Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

UNCOVERING COMPETITOR PRICING PATTERNS IN THE DANISH PHARMACEUTICAL MARKET VIA SUBSEQUENCE TIME SERIES CLUSTERING: A CASE STUDY

Ruhollah Jamali

Maersk Mc-Kinney Moller Institute
University of Southern Denmark
Campusvej 55
Odense, 5230, DENMARK

Sanja Lazarova-Molnar

Institute AIFB
Karlsruhe Institute of Technology
Kaiserstr. 89
Karlsruhe, 76133, GERMANY

ABSTRACT

Adopting data-driven decision-making approaches can significantly enhance profitability and foster growth in economic situations through quantitative analysis of market dynamics. One intriguing market that warrants examination is the price competition observed within the Danish pharmaceutical sector, where numerous companies are vying for a larger market share through the offering of diverse pharmaceutical products. This paper aims to shed light on this market by employing subsequence time series clustering techniques to identify pricing patterns among the players involved in the Danish pharmaceutical industry. The data analysis pipeline performed in this study allows for the identification of price patterns for clustering and discovering different agent groups, as well as providing a foundation for expanding the current agent-based model of the European pharmaceutical parallel trade market by analyzing the pricing behavior and patterns of players, facilitating the utilization of historical data to model agent behavior and advancing research in this area.

1 INTRODUCTION

The pharmaceutical industry is a highly dynamic and complex marketplace with unique challenges and opportunities. One such challenge arises from the exclusive rights granted to patent holders of a medicine, which allows them to prevent other companies from manufacturing the same medicine. However, in Europe, licensed wholesalers are permitted to resell medicines in countries where the patent holder is selling their medicines at a higher price, resulting in price competition. Another player involved in this market is parallel traders, who are individuals or companies that take advantage of price differences between different countries to purchase patented medicines in one country and sell them in another. Understanding the competitive landscape of the pharmaceutical market is crucial for companies to develop effective pricing strategies, identify potential opportunities for cost savings, and maintain their competitive edge. Moreover, modeling price competition in the pharmaceutical market can provide insights for economists and regulators to understand the behaviors of companies and how they compete with each other. This information can help regulators to design policies that promote competition, protect consumers from monopolistic practices, and ultimately lead to lower medicine prices.

In this paper, we focus on the Danish pharmaceutical market as a case study to explore the application of data science and, specifically, time series analysis in developing an agent-based model of the market. Here we demonstrate how analyzing historical data can provide valuable insights into market dynamics and competitive behavior. In Denmark, pharmaceutical companies, including patent holders and parallel traders, have to set medicine prices for 14-day periods (fortnights) and report their prices to the Danish Medicines Agency. At the same time, pharmacies must offer the medicine brand at the lowest available price each fortnight to promote fair competition and ensure affordable access to medicines for patients. For

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

instance, if one pharmaceutical company sets the price of a particular medicine at 10 monetary units, while another company sets the price at 15 monetary units, pharmacies are obligated to provide the medicine from the first company to customers at the price of 10. However, in the event that the first company is unable to fulfill the demand for the requested medicine, pharmacies are permitted to offer the medicine from the second company at its price of 15. Therefore, understanding competitors' pricing behavior in the Danish pharmaceutical market is crucial for companies to develop effective pricing strategies, identify potential opportunities for cost savings, and maintain their competitive edge.

Agent-based modeling and simulation (ABMS) is a powerful computational modeling technique for simulating individual agents' behavior and interactions within a system (Axtell and Farmer 2022). ABMS can capture the complexity and heterogeneity of economic systems that traditional econometric methods may not be able to capture. In economics, ABMS finds diverse applications, including exploring financial market dynamics (Feng et al. 2012; Samanidou et al. 2007) and labor market dynamics (Ng and Kang 2012). Furthermore, ABMS has been used to explore the behavior of financial markets and identify potential risks and vulnerabilities (Bookstaber 2012). By modeling individual agents and their interactions, ABMS can provide a more realistic representation of pricing behavior than traditional models that rely on simplified assumptions and generalizing players' characteristics. This approach enables economists to gain valuable insights into financial market dynamics by simulating the behaviors of individual traders and investors and studying how they interact with one another.

In addition, time series clustering (TSC) is a set of data analysis techniques that identify patterns in time series data, providing insights into economic phenomena such as business cycle phases, economic growth, and financial market volatility (Aghabozorgi et al. 2015). TSC has been used to identify business cycle phases by analyzing patterns in economic data, such as GDP growth (Yu et al. 2012), unemployment rates (Huang et al. 2020), and inflation rates (Hendrawati et al. 2021). By identifying patterns in economic data, economists can identify the phases of the business cycle, such as expansion, peak, contraction, and trough. TSC has also been used to identify patterns in financial market data, such as stock prices (Nair et al. 2017), exchange rates (Sun et al. 2018), and interest rates. Aghabozorgi et al. (2015) classify TSC methods into three categories: whole time-series clustering, subsequence time series clustering, and time point clustering. In whole time-series clustering, the goal is clustering a set of individual time series based on their similarity, while in subsequence time series clustering, the target is clustering segments from a single long time series. Time point clustering, on the other hand, clusters time points based on their temporal proximity and the similarity of corresponding values.

Subsequence time series clustering can also be applied to analyze pricing behavior in the Danish pharmaceutical market. By clustering similar subsequences of available pricing history for patented medicines, it is possible to identify distinct pricing patterns and trends. Subsequence time series clustering can be used to compare the pricing behavior of competing pharmaceutical companies, providing valuable insights into the competitive dynamics of the market. By utilizing subsequence time series clustering, not only pharmaceutical companies operating in Denmark can optimize their pricing strategies and gain a competitive advantage in the market. But also, the results of subsequence time series clustering can be used as inputs to an agent-based model of the Danish pharmaceutical market. In the Danish pharmaceutical market, participants do not engage in direct interactions. Instead, they rely on a hidden interaction, where they check each other's prices and medicine availability. The subsequence time series analysis enables us to understand the nuanced dynamics of the market, enhancing the realism and effectiveness of agent-based modeling of this market. This can help to evaluate the effectiveness of different pricing strategies in a controlled environment and identify the optimal pricing strategy to achieve desired outcomes, such as increased market share or profitability.

In this paper, we show how to apply subsequence time series clustering on historical pricing data of the Danish pharmaceutical market to get insight for further developing the available agent-based model of this market. The model enables us to simulate the behavior of competing pharmaceutical companies in response to changes in pricing strategies and evaluate the effectiveness of different pricing strategies in achieving desired outcomes such as increased market share or profitability. To accomplish this, we first review the literature on the application of agent-based modeling and simulation in economics and financial markets, followed by a discussion of the agent-based model of pharmaceutical parallel trade markets (Section 2). We then apply a subsequence time series clustering method to the available data of the Danish pharmaceutical market (Section 3) and present a case study that analyzes the pricing behavior of the players involved in this market for two popular medicines (Section 4). Finally, we summarize the findings and discuss how this data analysis can inform future research and contribute to the development of a more realistic agent-based model of the pharmaceutical parallel trade market (Section 5).

2 BACKGROUND

2.1 Agent-based Modeling in Economics

Agent-based modeling (ABM) has gained significant popularity in microeconomics, especially in modeling markets. Traditional microeconomic models, such as the supply and demand model, are based on the assumption that markets operate at the intersection point. However, ABMs provide a more realistic approach by modeling the behavior of individual agents in the market. ABMs can predict market efficiency by examining how agents behave in realistic market contexts, offering a bottom-up approach to understanding economic phenomena and providing a comprehensive understanding of the behavior of purposeful individuals in markets. A recent review paper, Axtell and Farmer (2022) investigated the application of ABM in economics and finance, highlighting the wide range of areas where ABM has been applied, including pricing. Research in this area has explored different subjects, such as the impact of pricing strategies on retailers in functional product markets (Arvitrida et al. 2019), examining the pricing strategies of adaptive retailers competing in the presence of complex consumer behavior (Du and Xiao 2019), and analyzing the response of consumers to real-time prices in electricity retail markets (Yousefi et al. 2011). ABM has also been used to analyze the impact of imbalance pricing mechanisms on market behavior in electricity balancing markets (Van Der Veen et al. 2012), explore dynamic pricing strategies (Cornacchione et al. 2023), evaluate water-pricing policies (Athanasiadis et al. 2005), and studying pricing mergers and acquisitions (Agarwal and Kwan 2017).

2.2 Subsequence Time Series Clustering

Agent-based modeling and simulation has shown to be an effective tool for modeling economic and financial markets. However, the multitude of elements involved in economic situations poses a significant challenge to the creation of such models. One way to this challenge is to make the model data-driven, based on observations from historical data. In order to achieve this, subsequence time series clustering can be used to identify patterns in the data and optimize clustering scores. Subsequence time series clustering is a technique used to partition a time series dataset into homogeneous groups based on the similarity of their subsequences. Subsequence time series clustering involves breaking down the time series data into smaller subsequences and using clustering algorithms to identify patterns in the data. The resulting clusters can then be used to achieve a better understanding of the data, leading to the development more insightful model. Subsequence time series clustering demonstrated application in multiple fields such as anomaly detection (Li et al. 2021; Blázquez-García et al. 2021), e-commerce (Aghabozorgi and Teh 2014), speech recognition (Fong 2012), and biology (Lonardi and Patel 2002).

2.3 Modeling Pharmaceutical Parallel Trade Market

The pharmaceutical parallel trade market has been studied from various perspectives, and researchers have employed different methods and techniques to understand the market dynamics. Previous studies have investigated the impact of parallel trade on medicine prices and welfare (Kanavos and Costa-Font 2005), the states involved in the market (Kanavos et al. 2004), the potential effects on innovation and research and development (Bennato and Valletti 2014), and the profitability of innovating pharmaceutical companies and retailers (Dubois and Sæthre 2020). Among the different methods in these studies, modeling and simulation methods have emerged as powerful approaches to understanding the complexity of the pharmaceutical parallel trade market. Game theory models have been widely used to study markets, including the pharmaceutical trade market (Pecorino 2002; Müller-Langer et al. 2007; Guo et al. 2013; Vandoros and Kanavos 2014; Gnecco et al. 2018). These models provide a framework for analyzing the strategic behavior of different players in the market and predicting their outcomes under different scenarios. By considering the interactions between pharmaceutical companies, distributors, regulators, and consumers, game theory can help identify the key factors driving market outcomes and inform policy decisions aimed at promoting efficiency and competition. However, agent-based modeling and simulation can offer several benefits over traditional game theory models for studying pharmaceutical parallel trade markets. Firstly, agent-based models can capture the complexity of individual behavior and interactions in the market, including the heterogeneity of agents such as patients, pharmacists, and wholesalers and how they respond to changes in market conditions. Secondly, agent-based models can incorporate real-world data and specific details about the market, such as regulatory policies and pricing structures, which can lead to more accurate and realistic predictions of market outcomes. Finally, agent-based models can allow for the exploration of "what-if" scenarios, which can inform policy decisions and guide strategies for market participants.

One of the first efforts to model the pharmaceutical parallel trade market is Pecorino's work (Pecorino 2002) where he made a game theoretic model to investigate the impact of the parallel trade in pharmaceutical companies considering two counties. Further, in another effort to model the European pharmaceutical market, an agent-based model of this market was developed (Jamali and Lazarova-Molnar 2022b) based on the game theoretic model to demonstrate the capabilities of agent-based modeling and simulation compared to game theory in this topic. The agent-based model was also considered to model a market involving two countries first to investigate the replication of game theoretic model equilibriums and secondly to demonstrate the ability of an agent-based model compared to the game theory model in further research (Jamali and Lazarova-Molnar 2022a).

The agent-based model of the pharmaceutical parallel trade market involves three types of agents - government, manufacturer, and parallel trader - and its environment consists of two countries. The model begins with the negotiation of the medicine price between the government and the manufacturer in the first country, which is modeled as a Nash bargaining game. Next, the parallel traders in the second country check the price of the medicine in the first country, taking into account a transfer cost, which includes the cost of repackaging and transferring the medicine to be sold in the second country. If a parallel trader can make a profit by selling the medicine in the second country, they will participate in the competition with each other and the manufacturer. The competition is modeled as Cournot competition in the game theory model (Pecorino 2002; Guo et al. 2013) and the first version of the agent-based model (Jamali and Lazarova-Molnar 2022b) of the market. This procedure repeats every step of the agent-based model to reach an equilibrium price in the second country.

In another research (Jamali and Lazarova-Molnar 2022a) to develop the agent-based model, the competition in the second country was changed further to investigate the abilities of the agent-based model. Here in this paper, we are aiming to focus on competition happening in the second country and model it as a price competition, considering the price competition over parallel imported medicines in Denmark. We will use available historical data on the price competition in Denmark and analyze them employing the subsequence time series clustering technique to investigate how to develop a data-driven agent-based model of the price competition.

3 ANALYZING PHARMACEUTICAL PRICING BEHAVIOR IN DENMARK

Danish Health and Data Protection Agency's data bank with health data is called eSundhed, and it aims to provide availability and visibility of health data to enable better insights into the Danish healthcare system. On eSundhed's website, data related to pharmacies' sales of medicines in Denmark is available for every quarter of the past five years. The data consists of the following properties for each medicine active substance: total quantity sold, total revenue, average shipping price per unit of quantity, the regions' health subsidy, pharmacies' total purchase price, pharmacies' purchase price per unit, and the number of people who bought the medicine. Generally, a medicine contains one or more active substances, which are the chemical components that produce a therapeutic effect on the body. In this case study, we focus on the most popular medicines in Denmark, which we selected as we explain in the following. Since the same medicine is sold under different names, we extracted the total quantity sold for all active substances in the eSundhed database to find the most popular medicines. Then using the most popular active substances list extracted from the database, we were able to list the top ten popular parallel traded medicines in Denmark.

Considering the extracted list of popular medicines in Denmark, we found that there were 36 different names or brands (sold by multiple companies) for the top ten medicines. To further our analysis, we extracted the price development for each of these 36 names since 1998. We utilized a database provided by the Danish Medicine Agency called (Medicinpriser.dk) to obtain this data. This database provides up-to-date information on the prices of medicines, updated every 14 days (fortnight). It is important to note that the length of price history for each name in this dataset varies, as some companies have been involved in the market for a longer period of time than others. The shortest price history in this dataset covers a period of 4 fortnights, while the longest history covers a period of 546 fortnights.

To cluster the pricing data of the 36 extracted medicine names, we first created a dataset from all available pricing data from all companies that participated in the market. Since medicines' price ranges vary in our dataset and we are aiming to find pricing patterns, we normalized the data into the range between 0 and 1. Then we divided the time series into subsequences with lengths ranging from 4 to 12 as expert knowledge suggested that pharmaceutical pricing in Denmark considers at least two months and at most six months of forecast. We used the tslearn (Tavenard et al. 2020) Python library to apply K-means clustering to the dataset, transforming the time series data into a suitable feature space and identifying clusters of similar time series using the k-means algorithm. To initialize the cluster centers, we specified the number of clusters, the metric to be used for cluster assignment, and a random seed. We ran the code for a range of 3 to 6 clusters, dynamic time wrapping as the metric, and a fixed random seed. Our dataset consisted of $51,950$ subsequence time series data, and the number of subsequences in each cluster is depicted in Figure 1.

In our analysis, we view a subsequence time series as a representation of the pricing behavior of a player over a specific period. Given that a time series of length 100 can generate over 800 subsequences, it is reasonable to assume that many of these subsequences provide a better representation of a specific behavior over their period. To identify the most informative subsequences, considering each of all 36 time series available, we utilized a greedy algorithm to generate the entire time series from the set of its subsequences. We did this by considering the distance of each subsequence from the center of its respective cluster and selecting the best set of non-overlapping subsequences that can cover the entire time series with the minimum distance from their respective cluster centers.

After regenerating the time series from their subsequences, we compared the results obtained using different numbers of clusters. Our observations revealed that when the number of clusters was set to 5, there was a limited selection of subsequences from one of the clusters during the regenerating process. Similarly, increasing the number of clusters to 6 resulted in a scarcity of selected subsequences from two clusters during regeneration. These observations suggest that these clusters do not adequately capture meaningful patterns in the time series data. Therefore, considering 4 clusters appears to provide more meaningful and representative subsequences for the regeneration process. It is important to acknowledge that the dataset may not be sufficiently large for a definitive conclusion and further research using a larger dataset is required to determine the optimal number of clusters that accurately represent the pricing behavior in the market.

In order to gain insights into the pricing behavior of pharmaceutical companies within each cluster, it is necessary to visualize some of the subsequences in each cluster. By representing each subsequence time series as a line plot, with the x-axis representing time and the y-axis representing the normalized price of the drug, we can obtain a visual representation of how the companies set medicine prices over time within each cluster. To achieve this, we selected the set of best subsequences representing companies' pricing behavior again by using the greedy algorithm. However, in order to plot them, we needed to ensure that

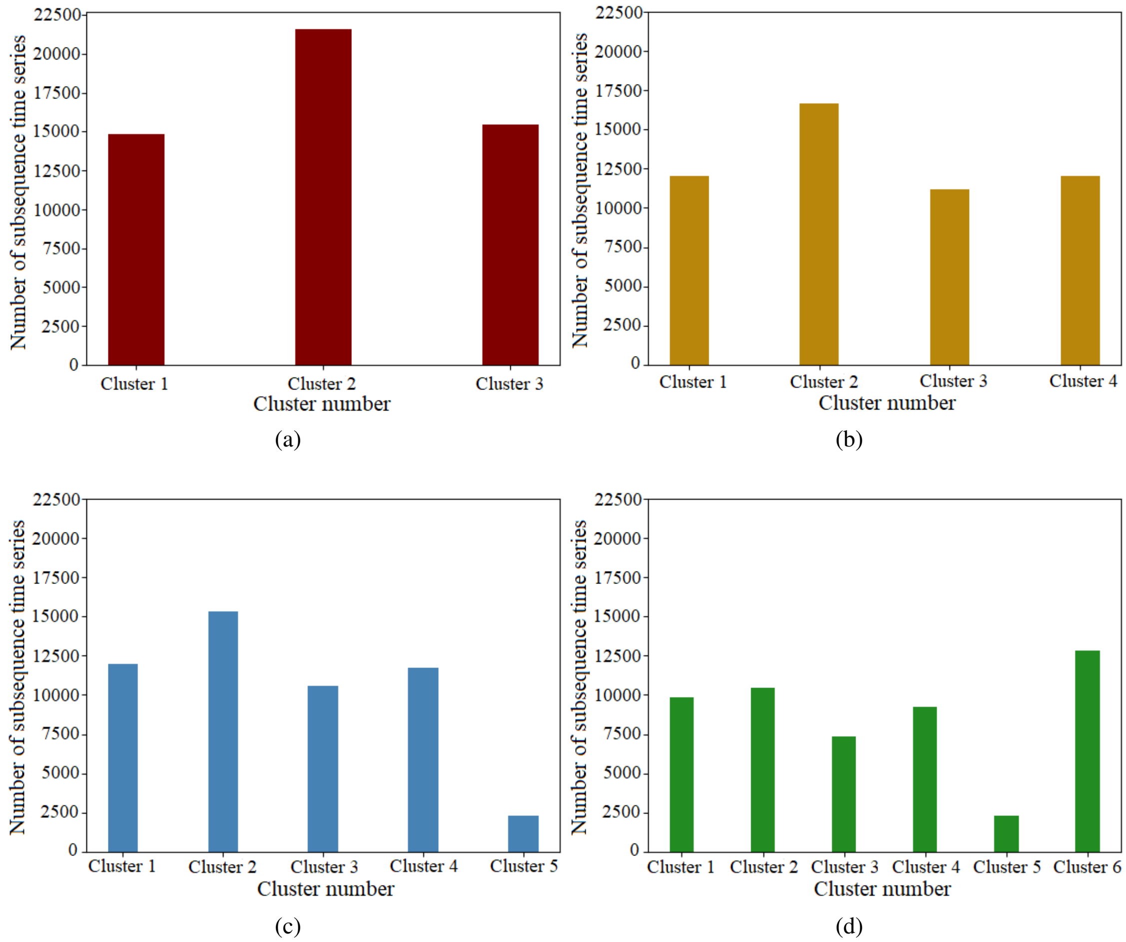


Figure 1: Cluster size of k-means clustering of subsequence time series Danish medicine price with (a) three, (b) four, (c) five, and (d) six clusters.

their lengths were equal. Therefore, we employed linear interpolation to monolithically increase the sample points for all subsequences to a length of 12. The resulting figure, shown in Figure 2, displays each cluster's subsequences, where each gray line represents a subsequence time series, and the red line represents the cluster center. Figure 2 demonstrates the density of the clusters and the distribution of the subsequence time series around their respective cluster centers. The plot shows that the clusters are well-formed and have a significant density of subsequence time series around their centers, indicating that the clustering algorithm has successfully captured the pricing patterns. Additionally, each cluster contains only a few anomaly behaviors, which suggests that the clustering is robust and can effectively differentiate between distinct pricing patterns.

Our analysis of the Danish pharmaceutical market reveals four main pricing patterns, which were identified through the subsequence time series analysis of historical pricing data of this market. Our visualization of the clustering results, presented in Figure 2, shows that cluster 1 (subfigure (a)) corresponds to the pricing pattern where the price of the medicine is high. This could happen because of a shortage of medicine, where companies increase their prices to maximize profit. This aligns with economic theory, where higher demand than supply leads to a price increase. Alternatively, this pricing pattern could demonstrate the pricing strategy of a non-competitive participant in the market who seeks a good profit margin rather than competing with other companies.

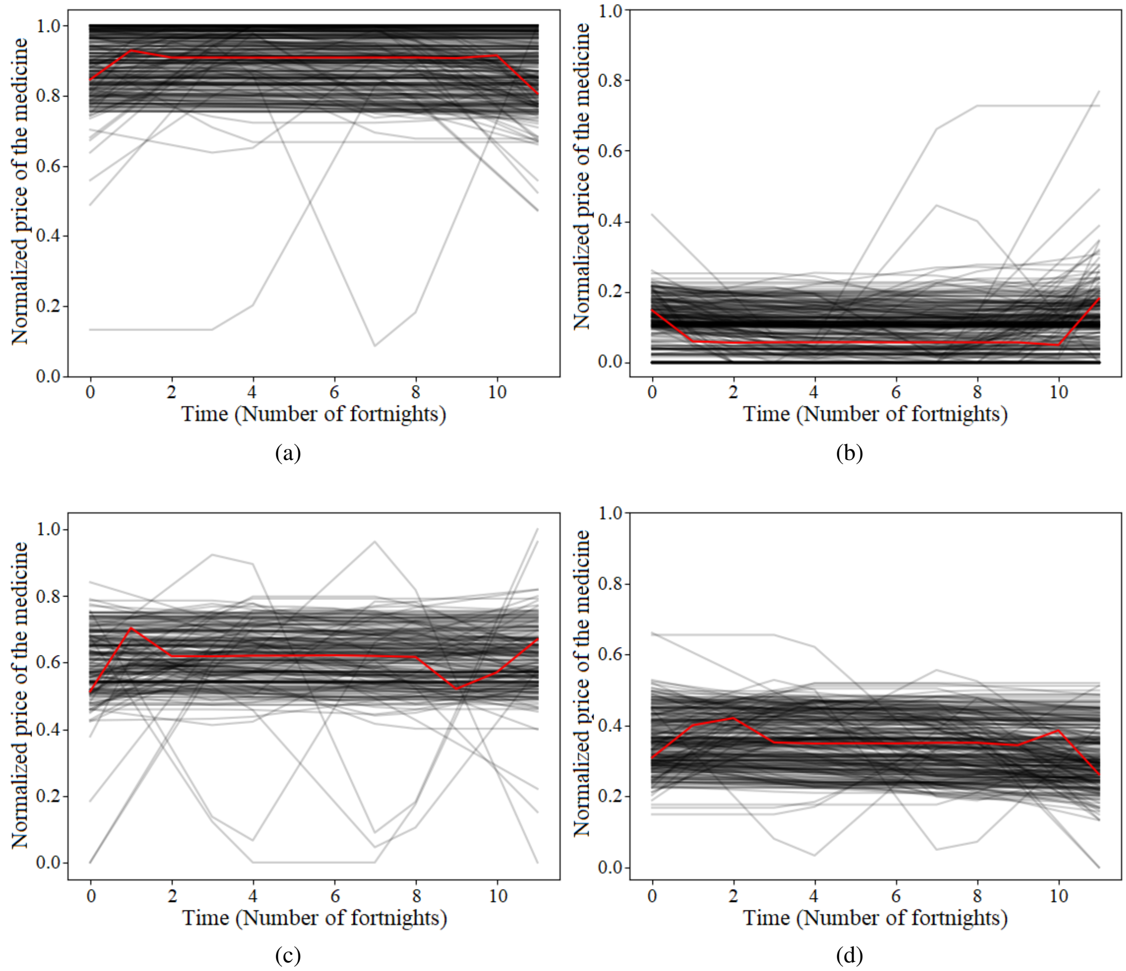


Figure 2: Illustration of the behavior of subsequences in each of the four clusters identified through subsequence time series clustering on pricing data. Each gray line represents a subsequence time series, and the red line represents the cluster center.

On the other hand, cluster 2 (subfigure (b)) represents the opposite pricing pattern compared to cluster 1, where the prices are at their lowest point. The reason for this could be low demand in the market, causing medicines to approach their expiry date and companies to reduce their price significantly to avoid financial loss. We observed that the pricing patterns of companies in clusters 1 and 2 are stable most of the time. In cluster 1, the pricing starts with a raise, followed by stability, and then a decrease at the end. The cluster center, which is in red, also approves this observation. In cluster 2, we observed a reverse pattern compared to cluster 1.

The other two clusters (subfigures (c) and (d)) exhibit competitive pricing patterns, with the cluster centers being around the middle. The cluster 3 (subfigure (c)) depicts competition in a market where demand is slightly higher than the supply of the medicine, resulting in higher prices than the average price. The cluster center suggests pricing fluctuations to gain a larger market share. This can also be an indicator of transferring from lower prices to higher prices in the market as the cluster center starts with a raise, followed by stability, and then an increase at the end.

Finally, cluster 4 (subfigure (d)) shows the same behavior as cluster 3, where the market is competitive. However, the pricing pattern shows a decrease at the end of the period. This could be an indicator of transferring to a market status similar to cluster 2.

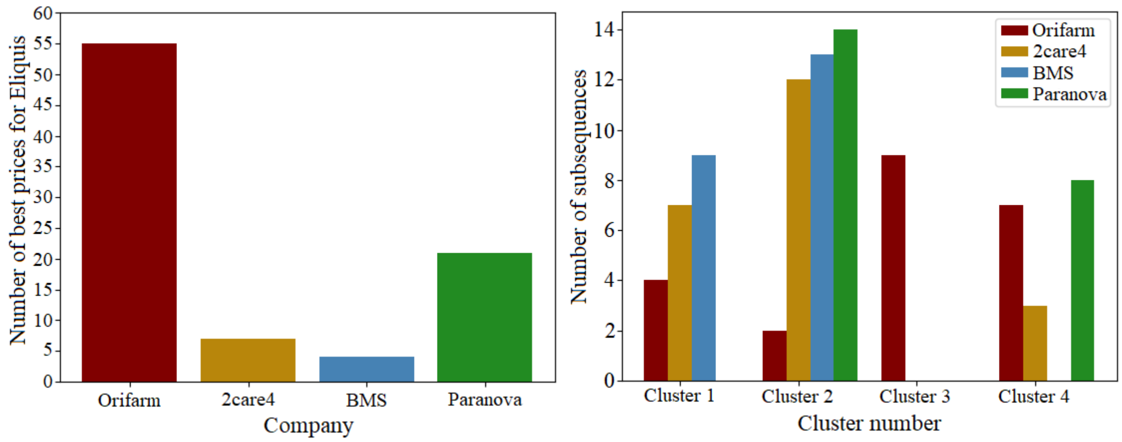
In summary, our analysis demonstrates that there are four major pricing patterns in the Danish pharmaceutical market, which can be useful for pharmaceutical companies, policymakers, and investors to understand and respond to market trends effectively. These pricing patterns resulted from various market conditions, including medicine shortages, demand fluctuations, expiration dates, and competition. The clustering results and subsequent time series analysis provide valuable insights into the dynamics of the market. They can aid in developing a more comprehensive agent-based model of price competition in the Danish pharmaceutical market. By incorporating these findings into the model, we can simulate different market scenarios and better understand the effects of various factors on pricing strategies and market outcomes.

4 ANALYSIS OF COMPETITIVE PRICING BEHAVIOR FOR ELIQUIS AND XARELTO IN THE DANISH PHARMACEUTICAL MARKET

This section aims to provide a detailed analysis of the pricing behavior observed in the Danish pharmaceutical market based on the subsequence time series clustering analysis performed in Section 3. Such an analysis can offer valuable insights into the pricing dynamics of the market and facilitate the development of an agent-based model that represents the players' behavior as a decision support system. We, specifically, focus on two medicines with high competition, involving more than three players, to investigate the most effective pricing behavior and whether a dominant pricing strategy exists. The purpose of this analysis is to enhance the understanding of the market and contribute to the development of an agent-based model that is reliable and effective.

In our analysis of historical pricing data, we selected Eliquis and Xarelto as our medicines of interest, which are popular medicines according to the available Danish Medicine Agency demand dataset. Eliquis and Xarelto are anticoagulant medications used to prevent blood clots in certain medical conditions such as atrial fibrillation and deep vein thrombosis. Specifically, for Eliquis we focused on the pricing behavior of packages containing 60 tablets with a strength of 2.5 mg and packages containing 42 tablets with a strength of 15 mg for Xarelto. Four companies compete over selling both medicines in the Danish market. BMS, Orifarm, 2care4, and Paronova compete to sell Eliquis, and Bayer A/S, Abacus, 2care4, and Orifarm are competing to sell Xarelto.

We analyzed the pricing behavior of companies involved in the Eliquis market using historical price data. All four companies were involved in the competition over selling this medicine in the last 87 fortnights ending the first week of April 2023, so we only considered those fortnights to analyze the competition. To compare the pricing behavior of the companies, we split their price histories into subsequences of 4 fortnights and we applied the K-means model from Section 3 to observe clusters representing their behaviors. Orifarm had the lowest price for 55 fortnights, followed by Paranova for 21 fortnights, 2care4 for 7 fortnights, and BMS for 4 fortnights. The number of best prices for each company is shown in a bar chart in Figure 3 (a). The clustering results in Figure 3 (b) show that Orifarm price history subsequences, which had the lowest prices for the most fortnights, were mostly clustered in clusters 3 and 4, indicating their competitive behavior. As we can see, Orifarm is the only company that used cluster 3 pricing patterns gaining them a higher market share on 55 fortnights. Paranova, the second competitive company, had pricing behavior mostly clustered in clusters 2 and 4, where both clusters relate to lower pricing patterns. On the other hand, 2care4 pricing subsequences, which only had the lowest price for 7 fortnights, were mostly clustered in clusters 1 and 2, indicating a lack of competitive behavior. 2care4 pricing patterns over half of the whole period clustered in cluster 2, which indicates they were struggling to sell the medicine during the period when the market was competitive, presenting as cluster 3 and 4. BMS, as a pharmaceutical manufacturer, had the best price for only 4 fortnights, and its pricing subsequences were all clustered in clusters 1 and 2, which also indicates non-competitive behavior.



(a) (b)

Figure 3: Comparison of pricing behavior for Eliquis among companies. (a) Bar chart displaying the number of best prices for fortnights for each company. (b) Cluster plot showing the behavior of companies based on K-means clustering of their price history subsequences.

In the last 49 fortnights ending in the first week of April 2023, there were four companies involved in the sale of Xarelto. Orifarm and Abacus, which are both parallel traders, had the lowest prices for the most number of fortnights, with 19 each. 2care4, another parallel trader, had the lowest price for 10 fortnights. Bayer A/S, a pharmaceutical manufacturer, had the lowest price for only one fortnight. The bar chart in Figure 4 (a) shows the number of times each company had the best price. To analyze the pricing behavior of each company for Xarelto, the same K-means model used for Eliquis was employed, and the subsequence results were studied. In Figure 4 (b), it can be seen that Orifarm's pricing behavior clustered mostly in cluster 2 and cluster 4, where the pricing patterns are low. Abacus had the same number of best prices as Orifarm, but their behavior mostly clustered in cluster 4, indicating that they were aiming to sell in competitive periods with low prices. 2care4 had the best prices in 10 fortnights, and their subsequence clustering analysis suggests that they mostly tried to keep their prices higher than average compete prices by having their patterns mostly clustered in clusters 3 and 1. Finally, Bayer A/S replicated a behavior similar to BMS in the case of Eliquis, and all their subsequences clustered in clusters 1 and 2.

In addition to the previously noted observations, this analysis provides valuable insights into the Danish pharmaceutical market. Our study reveals that pharmaceutical manufacturers such as BMS and Bayer A/S have adopted a pricing strategy that is closely tied to demand. When demand is high, these companies offer higher prices for their products, and when demand is low, they offer lower prices. This finding can help to refine our agent-based model and enable us to predict pricing behavior in the market accurately.

Considering parallel traders involved in the market, our analysis highlights the different types of pricing strategies. For example, 2care4 company for Eliquis and Xarelto had the lowest number of best prices among all parallel traders, and at the same time, their pricing behavior clustering analysis fit in different clusters compared to other parallel traders like Orifarm and Paranova. This observation not only provides valuable feedback to 2care4's pricing team to enhance their competitive performance but also emphasizes the significant variations in pricing strategies among market participants. This finding underscores the importance of incorporating rationality and competitiveness parameters in the development of an agent-based model of the market. By integrating such parameters, we can accurately model the general behavior of players and create a more realistic representation. Additionally, a comprehensive analysis of historical pharmaceutical pricing serves as a valuable resource in constructing a data-driven agent-based model of the market, further enhancing its accuracy and reliability.

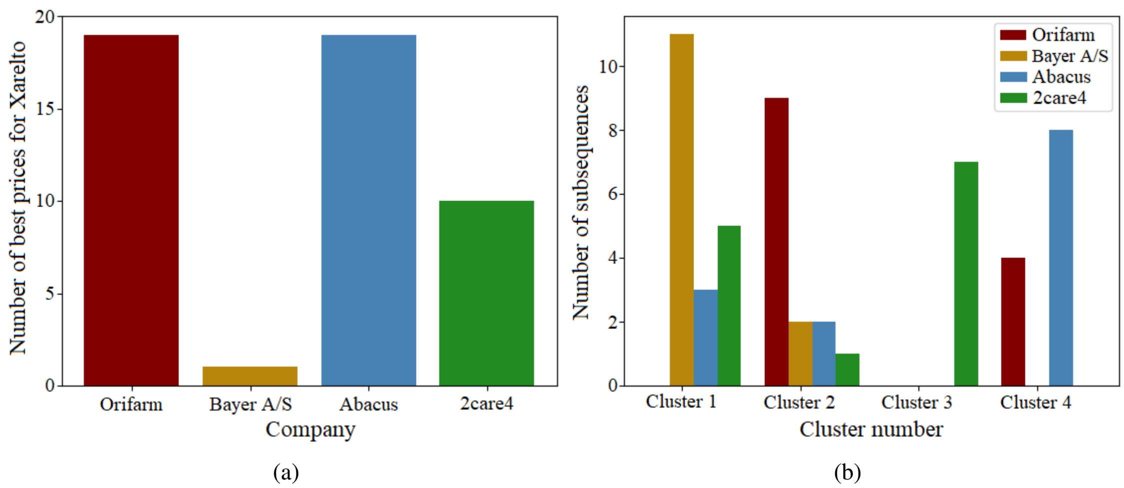


Figure 4: Comparison of pricing behavior for Xarelto among companies. (a) Bar chart displaying the number of best prices for fortnights for each company. (b) Cluster plot showing the behavior of companies based on K-means clustering of their price history subsequences.

The clusters identified in our data analysis offer valuable insights for designing pricing phases within the agent-based model of the market. When examining Eliquis and Xarelto, we observed distinct pricing behaviors among the clusters. Cluster 1 primarily represented passive pricing behavior from participants and manufacturers, with a majority of manufacturers' pricing behaviors falling within this cluster. Focusing on Eliquis, we discovered that Orifarm consistently offered the best price in over 50% of all fortnights. Their pricing behavior was primarily clustered in clusters 3 and 4, indicating competitive adjustments in pricing. Turning to Xarelto, we observed a close competition between Orifarm and Abacus. Figure 4 (b) illustrates that their competitive behaviors clustered in clusters 4 and 2. Cluster 4 demonstrated competition at low prices, while cluster 2 potentially represented competition at the lowest possible prices, highlighting the competitive nature of the Xarelto market.

In summary, our analysis provides valuable insights that can significantly enhance the design of the agent-based model of the market. However, it is important to note that for a comprehensive understanding, this data analysis procedure should be applied to the entire available dataset, encompassing a broader range of medicines beyond the current selection of 36 medicines. By incorporating these findings into the agent-based model, we can create a more accurate and realistic representation of the market, capturing the nuances of competitive pricing behaviors. This exploration lays the foundation for further research and provides a solid framework for understanding market dynamics and optimizing pricing strategies.

5 SUMMARY AND OUTLOOK

This paper aims to investigate the pricing patterns of parallel-traded medicines in Denmark using available historical price data. First, we identified the most popular parallel-traded medicines in the Danish market through a public dataset provided by the Danish Medicine Agency. Then, we employed subsequence time series analysis to examine the pricing pattern of the market, presenting a thorough analysis towards finding price patterns in the market. Furthermore, we applied the K-means clustering model to analyze the behavior of players involved in the competition for selling two popular medicines, Eliquis and Xarelto.

The data analysis pipeline presented in this study enables the identification of price patterns for clustering and discovering agent groups. Furthermore, it provides a foundation for expanding the current agent-based model of the European pharmaceutical parallel trade market by analyzing players' pricing behavior and patterns, facilitating the modeling of agent behaviors using historical data, and contributing to the advancement of research in this field. Furthermore, our analysis revealed that companies can adopt different pricing strategies for each medicine depending on market conditions and competition intensity. Consequently, to ensure the development of a comprehensive and accurate agent-based model of a pharmaceutical market, it is essential to conduct rigorous data analysis, as demonstrated in this paper. By incorporating the insights from a comprehensive analysis, we can design agents that accurately reflect the behavior, strategies, and patterns of participants and facilitate simulations that better capture the market dynamics. Moreover, the model can provide policymakers and industry players with a better understanding of the effects of different policies and market conditions on the behavior of players in the market. Finally, it can help stakeholders identify potential risks and opportunities associated with different policy options, enabling them to make informed decisions leading to a more stable and efficient pharmaceutical market.

Several avenues for research could build on the findings of this study. Firstly, future studies could focus on expanding the analysis to include other European countries to gain a broader understanding of the pharmaceutical parallel trade market in the region. Additionally, further research could investigate the impact of external factors, such as regulatory changes and market disruptions, on the behavior of players in the market. Furthermore, the development of an agent-based model for the European pharmaceutical parallel trade market could be a valuable tool for stakeholders to explore the potential outcomes of different policy scenarios. Finally, future research could explore the use of other clustering algorithms to analyze the behavior of players in the market and compare the results with those obtained from the K-means clustering model. Overall, the findings of this study provide a foundation for further research into the pricing patterns and behavior of players in the pharmaceutical parallel trade market.

REFERENCES

Agarwal, N., and P. Kwan. 2017. "Pricing Mergers & Acquisitions Using Agent-based Modeling". *Economics, Management and Financial Markets* 12(1):55.

Aghabozorgi, S., A. S. Shirkhorshidi, and T. Y. Wah. 2015. "Time-series Clustering–a Decade Review". *Information Systems* 53:16–38.

Aghabozorgi, S., and Y. W. Teh. 2014. "Stock Market Co-movement Assessment Using a Three-phase Clustering Method". *Expert Systems with Applications* 41(4):1301–1314.

Arvitrida, N. I., A. Candra, N. Siswanto, and L. Yuwana. 2019. "The Effect of Pricing Strategies on Retailers: An Agent-Based Modeling Approach". In *International Conference on Operations and Supply Chain Management*, 1–8.

Athanasiadis, I. N., A. K. Mentes, P. A. Mitkas, and Y. A. Mylopoulos. 2005. "A Hybrid Agent-based Model for Estimating Residential Water Demand". *Simulation* 81(3):175–187.

Axtell, R. L., and J. D. Farmer. 2022. "Agent-based Modeling in Economics and Finance: Past, Present, and Future". *Journal of Economic Literature (FORTHCOMING)*.

Bennato, A. R., and T. Valletti. 2014. "Pharmaceutical Innovation and Parallel Trade". *International Journal of Industrial Organization* 33:83–92.

Blázquez-García, A., A. Conde, U. Mori, and J. A. Lozano. 2021. "A Review on Outlier/Anomaly Detection in Time Series Data". *ACM Computing Surveys (CSUR)* 54(3):1–33.

Bookstaber, R. M. 2012. "Using Agent-based Models for Analyzing Threats to Financial Stability". *Working Paper no. 0003 (21 December). U.S. Department Treasury. Office of Financial Research*.

Cornacchione, E. B., L. Reginato, J. O. Imoniana, and M. Souza. 2023. "Dynamic Pricing Models and Negotiating Agents: Developments in Management Accounting". *Administrative Sciences* 13(2):57.

Du, H., and T. Xiao. 2019. "Pricing Strategies for Competing Adaptive Retailers Facing Complex Consumer Behavior: Agent-based Model". *International Journal of Information Technology & Decision Making* 18(06):1909–1939.

Dubois, P., and M. Sæthre. 2020. "On the Effect of Parallel Trade on Manufacturers' and Retailers' Profits in the Pharmaceutical Sector". *Econometrica* 88(6):2503–2545.

Feng, L., B. Li, B. Podobnik, T. Preis, and H. E. Stanley. 2012. "Linking Agent-based Models and Stochastic Models of Financial Markets". *Proceedings of the National Academy of Sciences* 109(22):8388–8393.

Fong, S. 2012. "Using Hierarchical Time Series Clustering Algorithm and Wavelet Classifier for Biometric Voice Classification". *Journal of Biomedicine and Biotechnology* 2012:1–12.

Gnecco, G., B. Tuncay, and F. Pammolli. 2018. "A Comparison of Game-theoretic Models for Parallel Trade". *International Game Theory Review* 20(03):1850003.

Guo, S., B. Hu, and H. Zhong. 2013. "Impact of Parallel Trade on Pharmaceutical Firm's Profits: Rise or Fall?". *The European Journal of Health Economics* 14:345–355.

Hendrawati, T., A. Wigena, I. Sumertajaya, and B. Sartono. 2021. "Clustering of Commodity Inflation Pattern based on Estimated ARIMA Model". In *Journal of Physics: Conference Series*, Volume 1863, 012058. IOP Publishing.

Huang, X., Z. Li, J. Lu, S. Wang, H. Wei, and B. Chen. 2020. "Time-series Clustering for Home Dwell Time During COVID-19: What Can We Learn From It?". *ISPRS International Journal of Geo-Information* 9(11):675.

Jamali, R., and S. Lazarova-Molnar. 2022a. "The Relationship Between Agent-based Simulation and Game Theory in the Case of Parallel Trade". In *2022 IEEE International Conference on Agents (ICA)*, 36–41. IEEE.

Jamali, R., and S. Lazarova-Molnar. 2022b. "Towards Agent-based Simulation of the Parallel Trading Market of Pharmaceuticals". In *2022 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, 173–181. IEEE.

Kanavos, P., and J. Costa-Font. 2005. "Pharmaceutical Parallel Trade in Europe: Stakeholder and Competition Effects". *Economic Policy* 20(44):758–798.

Kanavos, P., J. Costa-I-Font, S. Merkur, and M. Gemmill. 2004. "The Economic Impact of Pharmaceutical Parallel Trade in European Union Member States: A Stakeholder Analysis". *London: LSE Health and Social Care Special Research Paper*.

Li, J., H. Izakian, W. Pedrycz, and I. Jamal. 2021. "Clustering-based Anomaly Detection in Multivariate Time Series Data". *Applied Soft Computing* 100:106919.

Lonardi, J., and P. Patel. 2002. "Finding Motifs in Time Series". In *Proc. of the 2nd Workshop on Temporal Data Mining*, 53–68.

Müller-Langer, F. et al. 2007. "A Game Theoretic Analysis of Parallel Trade and the Pricing of Pharmaceutical Products". *German working paper in Law and Economics* 2007.

Nair, B. B., P. S. Kumar, N. Sakthivel, and U. Vipin. 2017. "Clustering Stock Price Time Series Data to Generate Stock Trading Recommendations: An Empirical Study". *Expert Systems with Applications* 70:20–36.

Ng, T. S., and Y. C. Kang. 2012. "The Inquiry of Labor Market Dynamics in Knowledge Economies: An Agent-based Approach". *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 43(2):416–427.

Pecorino, P. 2002. "Should the US Allow Prescription Drug Reimports from Canada?". *Journal of Health Economics* 21(4):699–708.

Samanidou, E., E. Zschischang, D. Stauffer, and T. Lux. 2007. "Agent-based Models of Financial Markets". *Reports on Progress in Physics* 70(3):409.

Sun, S., S. Wang, Y. Wei, and G. Zhang. 2018. "A Clustering-based Nonlinear Ensemble Approach for Exchange Rates Forecasting". *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 50(6):2284–2292.

Tavenard, R., J. Faouzi, G. Vandewiele, F. Divo, G. Androz, C. Holtz, M. Payne, R. Yurchak, M. Rußwurm, K. Kolar, and E. Woods. 2020. "Tslearn, A Machine Learning Toolkit for Time Series Data". *Journal of Machine Learning Research* 21(118):1–6.

Van Der Veen, R. A., A. Abbasy, and R. A. Hakvoort. 2012. "Agent-based Analysis of the Impact of the Imbalance Pricing Mechanism on Market Behavior in Electricity Balancing Markets". *Energy Economics* 34(4):874–881.

Vandoros, S., and P. Kanavos. 2014. "Parallel Trade and Pharmaceutical Prices: A Game-theoretic Approach and Empirical Evidence from the European Union". *The World Economy* 37(6):856–880.

Yousefi, S., M. M. Parsa, and M. V. Johari. 2011. "Agent-based Modeling of Day-ahead Real Time Pricing in a Pool-based Electricity Market". *Iranian Journal of Electrical Electronic Engineering* 7(3).

Yu, N., M. De Jong, S. Storm, and J. Mi. 2012. "Transport Infrastructure, Spatial Clusters and Regional Economic Growth in China". *Transport Reviews* 32(1):3–28.

AUTHOR BIOGRAPHIES

RUHOLLAH JAMALI is a PhD student in the SDU Software Engineering Section of the Faculty of Engineering at University of Southern Denmark. His research interests concern Modeling, Data Analysis, and Simulation, especially as applied in developing decision support systems. His email address is ruja@mmmi.sdu.dk

SANJA LAZAROVA-MOLNAR is a Professor at the Institute of Applied Informatics and Formal Description Methods, Karlsruhe Institute of Technology. She is also a Professor at the University of Southern Denmark, where she leads the research group Modelling, Simulation and Data Analytics. She is a Senior Member of The Institute of Electrical and Electronics Engineers (IEEE), and currently serving as Director-at-Large on the Board of Directors of The Society for Modeling & Simulation International (SCS). Furthermore, she is Chair of IEEE Denmark and Vice-Chair of IEEE Denmark Women in Engineering Affinity Group. Her email address is sanja.lazarova-molnar@kit.edu.