech Republic, miroslav.hudec@vsb.cz
[3] VŠB-TU Ostrava, Faculty of Economics, Dept. of Systems Engineering and Informatics, Ostrava, Czech Republic, miljan.vucetic@vsb.cz
[4] VŠB-TU Ostrava, Faculty of Economics, Dept. of Systems Engineering and Informatics, Ostrava, Czech Republic, marek.stepan.st@vsb.cz
[5] University of Economics in Bratislava, Faculty of Economic Informatics, Bratislava, Slovakia, nina.barcakova@euba.sk

method, see Alinezhad et al, 2019, is an example. This method adjusts the weights down (in case of the positive correlation between two criteria) or up (in case of the negative correlation). An approach based on expert opinions assumes that a decision maker is able to provide the input parameters possibly even without considering the decision matrix. Let us mention, for example the following ones:

- Analytic network process (ANP) Saaty et al, 2006) – the decision-maker expresses the strength of the dependence using the discrete Saaty's scale.
- Choquet integral (Grabisch, 1995) – the decision-maker expresses the dependence using the weights of "coalitions" of criteria.
- DEMATEL method (Alinezhad et al, 2019) – similarly to ANP, the adjustments are done using the discrete scale whose grades assess the dependence pair-wisely. The dependencies are averaged at the end.

All methods mentioned above can be reasonable for particular situations. If the dependence occurs, it is better to use any method which considers this fact than to ignore it. On the other hand, one must be aware of the weaknesses of each method. If the CRITIC method is used, one reduce the impact of subjectivity. On the other hand, it is necessary to take into account that the correlation is based only on the values in the decision-matrix (the real relationship can be biased by the selection of alternatives) and the direction of the dependence is missing (unlike, e.g., ANP or DEMATEL). Moreover, this method is more suitable for quantitative data, where no uncertainty or missing data occur, which is not always the case. The weakness of the expert opinion-based methods is clear. The quality of the output can be heavily influenced with the skills and knowledge of the experts.

This paper presents a new method which lies somewhere in between of two groups mentioned above. This means that it will also be, to some extent, driven by the data in the decision matrix. However, an additional assistance of the decision-maker and his or her subjective opinion will be required too. Namely, the method is based on the fuzzy functional dependencies (FFD) which use the fuzzy sets and fuzzy logic to map the mutual dependencies between the criteria, see Vučetić et al, 2020. By this method, the strength and direction of the dependence can be revealed. In addition, FFD is also very suitable for qualitative data or imprecise data. As for the role of subjectivity, the decision-maker is supposed to describe the values in the decision matrix using the ordinal scale by fuzzy sets. How alternative behaves considering other alternatives can be expressed numerically or even categorically. In the latter, it is harder to adopt traditional methods for adjusting weights, while FFD can be straightforwardly applied.

The main contribution of the proposed method is that it is based on the values in a decision matrix (not only on subjective evaluations by a decision-maker), it can easily handle qualitative or imprecise data, and considers also the direction of the relationship. On the other hand, its suitability for the quantitative precise data is lower (but it still can be used).

The remainder of the paper is organised as follows. Section 2 recalls the necessary theoretical background of the Fuzzy Functional Dependencies method. The core part of the paper is Section 3 where the algorithm of the proposed method for weights' adjustment, based on dependencies between criteria, is introduced. The applicability of the proposed method is checked using a numerical example and the results are then compared with some existing methods for a better comparison, see Section 4, while Section 5 concludes this work.

2 Theoretical background of Fuzzy Functional Dependencies

Fuzzy functional dependencies (FFDs) are an extension of classical functional dependencies Vucetic and Vujosevic 2012. While canonical functional dependencies deal with equality relation and precise matches between values in attributes, FFDs allow for a degree of uncertainty, fuzziness and approximate equality (i.e. *more or less equal*) of the attribute values. A canonical functional dependence (FD), denoted as $X \rightarrow Y$, indicates the existence of a relationship between two attributes (in our work criteria), X and Y. It states that if entities (or alternatives), t and t', share a common value within the domain of the attribute X, they will also share the same value within the domain of the attribute Y. In simpler terms, whenever $t\ [X]=t'[X]$, then $t[Y] = t'\ [Y]$. (for example, we might say that the employee ID determines the employee name indicating that if we know the ID, we can determine the name). However, there are situations where the classic FD is inadequate. For example, we can intuitively say that teachers show similar *practice length* when they present similar *technical certificates*. Next, we might notice that products with the similar prices correlate with similar customer satisfaction. If high score is considered over 85, then there is some similarity (or fuzziness) with a score of 83.

In FFDs, instead of requiring an exact match between attribute values, there is a notion of a degree of dependence between attributes. The equality relation is replaced by a measure of proximity between attribute values. This FFD is denoted as $X \xrightarrow{\theta} Y$ where θ represents the degree of dependence. This degree can range from complete dependence to partial or no dependence at all. Various definitions of FFDs were defined in Vucetic and Vujosevic

2012. FFDs are particularly useful in scenarios: (i) where data are uncertain or incomplete, such as categorical and fuzzy attribute values or multiple neighbouring values from discrete domains (i.e.{*little, medium*}) in a case of Likert scale surveys, (ii) when dealing with similar values in data processing tasks.

The FFDs offer a valuable tool for improving data analysis, query relaxation, record matching, and decision-making applications Caruccio et al 2015,Vučetić et al 2020. In this paper, we compute the degree of similarity between values over attribute (criterion) X as followsSözat and Yazici 2001:

$$C(X[t_i,t_j] = \min\left\{ \min_{x\in d_i}\{\max_{y\in d_j}\{s(x,y)\}\}, \min_{x\in d_j}\{\max_{y\in d_i}\{s(x,y)\}\} \right\} \tag{1}$$

where d_i represents the values (subset of attribute domain) of attribute X for entity t_i, d_j represents the values of X for entity t_j, both d_i and $d_j \in D$ (where D is a domain of acceptable values, e.g., categorical answers on the Likert scale or linguistic terms from a linguistic variable created over a numerical attribute), and $s(x, y)$ denotes a proximity relation between the values x and y within the domain D Shenoi and Melton, 1999.

The next step is recognising in which proportions of all pairs of entities t_i and t_j the fuzzy functional dependence is relevant. We calculate for which pairs of entities hold:

$$C(Y[t_i,t_j]) \geq \min(\theta, C(X[t_i,t_j]) \tag{2}$$

and finally, we compute the proportion of pairs of entities where fuzzy functional dependencies are satisfied. Design of algorithm for discovering FFDs using (1) is explained in Vucetic et al 2013.

3 FFD-based approach

At the beginning of the decision-making process, the domain expert assigns the initial weights to the criteria. However, the real values of the criteria for the entities considered influence the weights. When almost all values of alternatives for one criterion are similar, then the criterion becomes less relevant, i.e., the weight should be reduced. This can be solved by applying an entropy or standard deviation that adjusts the weights Aggarwal 2015. The next case influencing weights is the correlation between two criteria. A strong positive correlation weakens the importance of the criterion, meanwhile, a strong negative correlation strengthens it, e.g., the CRITIC method, see Alinezhad et al 2019. A weakness is in non-detecting the direction of influence in order to asymmetrically reduce the weights. Next, non-monotonic correlation is not revealed.

On the other hand, FFDs are able to record not only the existence of a nonlinear dependence, but also its direction. FFDs have shown their benefits in various data mining applications, for example Vučetić et al 2022. In this work we explore FFDs in adjusting weights of criteria. The classical functional dependencies are not able to reveal that a small modification in values keeps more or less significant dependence or a very high dependence. This is another reason to apply FFDs. Let us consider Tab. 1 of recognised FFDs for $\theta = 0.7$.

Table 1 FFD among criteria.

FFD	a	b	c	d	e	I
a	1	0.6	0.3	0.8	0.2	I_1
b	0.5	1	0.3	0.8	0.2	I_2
c	0.3	0.6	1	0.8	0.2	I_3
d	0.6	0.6	0.3	1	0.2	I_4
e	0.4	0.6	0.3	0.8	1	I_5
D	D_1	D_2	D_3	D_4	D_5	

A row indicates to what extent a criterion influences the others. The criterion is more relevant with higher values (value I_j). On the contrary, columns indicate how much the criterion is influenced by other criteria. The relevance of the criterion decreases (value D_j) with the higher values. We should formalise these observations in order to modify initially assigned weights. In Tab. 1, one can see that criteria c and e are almost symmetrically lightly influenced, while the relationship between d and e is heavily asymmetric.

Formalisation of values I_i. An averaging function is suitable in this case. Namely, we apply arithmetic mean as a logically neutral aggregation function, i.e., increased influence on one attribute is compensated by the decreased influence of another attribute. The geometric mean is not suitable, due to existence of an annihilator (one zero value causes that the total influence of a considered criterion becomes 0).

$$I_j = \frac{1}{n} \sum_{l=1}^{n} w_{jl} \tag{3}$$

Formalisation of values D_i. In this case, the disjunction function like the *max* function are reasonable to handle the case when a criterion is fully influenced by any other one, and becomes irrelevant, which is reflected by value 1, see 4.

$$D_j = \max_{i=1,\dots,n} (w_{ij}) \tag{4}$$

Finally, we need the difference (5). Parameter γ adjusts initial weight w_j, which has to be normalized at the end using (6).

$$\gamma_j = \theta(I_j - D_j) \tag{5}$$

$$w_j^* = \frac{(1+\gamma_j) * w_j}{\sum_{j=1}^{n}(1+\gamma_j) * w_j}. \tag{6}$$

A value higher than 0 increases weight, while a value lower than 0 decreases weight.

4 Numerical example

Data concerning the evaluation of 10 teachers (T1 to T10) at the second level of elementary school were selected for this paper. Each teacher is described using 7 specified criteria, including Experience (C1), Languages (C2), Subjects (C3), Hours (C4), Social Certificates (C5), Technical Certificates (C6), and Courses (C7). The decision matrix can be found in Tab. 2.

The criterion of Experience indicates the duration for which the teacher has been in the field. Languages specify how many languages the teacher masters. The criteria Subjects and Hours indicate the amount of subjects taught and the amount of hours taught per week. The criteria Social and Technical Certificates provide the quantity of certificates owned by the teacher related to the respective areas (social or technical aspects). Courses indicate which extracurricular activities the teacher organises.

The individual attributes are defined on 7 respective domains using linguistic expressions:

- Dom(Experience) = {beginner (B), little experienced (LE), moderately experienced (ME), highly experiencd (HE)},
- Dom(Languages) = {limited (L), significant (S)},
- Dom(Subjects) = {few (F), many (MA)},
- Dom(Hours) = {few (F), middle (MI), many (MA)},
- Dom(Social Certificates) = {few (F), middle (MI), many (MA)},
- Dom(Technical Certificates) = {few (F), middle (MI), many (MA)},
- Dom(Courses) = {easy (E), demanding (D), very demanding (VD)}.

The performance of each alternative is assessed using one or two neighbouring grades from the domains above. Two grades indicate the uncertainty in the evaluation, i.e., borderline case between beginner and little experienced. Based on the methodology described in Sec. 2 and 3, and under $\theta = 0.7$, FFD between criteria are calculated, see Tab. 3. The initial weights $\boldsymbol{x}$ are assumed equal, i. e., $\frac{1}{7}$ for all criteria.

4.1 Comparison with the CRITIC method

The CRITIC method is based on the stochastic analysis of the data in the decision matrix Alinezhad et al, 2019. It is a purely data-driven approach. Namely, the weights are given using the standard deviations of the performance values in the given criterion across all alternatives, and then adjusted by correlation coefficients, see (7). To make the results comparable with the proposed approach, equal weights are used instead of the ones given by standard deviations (and the corrections implied by correlations will be applied to these equal weights). Another correction of the original CRITIC method is due to the type of data. Since all input data in the considered numerical example are qualitative, Spearman's correlation coefficient is applied instead of Pearson's one

Table 2 Dataset expressing behaviour of each teacher by the considered criteria

Alternative	C1	C2	C3	C4	C5	C6	C7
T1	HE	L	MA	F	F, MI	MA	E
T2	ME	S	F, MA	F, MI	MA	F, MI	E
T3	HE	L	M	F, MI	F	F, MI	D
T4	B	L, S	MA	MI	F	F	E
T5	ME, HE	L	MA	MI	F	F	E
T6	ME	L	MA	F, MI	MA	F, MI	VD
T7	ME	L, S	F, MA	MI	F	F	E
T8	ME	L	MA	F, MI	F, MI	F	D
T9	ME, HE	L	MA	F, MI	F, MI	F, MI	E
T10	ME, HE	L	F, MA	F, MI	F	F	E

Table 3 FFD and resulting weights for the numerical example

	C1	C2	C3	C4	C5	C6	C7	I	w^*	Δ
C1	1	0.47	0.4	0.36	0.22	0.27	0.36	0.35	0.14	-1.02%
C2	0.11	1	0.33	0.24	0.16	0.16	0.18	0.20	0.13	-9.77%
C3	0.49	0.49	1	0.64	0.31	0.36	0.49	0.46	0.13	-11.62%
C4	0.53	0.49	0.78	1	0.31	0.36	0.49	0.49	0.14	-0.49%
C5	0.11	0.36	0.49	0.47	1	0.38	0.36	0.36	0.16	12.09%
C6	0.42	0.36	0.6	0.67	0.38	1	0.38	0.47	0.17	20.57%
C7	0.13	0.22	0.27	0.2	0.16	0.2	1	0.20	0.13	-9.77%
D	0.53	0.49	0.78	0.67	0.38	0.38	0.49			

$$w_i = \frac{\sigma_i \sum_{l=1}^{n}(1-\rho_{il})}{\sum_{j=1}^{n}\sigma_j \sum_{l=1}^{n}(1-\rho_{jl})} \tag{7}$$

The resulting weights can be found in Tab. 4. We see that the resulting adjustments of the original weights caused by dependencies differs for both methods, despite the values of correlation coefficients are similar for many comparisons with FFD in at least one direction (e.g., FFD $C6 \rightarrow C1 = 0.42$, $\rho_{C6,C1} = 0.378$, FFD $C1 \rightarrow C6 = 0.27$). The reason is two-fold. First, FFDs do not distinguish the polarity of the dependence. Contrary, a correlation coefficient does not distinguish the direction of the dependence. The logic, under which the adjustments are done, is also different. The proposed approach strengthens the weight with increasing ability of a criterion to explain other criteria, meanwhile the CRITIC method increases the weight with a stronger negative correlation between the criterion and other criteria. The 'dependence' is perceived completely differently by both measures. In spite of this, at least some similarities in the pattern in the results could be expected.

Generally, the results are promising, but the further investigation is necessary, because of the different nature of adjustment calculation in order to provide clearer recommendations of suitability for diverse real-life situations. For instance, CRITIC recognised partial negative correlation between $C1$ and $C2$, while FFD recognised that this partial correlation is stronger in direction $C1 \rightarrow C2$ than $C2 \rightarrow C1$. On the other hand dependencies between $C3$ and $C4$ exhibit a significant difference.

5 Conclusions

This paper presents a novel method for considering the dependencies between the criteria when setting their weights, based on the FFD concept. The effects of dependencies in both directions have been carefully investigated, and suitable, meaningful, and traceable ways of their aggregation have been proposed. The algorithm has been applied to a real-life numerical example, and the results have been compared with those of the well-established CRITIC method, which works with correlations between the criteria. This experiment showed that both methods can lead in some cases to similar, while in other cases to different results just because of the different perception of the dependencies there. The additional information, directing of dependence, is valuable for a finer tuning of weights.

Table 4 Spearman's correlation coefficients and weights by the CRITIC method

	C1	C2	C3	C4	C5	C6	C7	w^*	Δ
C1	1.000							0.144	0.78%
C2	-0.466	1.000						0.160	11.95%
C3	0.209	-0.553	1.000					0.136	-4.89%
C4	-0.427	0.275	-0.152	1.000				0.165	15.80%
C5	-0.104	0.213	-0.131	-0.110	1.000			0.123	-14.12%
C6	0.378	-0.217	0.132	-0.942	0.030	1.000		0.151	5.94%
C7	0.007	-0.364	0.452	0.000	0.645	-0.111	1.000	0.121	-15.46%

This gives rise to the need for further investigation to be able to provide a clear recommendation on which method is more suitable, under what conditions, and in which situations. Then, other competing approaches to modelling dependencies between criteria, like DEMATEL or Choquet integral, could be considered for comparison.

Acknowledgements

This article has been produced with the financial support of the European Union under the REFRESH – Research Excellence For REgion Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition and the SGS project no. SP2024/083.

References

Aggarwal C (2015) Data Mining: The Textbook. Cham: Springer

Alinezhad A, Khalili J, Alinezhad A, et al (2019) Critic method. New Methods and Applications in Multiple Attribute Decision Making (MADM) pp 199–203

Caruccio L, Deufemia V, Polese G (2015) Relaxed functional dependencies—a survey of approaches. IEEE Transactions on Knowledge and Data Engineering 28(1):147–165

Figueira JR, Greco S, Roy B (2009) Electre methods with interaction between criteria: An extension of the concordance index. European Journal of Operational Research 199(2):478–495

Grabisch M (1995) Fuzzy integral in multicriteria decision making. Fuzzy sets and Systems 69(3):279–298

Jahan A, Mustapha F, Sapuan S, et al (2012) A framework for weighting of criteria in ranking stage of material selection process. The International Journal of Advanced Manufacturing Technology 58:411–420

Saaty TL (2003) Decision-making with the ahp: Why is the principal eigenvector necessary. European journal of operational research 145(1):85–91

Saaty TL, Vargas LG, et al (2006) Decision making with the analytic network process, vol 282. Springer

Shenoi S, Melton A (1999) Proximity relations in the fuzzy relational database model. Fuzzy Sets and Systems 100:51–62

Sözat Mİ, Yazici A (2001) A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations. Fuzzy Sets and Systems 117(2):161–181

Vucetic M, Vujosevic M (2012) A literature overview of functional dependencies in fuzzy relational database models. Technics Technologies Education Management-TTEM 7(4):1593–1604

Vucetic M, Hudec M, Vujošević M (2013) A new method for computing fuzzy functional dependencies in relational database systems. Expert Systems with Applications 40(7):2738–2745

Vučetić M, Hudec M, Božilović B (2020) Fuzzy functional dependencies and linguistic interpretations employed in knowledge discovery tasks from relational databases. Engineering Applications of Artificial Intelligence 88:103395

Vučetić M, Brokešová Z, Hudec M, et al (2022) Financial literacy and psychological disaster preparedness: applicability of approach based on fuzzy functional dependencies. Information Processing & Management 59(2):102848

The 42nd International Conference on

MATHEMATICAL METHODS IN ECONOMICS

MME 2024

ČSOV

UNIVERZITA J. E. PURKYNĚ V ÚSTÍ NAD LABEM

UNIVERZITA J. E. PURKYNĚ V ÚSTÍ NAD LABEM
Přírodovědecká fakulta

PROCEEDINGS OF THE 42ND INTERNATIONAL CONFERENCE ON **MATHEMATICAL METHODS IN ECONOMICS**