Universität Hohenheim
Fakultät Wirtschafts- und Sozialwissenschaften

Consumer Prices: Effects of Learning Algorithms and Pandemic-Related Policy Measures

Dissertation zur Erlangung des Doktorgrades der Wirtschaftswissenschaften (Dr. oec.)

Eingereicht von
Katrin Buchali
aus Stuttgart

Institut für Volkswirtschaftslehre
Lehrstuhl für Mikroökonomie insb. Industrieökonomie

2023

Datum der mündlichen Prüfung: 15. Februar 2023
Dekan: Prof. Dr. Jörg Schiller
Prüfungsvorsitz: Prof. Dr. Sibylle Lehmann-Hasemeyer
Erstgutachter: Prof. Dr. Ulrich Schwalbe
Zweitgutachter: Prof. Dr. Alfonso Sousa-Poza

Acknowledgements

First and foremost, I want to express my gratitude to my supervisor Ulrich Schwalbe for giving me the opportunity to write this thesis and for the continuous support throughout working on it. But also for the freedom to pursue my own research interests and for the pleasant working atmosphere at our chair.

Special thanks go to Alfonso Sousa-Poza and Sibylle Lehmann-Hasemeyer for being my second supervisor and the chairwomen of my examination commission, respectively. Their valuable comments and feedback, also during the doctoral seminar in Hohenheim, were very important and helpful for improving the dissertation.

I am very grateful to my colleagues and co-authors Matthias, Jens and Marcel for the productive conversations and discussions and for the enjoyable atmosphere during our joint research and teaching.

Exceptional gratitude goes to my husband Tim for his unconditional belief in me and for the ongoing support. Thank you!

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

When it comes to product prices, two major topics have dominated the public debate in recent years: One is pricing with the help of artificial intelligence (AI), and the other is the price level, which has risen more than usual with the onset of the COVID-19 pandemic. While at first glance these two issues are independent of each other, they both contribute to a growing concern about higher prices. Higher prices create a loss of consumer surplus and possibly total welfare, which is the reason this topic has become ubiquitous in political discussions.

One subfield or application of AI is machine learning, which is the process of using statistical models or algorithms to help computer systems learn without explicit instructions and improve on their own based on past experiences. Concerns have been raised that by extensively collecting and analyzing a large amount of data, these learning algorithms may cause potential harm to consumers and a decline in social welfare. There is no doubt that an increasing amount of available data, combined with AI-related improvements and innovations, may affect firms' behavior in the market. For example, more information and better knowledge of consumers, as well as the use of price-setting algorithms, can facilitate the personalized pricing strategies of firms. Additionally, since firms are able to better observe and predict their competitors' behavior and algorithms are able to respond almost instantaneously to competitors' moves, collusive behavior may be facilitated. These two phenomena – algorithmic price discrimination and tacit collusion by price-setting algorithms – will be discussed in further detail in the present work.

Collusion occurs "when firms use strategies that embody a reward-punishment scheme which rewards a firm for abiding by the supra-competitive outcome and punishes it for departing from it" (Harrington, 2018, p. 336). The potential use of algorithms may facilitate such collusion by making it easier to detect and respond to competitors' deviations. Thereby, algorithms are deliberately used by firms as an instrument to form a cartel or to stabilize an existing one, which would be similar to explicit collusion. Moreover, a hub-and-spoke scenario might emerge when several firms (the "spokes") use the same algorithm provided by a third party (the "hub") that offers algorithmic pricing as a service. And finally, there is the possibility of algorithms learning to collude without being programmed to do so. This is what we refer to as algorithmic tacit collusion. These scenarios have been widely discussed among competition authorities, economic organizations, and professional experts in recent years (Autorité de la concurrence and Bundeskartellamt, 2019; Competition and Markets Authority, 2021; Ezrachi and Stucke, 2016; OECD, 2017; Varian, 2018). They state that AI technologies could enable firms to analyze and monitor the market on the supply side as well as on the demand side on an unprecedented scale. Algorithms are able to detect and scrutinize deviating behavior of competitors as well as to change prices more frequently in order to adapt to a changing environment. In this context, a growing number of researchers argue that pricing algorithms may learn how to collude based on reward and punishment schemes that occur with repeated interactions over a sufficiently long time horizon, resulting in setting prices at a supra-competitive level. Moreover, algorithms may be able to sustain these supra-competitive prices without human intervention, and in contrast to traditional cartels, do not require explicit agreements. Instead, learning algorithms may be able to autonomously induce collusive behavior without having been programmed to do so and without firms' intent. In their simulations, Calvano et al. (2020) and Klein (2021) have demonstrated that algorithms can effectively learn how to implement and sustain collusive strategies over a reasonably long period.

The second phenomenon discussed in this work is algorithmic price discrimination. Price discrimination means that companies charge different prices to different consumers or groups of consumers for the same or a similar product or service. In the context of the increasing use of big data by firms and the sometimes very careless disclosure of personal data by consumers online, which might be explained by the privacy paradox[1] (Norberg et al., 2007), concerns have been raised that learning algorithms might engage in price discrimination as firms have access to huge amounts of data. AI algorithms are able to use this data to create more accurate consumer profiles and to gain a better understanding of consumers' purchasing behavior with regard to their preferences or needs (Woodcock, 2019). Thus, access to large data sets on consumer behavior and the use of AI technology may enable firms to incorporate this information in marketing

[1] The privacy paradox describes the apparent dichotomy between individuals' intentions to disclose personal information and their actual privacy behavior: On the one hand, individuals express concerns about the handling of their personal data and report a desire to protect their data, whereas at the same time, they not only voluntarily disclose these personal data, but also rarely make an effort to protect their data actively (Acquisti and Grossklags, 2005).

or pricing applications resulting in targeted advertising, personalized pricing, or personalized product recommendations. These concerns are further supported by scientific research. In experimental studies, Shiller (2013), Ban and Keskin (2020), and Dubé and Misra (2017) have demonstrated that finer-grained price discrimination is possible with the use of machine learning. Despite the new opportunities to personalize prices that have been gained through digitalization and advances in pricing technology, this has only rarely been observed in practice. This could be due to the expectation of negative consumer reactions, as price discrimination is often seen as an unfair violation of consumer equality (Kahneman et al., 1986).

Even though price discrimination and collusive behavior of firms both require intensive use of data and advanced algorithms, they are unlikely to occur in the same type of market. This is also supported by the studies of Colombo (2010) and Helfrich and Herweg (2016) who have found that price discrimination is not a facilitating factor for collusive behavior. Moreover, the effects of the two phenomena on consumer surplus differ. While supra-competitive prices always reduce consumer surplus, the effect of price discrimination is not that clear. When firms engage in discriminatory pricing strategies, prices are closer to consumer's willingness to pay. On the one hand, consumers with a high willingness to pay will have a lower surplus as they will be charged higher prices. On the other hand, there is the market expansion effect, which results from the fact that consumers who would not have received an offer under a uniform price are now able to buy the product or service at a lower price. If the number of these consumers is sufficiently large, the increase in demand has a positive effect on overall consumer surplus. However, the global impact of price discrimination on consumer welfare depends on the relative magnitude of these two effects. As both collusive behavior and price discrimination may affect consumers, these topics are of particular interest for competition authorities as well as for regulatory bodies.

Another event that created an awareness among financial media, academics, and bankers regarding a significant increase in prices was the COVID-19 pandemic. The outbreak of the coronavirus disease at the beginning of 2020 led to a world health crisis of a type and magnitude never before experienced. In addition to the dramatic health and societal impacts, there were also economic challenges: In Germany, there was a rise in prices not seen since the beginning of the 1990s. Increasing global interdependencies and the provision of intermediate products according to the just-in-time approach have led to suppliers being very sensitive to disruptions in the value chain (OECD, 2020). The pandemic even spurred discussions emphasizing the risks and instability that is associated with global value chains leading to an international fragmentation of production (OECD, 2021). Disruption in these value chains has been caused by the closure of national borders and lockdown measures that were imposed by several governments to contain the spread of the virus. Another consequence of these measures was the limited availability of international workforces, which contributed to further aggravation of the situation (Nicola et al., 2020). Consumers on the demand side of the market were also affected by the measures. Social distancing and several lockdowns led to a change in the spending patterns of consumers (Andersen et al., 2020; Baker et al., 2020; Carvalho et al., 2020; Chen et al., 2021; Chronopoulos et al., 2020; Landais et al., 2020). These are just a few examples of the different factors that may influence demand and supply, and as a consequence, prices. In this context, it is important to know to what extent government measures, which actually pursue other goals, can have an effect on price developments. In the present work, we highlight the role of the stringency of government measures, and thus reduced mobility, as a driver of consumer prices. Since many production processes still rely on the physical presence of workers, labor-intensive products seem to have been particularly affected by government measures that restricted mobility. Thus, the increasing importance of AI, which could be used to (partially) automate production processes, is also evident in this area. However, it is obvious that humans cannot be completely replaced.

This dissertation contributes to the debate by extending the existing literature on algorithmic pricing and collusion and to enhance the general understanding of how government measures enforced during the COVID-19 pandemic contributed to (short-term) price developments. More specifically, the main questions under consideration are:

- Does the risk of collusive pricing by learning algorithms persist in real-world scenarios?
- How does a self-learning pricing algorithm perform when faced with inequity-averse consumers?
- To what extent does the stringency of pandemic-related government measures influence consumer price development?

The dissertation contains three papers that address the aforementioned research questions. In Chapter 2, the concern is addressed that tacit collusion might occur if firms employ learning algorithms, as several simulation studies have demonstrated that algorithms using reinforcement learning – a type of machine learning in which agents learn from interacting autonomously with their environment – are able to coordinate their pricing behavior and, as a result, achieve a collusive outcome without having been programmed for it. These studies, however, use very restrictive assumptions about the involved firms, the employed algorithms, and the environment in which they interact. Therefore, we are skeptical that such results can be transferred to more realistic settings. More specifically, in the first paper titled "Algorithmic Collusion: Fact or Fiction?", several conceptual challenges as well as challenges in the real-world application of algorithms are discussed, and we show by our own simulations that resulting market prices strongly depend on the type of algorithm or heuristic that is used by the firms to set prices. We conclude that the strategy combination of all firms employing a Q-learning algorithm, which is a crucial assumption in simulation studies showing collusive behavior of algorithms, is certainly not a Nash equilibrium.

In Chapter 3, the second paper titled "Price Discrimination with Inequity-Averse Consumers: A Reinforcement Learning Approach" introduces inequity aversion of consumers. This means that consumers respond with a lower acceptance probability of a firm's price bid if they feel they are being treated unfairly. We conducted experiments making use of a reinforcement learning algorithm and answer the question of whether it is possible for this self-learning algorithm to learn to engage in price discrimination on the basis of fairness to avoid upsetting customers but still maximize expected revenues by charging personalized prices. Compared to a scenario where inequity aversion is not considered, an improvement in fairness is seen in the situation where inequity-averse consumers are introduced, while at the same time, the algorithm is able to maintain the goal of maximizing revenue. We conclude from our simulations that consumers' sense of fairness, which has prevented firms from engaging in price discrimination, can be incorporated into firms' pricing decisions with the help of learning algorithms, making differential pricing strategies more feasible.

The discussion surrounding the above-average price levels in many countries during the COVID-19 pandemic is extended in Chapter 4 with the third paper titled "The Effects of Movement Restrictions on Consumer Prices During the COVID-19 Pandemic." The rapid increase in the number of COVID-19 infections prompted governments in affected countries to impose measures designed to contain the spread of the coronavirus, including border closures that severely restricted mobility between countries, so-called stay-at-home restrictions, and workplace closures. These restrictions were expected to affect both supply and demand. On the one hand, workers and goods could cross national borders only under more restrictive conditions, resulting in an abrupt worker shortage and decreased supplies of certain goods. On the other hand, consumers adjusted their spending behavior in response to the pandemic due to shifts in preferences, expected income or health risks, or higher economic uncertainty. In this chapter, we present empirical evidence of the impact of government-imposed restrictions and, as a consequence of their enforcement, reduced mobility on consumer prices during the COVID-19 pandemic. We show that the stringency of government measures has a positive and significant impact on the overall consumer price index as well as on the sub-index of the food category, which means that more stringent measures induce higher consumer prices in these categories. Regressions with actual mobility data instead of the stringency of government measures support these results.

Finally, Chapter 5 summarizes the key findings of the thesis and draws conclusions with respect to the research questions posed. In addition, the concluding remarks are briefly put into perspective from a legal and political point of view.

References

Acquisti, A. and J. Grossklags (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy*, 3.1, pp. 26–33.

Andersen, A. L., E. T. Hansen, N. Johannesen, and A. Sheridan (2020). Consumer responses to the COVID-19 crisis: Evidence from bank account transaction data. *Covid Economics*, 7, pp. 88–114.

Autorité de la concurrence and Bundeskartellamt (2019). Algorithms and competition. Report. URL: `https://www.autoritedelaconcurrence.fr/sites/default/files/algorithms-and-competition.pdf` (visited on 07/01/2022).

Baker, S. R., R. A. Farrokhnia, S. Meyer, M. Pagel, and C. Yannelis (2020). How does household spending respond to an epidemic? Consumption during the 2020 COVID-19 pandemic. *The Review of Asset Pricing Studies*, 10.4, pp. 834–862.

Ban, G.-Y. and N. B. Keskin (2020). Personalized dynamic pricing with machine learning: High dimensional features and heterogeneous elasticity. *Management Science*, 67.9, pp. 5549–5568.

Calvano, E., G. Calzolari, V. Denicolò, and S. Pastorello (2020). Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, 110.10, pp. 3267–3329.

Carvalho, B., S. Peralta, and J. Pereira dos Santos (2020). What and how did people buy during the Great Lockdown? Evidence from electronic payments. *Covid Economics*, 28, pp. 119–158.

Chen, H., W. Qian, and Q. Wen (2021). The impact of the COVID-19 pandemic on consumption: Learning from high-frequency transaction data. AEA Papers and Proceedings. DOI: `10.2139/ssrn.3568574`.

Chronopoulos, D. K., M. Lukas, and J. O. S. Wilson (2020). Consumer spending responses to the COVID-19 pandemic: An assessment of Great Britain. *Covid Economics*, 34, pp. 145–186.

Colombo, S. (2010). Product differentiation, price discrimination and collusion. *Research in Economics*, 64.1, pp. 18–27.

Competition and Markets Authority (2021). Algorithms: How they can reduce competition and harm consumers. URL: `https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers` (visited on 06/29/2022).

Dubé, J.-P. and S. Misra (2017). Scalable price targeting. Working Paper No. 23775. National Bureau of Economic Research. DOI: `10.3386/w23775`.

Ezrachi, A. and M. E. Stucke (2016). *Virtual Competition: the promise and perils of the algorithm-driven economy*. 2nd ed. Cambridge, Massachusetts: Harvard University Press.

Harrington, J. E. (2018). Developing competition law for collusion by autonomous artificial agents. *Journal of Competition Law & Economics*, 14.3, pp. 331–363.

Helfrich, M. and F. Herweg (2016). Fighting collusion by permitting price discrimination. *Economics Letters*, 145, pp. 148–151.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1986). Fairness and the assumptions of economics. *The Journal of Business*, 59.4, pp. 285–300.

Klein, T. (2021). Autonomous algorithmic collusion: Q-learning under sequential pricing. *RAND Journal of Economics*, 52.3, pp. 538–599.

Landais, C., D. Bounie, Y. Camara, E. Fize, J. Galbraith, C. Lavest, T. Pazem, and B. Savatier (2020). Consumption dynamics in the COVID crisis: Real time insights from French transaction and bank data. CEPR Discussion Paper DP15474. URL: `https://cepr.org/publications/dp15474`.

Nicola, M., Z. Alsafi, C. Sohrabi, A. Kerwan, A. Al-Jabir, C. Iosifidis, M. Agha, and R. Agha (2020). The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery*, 78, pp. 185–193.

Norberg, P. A., D. R. Horne, and D. A. Horne (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, 41.1, pp. 100–126.

OECD (2017). Algorithms and collusion: Competition policy in the Digital Age. URL: `https://www.oecd.org/competition/algorithms-collusion-competition-policy-in-the-digital-age.htm` (visited on 07/01/2022).

— (2020). Evaluating the impact of COVID-19 containment measures on activity and spending. OECD Economic Outlook, Volume 2020 Issue 1. OECD Publishing, Paris. DOI: `10.1787/fe40a82a-en`.

— (2021). Global value chains: Efficiency and risks in the context of COVID-19. OECD Publishing, Paris. DOI: `10.1787/67c75fdc-en`.

Shiller, B. R. (2013). *First Degree Price Discrimination Using Big Data*. Waltham, Massachusetts: Brandeis University, Department of Economics and International Business School.

Varian, H. (2018). Articial intelligence, economics, and industrial organization. Working Paper No. 24839. National Bureau of Economic Research. DOI: `10.3386/w24839`.

Woodcock, R. (2019). Personalized pricing as monopolization. *Connecticut Law Review*, 51.2, pp. 311–373.

Chapter 2

Algorithmic Collusion: Fact or Fiction?

This chapter is joint work with Jens Grüb, Matthias Muijs, and Ulrich Schwalbe.

Abstract

Concerns are growing that learning algorithms will harm competition as they are said to be able to coordinate their pricing behavior to achieve a collusive outcome. Simulation studies have shown that learning algorithms are indeed able to autonomously induce collusive behavior without having been programmed for it. These studies, however, use very restrictive assumptions about the involved firms, the employed algorithms, and the environment in which they interact. Therefore, we are skeptical that such results can be transferred to more realistic settings. In our paper, we address conceptual challenges as well as challenges in the real-world application of algorithms, and we show by our own simulations that market prices strongly depend on the type of algorithm or heuristic that is used by the firms to set prices. We can conclude that the strategy combination of all firms employing a Q-learning algorithm, which is a crucial assumption in simulation studies showing collusive behavior of algorithms, is certainly not a Nash equilibrium.

2.1 Introduction

Algorithmic pricing is not a new phenomenon, as it has been common in some industries for many years, such as airlines, hotels, and cruises. With the growth of online trading, the use of such algorithms is increasing. Algorithms offer a number of advantages, such as adjusting prices better and more quickly to supply and demand, gathering and processing increasingly larger amounts of data, and freeing up resources through higher levels of automation. However, fears have been expressed, especially in the last five to six years, that such algorithms may have negative effects on consumers and on competition. For example, consumers could be disadvantaged by algorithms in terms of higher prices.[1]

There are also concerns that learning algorithms can harm competition by facilitating collusive behavior. In markets with high transparency, firms might deliberately use algorithms as an instrument to form a new cartel or to stabilize an existing one, which would be similar to explicit collusion. Moreover, several studies indicate that learning algorithms are able to autonomously induce collusive behavior without having been programmed for it and without such behavior having been intended by the firms, which is known as tacit collusion. It is feared that this will lead to additional and novel competition problems – i.e., that collusive behavior will increase and that it will also take place in markets where collusion would not have occurred without such algorithms. However, the assumptions made in these studies are often based on a shallow understanding of machine learning and rather intuitive considerations and assumptions.

The first papers on competitive concerns caused by self-learning, price-setting algorithms – written mostly by legal scholars – pointed out that algorithmic collusion would arise very quickly, that it would be virtually unavoidable, and that it would also occur in markets with many firms that are usually not prone to collusion. This view has been criticized with the claim that such a collusion would never happen tacitly, even in laboratory experiments, as it is generally rather difficult to achieve arrangements or agreements, particularly without any communication. In the last few years, these extreme positions seem to have somewhat converged. The main reason for this is that some simulation studies have demonstrated that comparatively simple learning algorithms are able to learn collusive behavior without having been programmed to do so. The model structure usually employed is a simple Bertrand model. The collusion that the algorithms achieve is generally not perfect, i.e., they do not replicate a perfect cartel and the

[1] For a detailed analysis on consumer harm caused by pricing algorithms, see MacKay and Weinstein (2022).

learning often takes many periods, but nevertheless these studies have demonstrated that learning algorithms are very well able to learn collusive behavior and set prices that are significantly higher than those that would result from effective competition in the Bertrand model. This demonstrates that algorithmic collusion is indeed possible. However, the simulation studies in which collusive behavior has occurred use very restrictive assumptions about the firms involved, the algorithms employed, and the environment in which they interact. For example, it is assumed that all the firms operating in that market use the same price-setting algorithm. Since these assumptions do not reflect realistic market conditions, we are skeptical that such results can be transferred to economic reality. For this reason, we have run additional simulations under more realistic assumptions – e.g., considering firms which produce differentiated products and allowing for the use of algorithms as well as simple heuristics to set prices.[2] We then examined whether collusive behavior still occurs under these conditions.

The paper is organized as follows. In Section 2.2, we review the literature on the recent debate over whether artificial intelligence algorithms are able to autonomously learn to collude. Section 2.3 describes the basic machine learning methods underlying the algorithms, providing a broader understanding of the context. Additionally, the limitations of the previous approaches are considered. In Section 2.4, conceptual challenges and challenges in the real-world application of price-setting algorithms are described. Our arguments are supported by our own simulations, which we provide in Section 2.5. Section 2.6 concludes.

2.2 Literature Review

The Competition and Markets Authority (2021) in the United Kingdom identified three main competition concerns with regard to the use of algorithmic pricing software. First, algorithms may facilitate explicit collusion by making it easier to detect and respond to competitors' deviations. Second, a so-called hub-and-spoke scenario might emerge if firms use the same third-party pricing software or service. Third, it may be possible that algorithms will learn to collude without being programmed to do so. In this study, we focus on the concern that algorithms autonomously learn to collude, which we refer to as tacit collusion.[3]

Tacit collusion occurs when firms coordinate prices, quantities, or any other variable and achieve supra-competitive profits, without any communication or explicit agreements between them. Consequently, the outcome deviates from the

[2] Differentiated products are also considered by Calvano et al. (2020).

[3] Other terms used for tacit collusion include *tacit coordination* and *conscious parallelism*.

competitive equilibrium and, as a result, leads to a reduction in consumer welfare that is similar to the outcome of explicit collusion or a cartel. The necessary requirements for tacit collusion to occur are firms' abilities to detect and to punish competitors' deviations from the collusive equilibrium. Otherwise, competitors will always have the incentive to deviate from the collusive equilibrium by undercutting rivals' prices in order to serve a larger share of the demand. If this happens, the collusive equilibrium breaks down, and prices return to the competitive level.

Whether algorithms are able to tacitly collude, not only in theoretical models but also in the real world, is a controversial discussion in the economics literature. Several simulations have demonstrated that algorithmic collusion might occur. Contributing to this literature, Waltman and Kaymak (2008) studied the use of Q-learning algorithms in repeated Cournot oligopoly games. Drawing on their computer simulations, the authors showed that firms may learn to collude, although full collusion usually does not occur. They also shed light on the fact that Q-learning may explain the emergence of collusive behavior in settings in which punishment mechanisms and communication between firms are absent. Using the framework of a simple two-firm Bertrand model, Calvano et al. (2020) demonstrated that Q-learning algorithms are able to learn collusive strategies when competing algorithms update their prices simultaneously and rivals' prices are perfectly observable. In this setting, the authors also observed that deviating behavior – i.e., setting a lower price than the collusive price – was sanctioned. Typically, collusive strategies were followed by a finite phase of punishment, with a gradual return to prices that were set prior to the deviation. Klein (2021) introduced a sequential setting and showed that in a sequential pricing duopoly wherein firms offer a homogeneous good, competing Q-learning algorithms learn to converge to collusive equilibria when the set of potential prices the algorithm can choose is limited. When this set of potential prices expands, the algorithm increasingly approaches supra-competitive asymmetric price cycles.

However, Schwalbe (2018) argued that these studies used extremely stylized settings. Moreover, according to the author, the settings in which algorithms achieved collusive outcomes correspond to those in which humans also colluded. Contributing to this point of view, Ittoo and Petit (2017) argued that it is not very likely that algorithms will be able to autonomously collude. The authors studied different types of reinforcement learning technologies, such as Q-learning, and tried to determine whether the use of such algorithms can lead to tacit collusion. They concluded that several significant existing technological challenges, such as specifying an appropriate objective function, observing data on the competitive environment, or dealing with a non-stationary environment, undermine the capabilities of Q-learning algorithms to approach a tacit collusion equilibrium. A further aspect that was noted by Harrington (2018) is that collusive behavior is more likely to occur if competitors use a symmetric price-setting technology, which can be achieved by using the same pricing algorithms from the same provider. Harrington further argued that even in this case, the competing firms would be required to train their algorithms with the same data and use the same hyperparameter values for the algorithm.[4] However, this does not sound like a realistic scenario.

In contrast to these findings, Brown and MacKay (2021) introduced a model of price competition that allows for asymmetric technology among firms, whereby firms are able to change prices after short time intervals. In addition, the model incorporates short-run commitment through the use of algorithms. The authors showed that the use of superior pricing technology leads to an increase in markups of all firms: if all firms adopt algorithms with high pricing frequency, collusive prices are observed. Furthermore, the authors concluded that asymmetries in pricing frequency and commitment allows firms to set supra-competitive prices even in a competitive equilibrium by using algorithms that are simple linear functions of rivals' prices.

As shown above, several studies have indeed demonstrated that algorithms are capable of collusive behavior in very stylized settings. The authors of these papers performed several real-world robustness checks to prove that their results are robust regardless of their theoretical study designs. For example, Calvano et al. (2020) carried out a series of robustness checks in their work: the number of firms was increased from two to three or four firms, and asymmetries between firms (such as different marginal costs) were considered. Furthermore, they examined whether the results change fundamentally if fluctuating demand is assumed or if the firms use different algorithms, while maintaining the same type of learning behavior, i.e., Q-learning. Moreover, robustness checks were carried out for various forms of uncertainty. The authors introduced demand shocks, which were modeled by stochastic entry and exit. They found that demand variability constrains collusion among firms but does not eliminate it. Additionally, random entry and exit of an "outsider" firm, which alters the market structure, as well as changes in substitutability of the product, do not lead to significantly different results. Changes in the hyperparameters of the model – e.g., changing the initial

[4] Further details on hyperparamter values of algorithms are provided in Sections 2.3.1 and 2.3.2.

values the algorithm must update, and enlarging the action set of firms by increasing the number of feasible prices – have only limited impact on the collusive strategies. Overall, even with modified assumptions, the authors concluded that such collusive behavior is essentially robust, apart from minor changes in the level of the collusive price. We are, however, skeptical that these robustness checks are sufficient to account for the complexity of economic reality.

A further robustness check was introduced by Calvano et al. (2021) by analyzing the case of imperfect monitoring where firms compete in quantities and observe the competitors' price level but cannot perfectly infer their outputs because demand is stochastic. Perfect monitoring means that each seller is able to monitor the competitors' prices in real time, and is assumed to be present in online marketplaces, such as Amazon. However, perfect monitoring does not appear to be a necessary condition for collusion, as theory shows that collusion may also occur in markets where the strategies of competing firms are not easy to observe (Green and Porter, 1984; Tirole, 1988). Calvano et al. (2021) again used Q-learning pricing algorithms and assumed that all firms use similar algorithms. The results indicated that if those algorithms were provided with sufficient time to complete the learning process, they colluded even under the assumption of imperfect monitoring. The prices set by the firms still yielded supra-competitive profits, but no perfectly collusive outcome was observed. Moreover, the algorithms required hundreds or thousands of periods to stabilize their behavior and converge to a certain strategy. Consequently, the practical significance of this result is questionable. Hansen et al. (2020) contributed to this strand of literature as well, studying market outcomes in an oligopoly setting wherein two competing firms independently employ upper confidence-bound algorithms that are not able to observe competitors' choices. Simulation results suggest a relationship between the price level and the information value of the underlying pricing experiments: more informative pricing experiments result in correlated price experiments across firms, thereby competitors' pricing becoming correlated unobservables in each firms' pricing algorithm. Thus, the firms' misspecified models overestimated own price sensitivity, resulting in supra-competitive prices.

The first empirical evidence for the appearance of collusive price levels due to algorithmic pricing was presented by Assad et al. (2020), who investigated the impact of algorithmic pricing software adoption on competition in the German retail gasoline market by comparing the retail margins of adopting and non-adopting stations. This software, which became widely available by mid–2017, is able to perform high-frequency analysis of data on competitors' price-setting in order to provide fast, intelligent, and agile pricing decisions in reaction to current market conditions, however, it does not include learning algorithms. The authors were able to show that the pricing software has been used as a tool to achieve higher prices and thus higher profits: Regression results indicated that, in non-monopoly markets, margins of gasoline retailers adopting the algorithmic pricing software increased by 9%. Restricting their attention to duopoly markets, the authors found that adoption of algorithmic pricing led to an average margin increase of 28% when both gas stations adopted algorithmic pricing, while only one station adopting algorithmic pricing did not lead to an increase in margins. The authors also gave an explanation for why pricing algorithms could reach margins above competitive levels by examining the timing of adaption effects and by looking at the average number of price changes in duopoly markets. They argued that the algorithms did not fail to learn to compete effectively but rather actively learned how not to compete, i.e., how to tacitly collude.

The possibility of the occurrence of algorithmic collusion is well known to competition authorities. The legal debate on pricing algorithms and collusion has been waged by academics for several years (e.g., Ezrachi and Stucke (2015, 2016), Gal (2017), and Mehra (2016)), but competition authorities have also started to develop measures to combat algorithmic collusion. To this end, the Federal Trade Commission in the United States has issued a guidance paper on the use of artificial intelligence (AI) in markets, providing guidance on the desirable features of AI tools to avoid unintended consequences (Federal Trade Commission, 2020). Currently, however, the Federal Trade Commission believes that the existing US regulatory framework sufficiently addresses the risks associated with the increasing use of AI systems (Federal Trade Commission, 2021). In April 2021, the European Commission presented a new proposal for an EU legal framework for AI, the "Artificial Intelligence Act" (AI Act)[5]. The draft AI Act is the first attempt at horizontal regulation of AI and focuses on the specific usage of AI systems and the associated risks to people's safety or fundamental rights. It proposes a new definition of AI systems in European law and suggests a new classification for AI systems with different requirements and obligations, following a risk-based approach.

In addition, in a joint study the competition authorities of France and Germany addressed potential competitive risks that are associated with the use of algorithms (Autorité de la concurrence and Bundeskartellamt, 2019). With regard to alignment of pricing algorithms, the competition authorities stated that under current case law, Art. 101 of the Treaty on the Functioning of the European

[5] European Commission, Document 52021PC0206, 21 April 2021, `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206`.

Union (TFEU) does not prohibit conscious parallel behavior. According to the authors, a situation in which "an algorithm merely unilaterally observes, analyses and reacts to the publicly observable behavior of the competitors' algorithms [...] might usually have to be considered as an intelligent adaptation to the market rather than a coordination" (Autorité de la concurrence and Bundeskartellamt, 2019, p. 56). Another legal question that was addressed in the study concerns the issue of assessing a firm's responsibility for collusive algorithmic behavior. The authors showed that proposals range from releasing developers from all liability to considering the behavior of an algorithm as the action of a firm's employees, which is to be held completely accountable. However, the authors did not propose committing to a particular standard. Currently, without the ability to observe algorithmic tacit collusion in real-life, the authors see no need for adapting the existing legal regime or their methodological toolkit.

To sum up, the main challenge or novelty that arises with algorithmic collusion is algorithms' ability to autonomously learn to collude, without being programmed to do so. Furthermore, algorithmic collusion – if detected at all – currently does not constitute a violation of competition law.

2.3 Technical Aspects of Reinforcement Learning

To put the topic of machine learning into a broader context, we will first differentiate between different types of learning. Machine learning is a subfield of AI and refers to the extraction of knowledge from data. We can differentiate between several models of machine learning, all of which use different approaches. What most of these models have in common is that they are based on several parameters and an objective function that indicates the performance of the model. In this way, we can distinguish between hyperparameters and parameters. A hyperparameter is a configuration that is external to the model and whose value cannot be estimated from the data. Since the best value for a hyperparameter on a given problem is unknown, they are often set by using rules of thumb, or by searching for the best value through trial and error. Examples of hyperparameters are the choice of a particular optimization algorithm, the learning rate, the number of hidden layers in a neural network, or the number of iterations in training a neural network. In contrast, parameters are internal to the model. The algorithm tries to learn or estimate from the data the parameter values that lead to the best possible performance of the model for a given dataset. Examples of parameters are the coefficients of linear and logistic regression models or the weights and biases of a neural network.

Every machine learning method needs some kind of data, which in turn determines which kind of algorithm is appropriate. Data may be structured in a table-like format or unstructured, which can be text, images, videos, or audios. Moreover, we distinguish between data that is labeled with a certain tag or data that is unlabeled.

In the following, some types of machine learning are briefly discussed. For a comprehensive overview of machine learning methods used in economic contexts, please refer to Athey and Imbens (2019).

The first type of machine learning that we consider is *Supervised Learning*. Supervised learning requires labeled data, which can be used as input, and the output of these methods is some kind of a prediction. To assess the performance of the algorithm, the predictions are compared to the actual output that is depicted by the labels. The goal of supervised learning algorithms is to minimize the discrepancy between actual output and predicted output. One example of a supervised learning problem is predicting house prices: this requires data about other houses, such as square footage, number of rooms, features, whether a house has a yard or not, and so on, as well as the corresponding labels (i.e., the houses' prices). By using data on those houses, their features, and their prices, a supervised machine learning model can be trained to predict a new house's price based on the examples observed by the model.

In contrast to these algorithms, *Unsupervised Learning* algorithms do not rely on labeled data, and no actual output is available. Instead, the performance of this kind of algorithm is evaluated on the input data. Clustering is an unsupervised learning approach that tries to find groups or clusters in a featured space and interpret the input data. Clustering is commonly used in determining customer segments in marketing data, for example, in order to approach these customer segments in more targeted ways.

Another type of machine learning is *Reinforcement Learning*, which cannot be classified as either supervised or unsupervised learning and is therefore often considered its own category. In this method, the algorithm learns how to map situations to actions in order to maximize a numerical reward signal, which is comparable to the above-mentioned objective function. The algorithm is not explicitly told what actions to choose but instead has to learn which actions yield the highest reward. Reinforcement learning methods use the formal framework of Markov decision processes to define the interaction between the algorithm (called the *agent*) and its environment in terms of states, actions, and rewards.[6]

[6] For a comprehensive introduction to single-agent reinforcement learning, we refer the reader to the textbook by Sutton and Barto (2018).

In the following paragraphs, we introduce some formal aspects of reinforcement learning. The agent interacts with its environment at discrete time steps $t = 0, 1, 2, \ldots$. In each time step, the agent observes state $s_t \in \mathcal{S}$, where $\mathcal{S}$ denotes the set of all states, and selects an action $a_t \in \mathcal{A}$, where $\mathcal{A}$ denotes the set of all actions that are available. After choosing this action, the agent obtains an immediate reward $r_{t+1} \in \mathbb{R}$. The agent can then observe the resulting state s_{t+1} and select a subsequent action a_{t+1}, and so on. The agent's policy π at time step t is a mapping from states to action probabilities and is described as

$$\pi_t(s, a) = P(a_t = a | s_t = s).$$

The agent's objective is to receive as much of a reward as possible in the long run. Therefore, the agent's objective function is given by the discounted return for m time steps:

$$R_t = r_{t+1} + \delta r_{t+2} + \delta^2 r_{t+3} + \cdots = \sum_{m=0}^{\infty} \delta^m r_{t+m+1},$$

where $\delta \in [0, 1)$ is the discount factor. During the learning process, the agent changes its policy as a result of the experience it has gained.

One of the challenges that arises in reinforcement learning is the trade-off between the strategies of *exploration* and *exploitation* when choosing an action. To obtain a high reward, an agent with reinforcement learning exploits what it has learned so far by preferring actions it has already tried that have led to a high reward in the past. However, to discover such actions, it has to try (or explore) actions that have not been chosen thus far. Therefore, the algorithm should employ a dynamic policy of action selection that balances exploitation – i.e., choosing the optimal action as currently perceived – and exploration – i.e., randomly choosing another action to improve future performance.

An algorithm receives feedback on its performance in different ways. One way is to apply a learning method that uses training information to evaluate chosen actions instead of relying on instructions that give the correct actions. To do so, algorithms have to explore which action leads to the highest reward. Purely evaluative feedback depends on the action taken by answering the question of how good the action was in receiving a high reward. In contrast, purely instructive feedback does not depend at all on the action chosen by the algorithm but instead indicates the correct action the algorithm should choose. In supervised learning, instructive feedback is given to the algorithms, whereas in reinforcement learning, evaluative feedback is present. In more complex environments, evaluative and instructive feedback can be combined.

We can further distinguish between associative and non-associative settings. In associative settings, inputs are mapped to outputs, and thus algorithms learn the best output for each input. In non-associative settings, algorithms learn (or discover) one best output.

2.3.1 Multi-Armed Bandit

The first reinforcement learning algorithm we consider is the *Multi-Armed Bandit (MAB)*, which operates in a non-associative setting with evaluative feedback. This algorithm is very simple, since the setting does not involve learning to act in more than one state or situation.[7]

Let us consider an algorithm that chooses repeatedly from k actions (or arms). After each action a_t, a reward r_t is obtained. This reward is chosen from a stationary probability distribution that depends on the selected action, and it is denoted by the action value $Q^*(a_t)$. The expected reward given that an arbitrary action a_t is selected is described as

$$E\{r_t|a_t\} = Q^*(a_t).$$

The objective of the MAB is to maximize the expected total reward, which is also called the action value, over some period of time. However, the action values are not known with certainty. Therefore, the estimated value of an action a at time step t is denoted as $Q_t(a)$. These estimates of action values are stored, so at any time step t there is at least one action whose estimated value is the largest compared to other actions that were chosen. Actions that yield the highest estimated value are called greedy actions. If the algorithm decides to choose one of these greedy actions, we say that the algorithm exploits its current knowledge of the action values. If the algorithm chooses one of the non-greedy actions that has not been selected thus far, the algorithm is exploring, since this selection enables the agent to improve the estimates of non-greedy action values.

There are different ways to estimate the values of actions and to use these estimates for future decisions about which action to take. Sutton and Barto (2018) identify these methods as *action-value methods*. Here, we describe two of them. Given that an action was chosen at multiple different time steps, the authors describe the true value of an action as the mean reward of this action. This value can be calculated by averaging all rewards received in past periods

[7] For a more detailed look at multi-armed bandit algorithms, see Sutton and Barto (2018) or Slivkins (2019).

(up to time step m_a) after choosing this specific action. The estimated value of choosing action a is then

$$Q_t(a) = \frac{r_1 + r_2 + \cdots + r_{m_a}}{m_a}.$$

The law of large numbers entails that the denominator m_a goes to infinity. As a result, the estimated value $Q_t(a)$ converges to the true value $Q^*(a_t)$. Since each estimated value is an average of the relevant sample that includes all rewards received when choosing a specific action, this form of estimating action values is described as the *sample-average method*.

We now consider a single action and compute the estimate of its action value Q_m after this action has been selected $m-1$ times. We do this by calculating the average of all $m-1$ rewards received after this action was selected. This average is given by

$$Q_m = \frac{r_1 + r_2 + \cdots + r_{m-1}}{m-1}.$$

Given Q_m and the mth reward (indicated by r_m), the new average of all m rewards can be calculated by accumulating the sum of all rewards and dividing by their number. We then get

$$Q_{m+1} = Q_m + \frac{1}{m}\left[r_m - Q_m\right], \tag{2.1}$$

which is a common form for update rules and, in general terms, can be displayed as

$$NewEstimate \leftarrow OldEstimate + StepSize\left[Target - OldEstimate\right].$$

The sample-average method is used primarily for stationary problems, meaning that the probability distribution of the reward does not change over time. However, for non-stationary reinforcement learning problems, it makes sense to give more weight to recent rewards than to rewards obtained in past periods. One way to do this is using a constant step-size parameter. We denote this parameter by α. Consequently, the update rule (2.1) for updating the average Q_m of the $m-1$ past rewards is modified and now given as

$$Q_{m+1} = Q_m + \alpha\left[r_m - Q_m\right],$$

with $\alpha \in (0, 1]$ being a constant step-size parameter.

In addition to different ways to estimate the Q-value, the agent has different ways to choose its action. One rule the agent might follow is to select one of the greedy actions, i.e., one of those actions with the greatest estimated value. If several greedy actions are observed at one time step, the agent is assumed to select one of these actions at random. This method to select an action can be formulated as

$$a_t = a_t^* = \arg\max_a Q_t(a).$$

An alternative method is to choose the greedy action most of the time but to randomly choose with probability ε an action from all possible actions, independent of the value of these actions. Methods that are based on this action selection approach are called *ε-greedy methods*. If the number of time steps goes to infinity, every action will be sampled an infinite number of times. Consequently, the estimated values $Q_t(a)$ converge to $Q^*(a_t)$.

In the general case of reinforcement learning, the observation s_{t+1} following the selection of an action depends on the previous state s_t and the action a_t taken by the policy π. In the case of MABs discussed here, the following state, which is observation s_{t+1}, does not depend on the action chosen by the agent. Therefore, MABs are called single stage or stateless.

2.3.2 Q-Learning

A different type of reinforcement learning that allows for multiple states is the independent *Q-learning* (Watkins and Dayan, 1992) algorithm. By interacting with its environment, the algorithm learns to maximize a reward according to the function $Q(s, a)$ that matches the optimal long-run value of choosing any action $a \in \mathcal{A}$ when faced with any given state $s \in \mathcal{S}$. During this interaction, the algorithm uses the above-mentioned dynamic action selection policy, which balances actions exploiting what has been previously learned with those exploring what has not been tried before.

The Q-function can be represented as a $|\mathcal{S}| \times |\mathcal{A}|$ matrix. If this Q-matrix is known, the algorithm can easily choose the optimal action for any given state. However, as this matrix is unknown, the Q-learning algorithm first has to estimate the values of the Q-matrix through an iterative procedure without knowing the underlying model. Starting from an arbitrary initial matrix Q_0, the algorithm chooses action a_t in state s_t, observes the reward r_t and subsequent state s_{t+1}, and updates the corresponding cell of the matrix $Q(s_t, a_t)$ according to the following recursive relationship:

$$Q(s_t, a_t) \leftarrow (1-\alpha)Q(s_t, a_t) + \alpha \left(r_t + \delta \max_a Q(s_{t+1}, a) \right),$$

where the updated value $Q(s_t, a_t)$ is a convex combination of the previous value $Q(s_t, a_t)$ and the reward obtained after performing action a_t in state s_t plus the discounted Q-value that is reached in state s_{t+1}. For all other cells of the matrix, the Q-value does not change. The step-size parameter $\alpha \in (0, 1]$ is the learning rate, and it determines how quickly new information replaces old information. The parameter $\delta \in [0, 1)$ describes the discount factor. Action a denotes the optimal strategy (i.e., the action leading to the highest reward) until this time step.

To balance exploration and exploitation, the Q-learning algorithm adopts a policy of selecting an action with some probability. Using a ε-greedy strategy, the algorithm follows a random explorative action within a given interval $[a_{min}, a_{max}]$ with probability $\varepsilon_t \in [0, 1]$ and an exploitative action with probability $1 - \varepsilon_t$. If several actions yield the same highest Q-value under exploitation, the algorithm selects one of these actions randomly.[8]

$$a_t = \begin{cases} [a_{min}, a_{max}] & \text{with probability } \varepsilon_t \\ \arg\max_a Q(s_t, a) & \text{with probability } 1 - \varepsilon_t \end{cases}$$

Therefore, the probability of exploration in period t is determined as

$$\varepsilon_t = \varepsilon_0 (1 - \theta)^t,$$

where $\varepsilon_0 \in [0, 1]$ is the initial exploration probability and $\theta \in [0, 1]$ is a decay parameter that ensures convergence to a deterministic strategy. If we want the exploration rate to, say, decrease to a value of 0.1% after 100,000 periods, the parameter θ is calculated as follows:

$$\theta = 1 - 0.001^{\frac{1}{100{,}000}}.$$

The resulting exploration probabilities of such an experiment are displayed in Figure 2.1, which shows the convergence of the exploration probability to the value of 0.1% after 100,000 periods.[9]

[8] This approach is based on the textbook of Sutton and Barto (2018) and is also used by Calvano et al. (2020) and Klein (2021).

[9] There are similar approaches to model the exploration probability. For example, Calvano et al. (2020) determine the exploration probability by $\varepsilon_t = e^{-\beta t}$ with $\beta > 0$. Using this approach, exploration decreases faster with greater β.

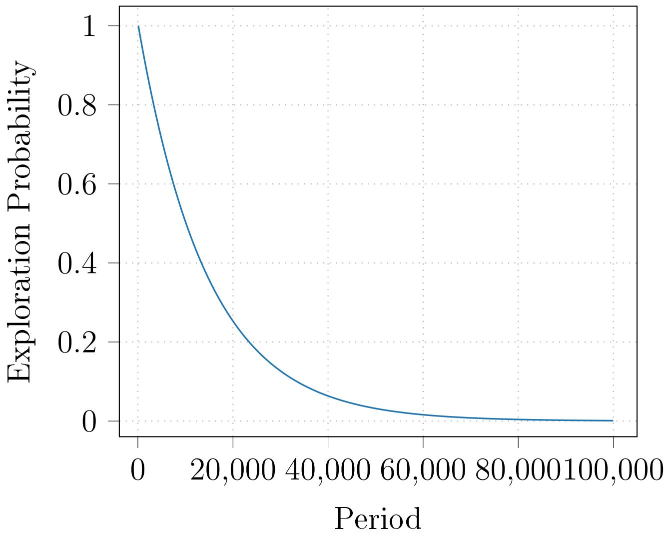

FIGURE 2.1: Convergence of the Exploration Rate

2.3.3 Limitations of Reinforcement Learning

Many constrained optimization problems that are well known in economics can be expressed within the framework of reinforcement learning, and several economic studies have provided simulations utilizing reinforcement learning algorithms to solve such problems. One example is the use of reinforcement learning and its specifications for optimization and control of modern power and energy systems.[10] However, in particular, the widely-used Q-learning has some shortcomings that we address in this section.

One feature of Q-learning is that the size of the Q-matrix is determined by the number of possible actions and the number of possible states, i.e., $|A| \times |S|$, wherein $|A| = k$ and $|S| = k^n$, and n denotes the number of agents. If the Q-learning algorithm is transferred to a more realistic environment, especially one with a higher number of agents, this leads to a vast increase in the size of the Q-matrix. For an example, let us consider the study by Calvano et al. (2020), where the Q-matrix for $n = 2$ agents and a restricted action set with $k = 15$ possible actions has $3,375$ entries. If the number of agents is increased to $n = 3$ or $n = 4$, the number of entries in the Q-matrix increases to around $50,000$ or $750,000$, respectively. In an oligopolistic setting with action sets that are not restricted[11], this implies huge storage capacities and significant processing time, which leads us to our next point: the time factor.

[10] Cao et al. (2020) provide a detailed application overview of single-agent as well as multiple-agent reinforcement learning algorithms in power and energy systems.

[11] In the study of Calvano et al. (2020), the number of actions in compared scenarios is restricted to 15, 50 and 100 possible actions. Klein (2021) compares scenarios with 6, 12, and 24 possible actions that can be chosen by the algorithm.

Depending on the dimensions of the Q-matrix, learning can represent a large or prohibitive investment of time, even if the price-setting intervals are quite short. For example, if prices can be changed every 5–10 minutes[12], the quite simple algorithm employed by Calvano et al. (2019) would take roughly half a year to realize that collusion may be profitable, as their simulation results indicated that the learning agents converge to collusive prices after approximately 70,000 iterations. In 2020, the authors conducted another study, in which the required number of iterations to ensure convergence was even higher at 500,000, and therefore would require an even longer time to learn a profit-maximizing collusive strategy (Calvano et al., 2020). During this time span, conditions such as the market environment might change, rendering obsolete the knowledge that the algorithm has acquired up to that point. In this case, even previously learned Q-values may have to be re-learned, which limits the application possibilities of Q-learning.

However, the speed of the algorithm's learning process can be controlled by the choice of certain hyperparameters, such as the learning rate (or step-size parameter) α, which determines how quickly old information is replaced by new information, and the decay parameter θ, which determines the speed at which the exploration probability ε converges to zero. In terms of the learning rate, high values of α indicate extensive experimentation, as the algorithm very quickly forgets the action values it has learned in the past. Thus, values of α close to 1 may disrupt the learning process. In the experimental studies of Calvano et al. (2020) and Klein (2021), the authors used learning rates that were close to zero in order to ensure that the learning process was persistent. Additionally, the learning process is affected by the probability of exploration ε. The smaller ε, the smaller the chance that the algorithm randomly selects any price to learn the respective Q-value. If $\varepsilon = 0$, the algorithm will no longer explore but instead will exploit what it has learned so far, meaning that in any state the action leading to the highest Q-value in the given state will be chosen. As mentioned above, the speed of the probability of exploration ε converging to zero is determined by the decay parameter θ. Figure 2.1 depicts the convergence of the exploration probability for an example of 100,000 periods. Note that the value of the decay parameter θ is set in a way that convergence to a deterministic strategy is achieved (or even forced) as the exploration probability decreases to a value of 0.1% after 100,000 periods. Consequently, there is always a trade-off between forcing the algorithm to converge to a deterministic strategy and ensuring that it is provided

[12] Amazon changes the prices of its own products up to 8 times per hour, see `https://www.sellerlogic.com/en/blog/repricing-on-amazon/`.

with sufficient time to learn the optimal strategy, which has to be considered when choosing the respective hyperparameter values.

A further challenge of reinforcement learning algorithms is that in multi-agent settings, the choice of an action of one agent affects the reward signals of the other agents, which in turn affects the agent's own learning process and may result in already-learned Q-values no longer being valid. This type of problem is known as a "moving target" problem, implying a non-stationary environment. Contrary to the context of single-agent reinforcement learning algorithms, where convergence is guaranteed, no proof has been provided as yet that convergence to either a collusive or competitive equilibrium will occur in such a non-stationary, multi-agent setting. For this reason, Calvano et al. (2020) made use of the above-mentioned hyperparameter settings and applied heuristics to solve the problem, assuming convergence when the optimal strategy has not changed for 100,000 consecutive periods.

Returning to the time factor, Calvano et al. (2020) argued that the time needed to converge to a deterministic strategy is unproblematic if algorithms can be trained offline before being employed in the market. However, offline training does not seem to be representative of the real world since it might be very difficult to replicate realistic market conditions of online marketplaces in an offline setting. Strategies learned in the simulated environment may be unsuitable, making offline training in simple, stylized simulated market environments insufficient for making satisfactory decisions in real-world markets. This problem will be discussed in more detail in the following section.

2.4 Reinforcement Learning in Economic Applications

As described in Section 2.2, several studies have demonstrated that learning algorithms are able to learn collusive behavior and set prices that are significantly higher than the competitive prices in a standard Bertrand model. In these studies, many assumptions about market conditions are unrealistically restrictive. To investigate whether learning algorithms are able to learn collusive behavior in a more realistic setting, it seems sensible to first take a look at the assumptions made within the frameworks of the simulation models available:

- Firms produce one homogeneous good;
- firms charge a uniform price;

- firms compete exclusively on prices;
- the price a firm charges is determined only by its costs, the prices set by the competitors and the own past price;
- firms use the same (or even an identical) type of algorithm, and the algorithms are trained in the same scenario (sandbox);
- the environment is static, or only demand fluctuates, and the algorithms used are not changed externally over time.

In the following, the central explicit and implicit assumptions made in the simulation models are discussed in more detail and compared with the conditions that exist in reality.

An essential assumption made in the model by Klein (2021) as well as in the contribution by Calvano et al. (2020) concerns the number of goods offered by the firms. In both models it is assumed that **each firm offers exactly one product**. However, this is rather restrictive, in particular in consideration of firms such as online retailers, which are likely to employ price-setting algorithms. These firms generally offer several products that could be substitutes or complements, and algorithms have to learn which relationship between the products holds and have to take that into account when setting prices. For example, if the price of game consoles is changed, this usually has an impact on demand for video games. Firms also differ in the range of products they offer, which affects pricing as well. This can be illustrated by the following example of a duopoly: Firm 1 offers two complementary products, a and b, while firm 2 offers only product a. Firm 1 will charge a different price for product a than firm 2, as it takes into account the negative effect of a high price for product a on the demand for the complementary product b. Thus, it will charge a lower price for product a than firm 2, in order to stimulate demand for product b. Consequently, firm 2 arrives at a different profit-maximizing price than firm 1. In addition, when a decision is made to include a new product or brand in the range, the substitution relationships between all the products change, and these new relationships then have to be re-learned by the algorithm. If this happens online, it generally takes some time, because in order to observe and learn consumer reactions, demand must be given time to respond to the changed prices. This sets an upper limit to the speed at which the algorithm can change prices. In online learning, assumptions about the new substitution and complementarity relationships need to be integrated into the market model used to train the algorithm. Similar pricing issues occur when some firms are able to offer personalized products, but others cannot (simple examples include personalized fountain pens or laptops).

Next, we examine the extent to which the assumption that **firms set a uniform price** for their product is a reasonable one. The discussion on price-setting algorithms relates not only to the question of possible collusive behavior but also to the question whether, given the information acquired about their customers, firms will price discriminate and set personalized prices. This has already been observed in the United States, where companies used information about the ZIP code of their customers' residential address[13] or their operating system[14] to charge them different prices. Additionally, it has been demonstrated in experimental studies by, e.g., Shiller (2013), Ban and Keskin (2020), and Dubé and Misra (2017), that finer-grained price discrimination is possible with the use of machine learning.

A key point that is often raised in discussions about algorithmic pricing is that the large amount of customer data that is available to online firms makes it possible to divide them into different customer groups that differ in their willingness to pay. The firms would then have an incentive to demand different prices from different groups of customers for the same product, according to their respective willingness to pay (Acquisti and Varian, 2005; Reinartz, 2002; Wertenbroch and Skiera, 2002). In extreme cases, the classification of customer groups would be so fine that this type of price discrimination would lead to personalized pricing. If this were the case, a firm could skim off its customers' entire willingness to pay by charging personalized prices. On the one hand, this would lead to some tension between price discrimination and collusive pricing: if a firm is able to siphon off its customers' entire willingness to pay through personalized prices, why should it still have an interest in demanding a uniform collusive price? On the other hand, one could argue that firms could also collude on personalized prices. However, this could prove difficult for several reasons. For example, firms may have collected different data because the purchasing behaviors of their customers in the past were not the same; thus, the classification into different customer groups varies between firms – e.g., one firm could use a finer classification, and another a coarser one, so that one firm price discriminates while another firm charges a unit price because of a homogeneous customer group. Moreover, some firms may not price discriminate at all, because they do not want to be considered as unfair by their customers. In addition, personalized prices could be hard to detect by other firms, giving rise to asymmetric information that makes collusion more difficult.

Other forms of nonlinear pricing that make collusion more difficult include

[13] Websites Vary Prices, Deals Based on Users' Information, in: *Wall Street Journal*, December 24, 2012.

[14] On Orbitz, Mac Users Steered to Pricier Hotels, in: *Wall Street Journal*, August 23, 2012.

various forms of quantity or loyalty rebates. Here, the algorithm must not only calculate the corresponding prices but also the respective sales quantities for which the corresponding discount levels apply.[15] A further complication is posed by special promotions that are only valid for a short time span, such as Black Friday or Cyber Monday. The time span is generally too short for the algorithm to learn collusive behavior. In addition, competitors' actions on these days may be misinterpreted by the algorithm as a deviation from collusive behavior, because those actions do not represent a reaction to the firm's pricing decision but are instead determined by the external factor of these special events. Consequently, the occurrence of irregular special promotions and discounts represent a non-stationary environment, making it necessary to re-evaluate or re-learn certain Q-values.

To sum up, if firms do not charge a uniform price, then it is not immediately clear which of the different prices their own algorithm should be conditioned on. In the case of personalized prices, information would be needed not only about the prices, but also about the respective customers to whom these prices were charged.

In the simulation models, but also implicitly in many other contributions to the topic of algorithmic collusion, it is assumed that **firms compete exclusively on prices**. Stated otherwise, other competitive parameters, such as product quality, service, innovation, etc., are not considered. It is questionable whether this is a realistic assumption for markets where price-setting algorithms are used, or whether parameters other than price are also essential to competition. Particularly in online trading, a number of competitive parameters are present that are likely to hold similar significance for customers as the product's price, including the length of delivery times, the shipping costs, how easy it is to return products, the security of the transaction, the existence and quality of recommender and reputation systems, and the overall service the firm offers to the customer (Fagerstrøm and Ghinea, 2011; Jun et al., 2004; Zhao et al., 2015). In this context, the question Stigler (1968) posed at the end of the 1960s about the significance of price and non-price competition is relevant: "Will any monopoly profit achieved by suppressing price competition be eliminated by non-price competition?" Even if algorithms actually arrive at a collusive price, the firms would continue to compete on other parameters, and this non-price competition could be so intense that no noticeable effects would result from mere price collusion. Stigler (1968) pointed out that the result depends on the marginal costs of the respective competitive parameters. As this question can ultimately only be decided

[15] For various forms of online personalized sales promotion, see Changchien et al. (2004).

empirically, it is necessary to examine these costs in the digital economy, in particular in online trading. Only then can a statement be made as to whether the effects of a restriction of price competition can be compensated by competition on other parameters.

The formulation of the algorithm in the simulation models also implies that **a firm's price depends only on its costs, the prices of competitors, and possibly the firm's own price in the previous period**.[16] This is a rather restrictive assumption, because it does not consider that several other factors are also essential to pricing. One of these factors is a firm's inventory.[17] It is not only the size of the stock that is decisive, but also the age of the products – especially if the products are perishable goods, such as food, airline seats, or hotel rooms.[18] For example, if the stock of a perishable good is still quite high shortly before its expiry date, there is a strong incentive to lower the price of that product, so that the stock is reduced as quickly as possible before the product loses its value.[19] This indicates that not only are the amount of stock and the age of the products important, but also the speed with which the stock changes. If the inventory decreases too quickly, there is an incentive to raise prices so as to not run out of stock, to prevent supply problems and to keep customers who do not receive the product from switching to competitors. On the other hand, if inventory increases and there is a risk of exhausting storage capacity, there is a significant incentive to reduce prices. Therefore, the algorithm would have to take both pricing and inventory management into account simultaneously.[20] Since firms are likely to differ in their storage capacity, this results in a further asymmetry that makes collusion more difficult (Compte et al., 2002).

Of course, prices always depend on costs as well, but for retailers the costs are mainly determined by the purchase prices of the goods. These prices can develop very differently for different retailers, depending on which manufacturer they purchase their products from and how much they buy. Consequently, the assumption that firms are symmetric with respect to their cost is not very realistic. Similar to the afore-mentioned asymmetric capacity constraints, cost asymmetries also hinder collusion (Vasconcelos, 2005).

---

[16] The conditioning of the price on the own past price results from the design of the employed reinforcement learning algorithm.

[17] For a detailed overview of price optimization models that consider inventory replenishment, see Chen and Simchi-Levi (2012).

[18] Airline seats and hotel rooms are considered perishable because they cannot be "stored" for sale at a future date.

[19] Similar problems arise with the demand of seasonal items, which is prone to fluctuations within a year, but also within a week or even a day. Examples are fashion or gasoline, the latter of which is mostly demanded by commuters in the morning on their way to work.

[20] Initial approaches to this have been made, such as the work of Schlosser et al. (2018).

In terms of price-setting, the extent to which a firm should take its competitors' prices into account must be considered, in particular when firms produce differentiated goods: If competitors produce a close substitute, the algorithm should take into account that the price of that product is significantly more important than the price charged by a competitor with a distant substitute. In the former case, a price reduction by the competitor would lead to a significant decrease in the firm's demand, while a price change by a competitor with a distant substitute is unlikely to have a significant impact on the firm's demand, which therefore would not lead to a significant price change. In general, firms try to hinder substitutability by competitors' products through product differentiation.[21] This includes innovations, offering the product with different qualities, providing more (personalized) services, investing in advertising campaigns, and more. As a consequence, products in real markets are often more complex and can be personalized by giving the customer the opportunity to set certain configurable parameters (Dewan et al., 2000).

To sum up, in contrast to the assumptions made in the simulation models of price-setting algorithms, firms in markets with differentiated products will decide both on a pricing strategy and on the configuration of their goods and services, while consumers will base their purchasing decision on more complex utility functions that also take into account the product's attributes.[22] In the context of differentiated products, firms may also differ in their customer bases: one firm's customers may have quite distinct preferences for the specific product(s) offered, so that the price elasticity of demand for this group is relatively low, while other firms have customers who are more willing to buy a substitute, i.e., are characterized by a comparatively high price elasticity of demand. In such a case, a firm that faces price-inelastic demand function tends to charge a higher price than a firm facing price-elastic demand. Furthermore, asymmetries in firms' cost structure and storage capacity make collusion more difficult.

In the simulation studies, it is generally assumed that **all firms use exactly the same algorithm or at least the same type of algorithm**, e.g., Q-learning, but with different hyperparameter values such as learning rates and exploration probabilities. Besides Q-learning, there are many machine learning methods that can be used, such as Optimal Adaptive Learning or Deep Neural Networks, and it is not clear a priori why firms would independently choose the same algorithm or the same type of algorithm when they decide about which

[21] For a comprehensive overview of product differentiation see, e.g., Church and Ware (2000).

[22] For a detailed discussion of this issue we refer to Kephart et al. (2000), who examined how agents might deal with complex multi-attribute goods and services by investigating a variety of models that emphasize different aspects of product differentiation.

algorithm to employ for pricing. It is also not clear whether all firms use learning algorithms at all, or if they instead adopt comparatively simple adaptive rules, such as *undercut the market price by x% as long as it is above your cost.* In the latter case, collusion would hardly result. Therefore, the question arises whether different algorithms, employed in the same simulated market environment, converge to the same strategy and, if not, what would happen if these algorithms, having learned different strategies, then interact in a market. The simulation models implicitly assume that firms have somehow coordinated about the algorithm or the type of algorithm they use for price-setting. Fundamentally, a firm's choice of a certain algorithm is itself a strategic decision, because its profit depends on their competitors' strategies, be it employing algorithms or utilizing pricing rules. The decision as to which algorithm or method to use for pricing would have to be modeled as a strategic game, requiring examination as to whether coordination on the use of the same algorithm represents a Nash equilibrium in this game. Of course, it would have to be assumed that each firm knows which algorithms are available for the firm and its competitors, and which payoffs are associated with each strategy combination. It is not clear whether in such a game any coordination on a certain algorithm or learning procedure takes place, or whether an equilibrium exists at all in this game, i.e., a combination of algorithms or price-setting procedures that are mutual best replies.[23]

A similar problem is related to the simulated market environment that firms use for the offline training of their algorithms.[24] Even if firms employ the same algorithm, it is not clear whether each firm uses the same or at least a very similar simulated environment, nor the extent to which the simulated environment corresponds to the actual conditions. Thus, firms' perceptions of the market situation, the number and importance of their competitors, the way in which competitors react to their own price-setting, the relevance of other competitive parameters, such as service or quality, and further factors that are taken into account and others that are deemed to be of no importance – all this can change. Especially in markets with differentiated products, these perceptions may differ between firms, and it is likely that asymmetric information exists about the simulated environments other firms use to train their algorithms. Due to this asymmetric information, collusion is less likely to occur. It has been pointed out in the literature that without coordination between firms about the algorithm to use, it should be assumed that asymmetric information about competitive strategies will prevail: "Regarding a self-regulatory option, note that at any point in time each

[23] This topic will be discussed further in Section 2.5.

[24] See also Section 2.3.3.

firm may only know the algorithms it employs to carry out its market strategies, but not necessarily the algorithms employed by other firms operating in the market. This means that each firm will not be able to simulate with a high enough degree of accuracy how its own algorithms will behave period after period as they interact with other algorithms" (Gata, 2021, p. 87). Thus, if firms have different perceptions of the market situation and train their algorithms in different simulated environments, it is not clear whether even identical algorithms will converge to the same strategy and what the market outcome would be in such a case.

Alternatively, the assumption that all firms use the same algorithm or very similar algorithms and train them in an identical simulated market environment can be seen as raising the coordination issue to a higher level: firms no longer discuss prices, but they coordinate on the type of learning algorithm they will use, and in which simulated market environment they will train this algorithm. This would essentially be a rather complex way for people to use algorithms as a facilitative device to deliberately bring about collusive behavior, similar to the Topkins case.[25] In principle, however, this would not raise any novel concerns related to competition law – only the problem of proof could prove to be difficult.

In the simulation models, it is further assumed that **both the algorithms and the market environment remain essentially unchanged**. This assumption entails that collusive behavior is more likely to be observed in these simulations, since theory shows that tacit collusion is more difficult if the environment changes due to, for example, an increase in the number of market participants or the introduction of a new, innovative product (Frass and Greer, 1977; Ivaldi et al., 2003). In this paragraph, we discuss which changes of the algorithm and the market environment are likely to be observed in reality and explain what these changes entail in the algorithmic setting.

If the algorithms lead to a collusive price, it is likely that this would have an effect on the colluding firms as well as on potential competitors, leading to firms potentially reprogramming the algorithm, whereby it is no longer static. This follows from the fact that, as the collusive firms would now charge higher prices, sales volumes would decrease, inventory would increase, and some of the production capacity would go unused. Similar to the case of traditional cartels, it is likely that the firm, especially in cases of perishable or seasonal goods, would charge lower prices to clear the inventory or to exploit production capacity. To do so, the firm would have to override their algorithm, thereby undermining and

[25] In this case, the U.S. Department of Justice prosecuted two retailers for aligning their pricing algorithms to increase the price of posters online. Although this practice involved pricing algorithms, the case is similar to explicit collusion, as both firms agreed on a price-fixing strategy via information sharing.

eliminating the collusion brought about by the algorithms.[26] Additionally, if a firm's management noticed that the pricing algorithms had reached a collusive equilibrium at higher prices, it might be incentivized to override the price set by the algorithm and charge a lower one to avoid antitrust sanctions.

Moreover, it can be expected that the introduction of new and improved algorithms would lead to the replacement of the previous ones, which is also in contrast to the assumption in the simulation models of the static character of algorithms. Given the assumption that firms modify the algorithms they use or introduce new ones at different points in time, it is not clear what effect this has on pricing, but it is likely that this will not facilitate collusive behavior.

We are also critical of the assumption that market conditions do not change. If algorithms have led to a collusive equilibrium with correspondingly higher prices and profits, potential competitors are likely to observe these higher profits and enter the market, at least if barriers to entry are low. Technically speaking, this would represent a change in the environment. Moreover, changes in the number of firms also lead to changes in the dimensions of the Q-matrix, which is especially challenging in the case of market entry as this dramatically increases the dimensions of the Q-matrix, as pointed out in Section 2.3.3. In addition to learning new Q-values, old Q-values may need to be re-learned, which is quite time-intensive.

Although the model of Calvano et al. (2020) considers the possibility of market entry and exit, both are considered as stochastic and not as systematic, which one would expect in markets with above average profits and low entry barriers wherein the number of firms would continuously increase. However, this would cause a reaction from the incumbents to protect or extend their market share against the entrants (Harrington, 1989; Ordover and Saloner, 1989). For example, firms might engage in predatory pricing strategies, expand output, introduce new products, or redesign existing ones, which would again entail changes in the market environment.

Further changes in the environment could be caused by governmental regulations: for example, firms' action spaces could be restricted by regulation such as binding upper or lower price limits. A real-world example of lower price limits can be observed for some agricultural products. Another example is a regulatory

[26] One real-life example of halting and overriding an algorithm is that of Uber, who's pricing algorithm caused price surges of more than 200% in some districts of London since demand for a taxi drive increased as a response to reports of a terrorist attack at London Bridge on June 3, 2017. The company had to manually stop this pricing mechanism to avoid public outrage(Bertini and Koenigsberg, 2021).

measure implemented by the Austrian government in 2011, under which gas stations are only allowed to increase their prices once a day, at 12 a.m., although a price reduction is allowed at any time. In the simplest case, such regulatory measures only lead to a limitation of the action set and the set of states. However, in the presence of price regulations, the applicability of the pricing algorithms employed in the simulation studies are limited since exploration is difficult or not useful.

2.5 Price-Setting Behavior in More Realistic Environments

Part of the above arguments will be substantiated through our simulations. While Calvano et al. (2020) and Klein (2021) have shown that simple Q-learning algorithms are capable of generating supra-competitive prices, we argue that this result holds only in very simple market environments. Consequently, it is uncertain whether these results can be transferred to a more realistic setting. To examine whether the results of the simulation studies still bear up under more realistic conditions, we slightly extend the model of Calvano et al. (2020) by considering a duopoly where the two firms offer differentiated products. We then compare scenarios wherein different types of pricing rules and algorithms are used by the firms. The two reinforcement learning algorithms we consider are Q-learning and an MAB algorithm which are explained in further detail in Sections 2.3.1 and 2.3.2. Since not every firm is able to put the same effort into developing a complex pricing algorithm, and since price-setting on popular platforms like Amazon is usually done using simple if-then functions, we also introduce firms that set their prices based on a meeting competition clause (MCC) or a simple price heuristic (H). An example of an MCC is a price guarantee, i.e. consumers can claim a price discount up to the difference to the lowest price in the market.[27] An example of a price heuristic employed by some suppliers is the automatic pricing option at Amazon. The *Amazon Seller Central Europe* website states that "[f]or example, you can create a rule that stays 0,10 EUR below the Buy Box price".[28] The MCC implemented in our experiments is designed to always set a price that is closest to the perfectly collusive price and automatically match any lower price set by a competitor. The heuristic is implemented such that it always undercuts

[27] For a survey on different types of price relationship agreements see Office of Fair Trading (2012).

[28] `https://sellercentral.amazon.de/help/hub/reference/external/201995750?ref=efph_201995750_cont_202166010&locale=en-DE`.

the competitor's price by the smallest monetary unit provided that the price is not below marginal cost. The heuristic matches the competitor's price if it is the lowest price in the action set.

To illustrate how Q-learning algorithms compare to multi-armed bandits and simple pricing rules, we conduct experiments in which two agents act in a market. Demand for firm $i = 1, 2$ is characterized by the demand function

$$D_i(p_i, p_j) = 2 - p_i + \frac{1}{2}p_j,$$

where $j = 1, 2$ and $j \neq i$. Firms produce heterogeneous goods with marginal costs of $c = 1$ and compete in prices. The prices and profits under oligopolistic competition are $p_i^N = 2.0$ and $\pi_i^N = 1.0$. Under perfect collusion, firms charge a price $p_i^C = 2.5$ and earn a profit of $\pi_i^C = 1.125$.

It is assumed in the simulations that the action sets have a lower bound of 10 percent below the price in a non-cooperative Nash equilibrium and an upper bound of 10 percent above the price under perfect collusion. This interval is then equally divided into $k = 6$ prices. Thus, the action sets of both agents are given as $A_i = \{1.8, 1.99, 2.18, 2.37, 2.56, 2.75\}$, describing the prices the firms may set. Exploration follows an ε-greedy strategy wherein the exploration probability is 1 in the first period (i.e., $\varepsilon_0 = 1$) and decreases to a value of 0.1% after 700,000 periods. The learning rate is $\alpha = 0.1$, and the discount factor is $\delta = 0.95$.

Each experiment consists of ten runs, each lasting for a maximum of 2,000,000 periods. A run is terminated earlier if both agents choose the same action or repeat the same cycle of actions for 10,000 consecutive periods.

In Table 2.1, simulation results are reported for the averages over the final 1,000 periods and of all runs.[29] The first column indicates the experimental setting, or what type of agents interact in the environment. For example, Q vs Q indicates that two independent Q-learning agents are used. In all settings, firm 1 is the first type of agent and firm 2 is the second one. In the second scenario, for example, firm 1 employs a Q-learning agent and firm 2 employs an MAB.

Comparing simulation results of the different experimental settings, it becomes clear that the resulting prices heavily depend on the type of algorithm or pricing rule that is applied by the firms, since observed prices range from below the competitive to above the collusive price level. Prices above the collusive level result from the choice of the action set. Due to the coarse discretization, it is not

[29] In these periods, the learning process is complete and the agents repeatedly choose the same exploitative action and explore with a probability of far below 0.1%. Prices and profits in italicized settings were not determined using simulations, but rather based on economic reasoning.

Experimental Setting	p_1	p_2	π_1	π_2
Q vs Q	2.180	2.180	1.074	1.074
Q vs MAB	2.076	2.161	1.073	1.015
MAB vs MAB	1.990	1.990	0.995	0.995
Q vs Heuristic	2.370	2.180	0.986	1.186
Q vs MCC	2.560	2.560	1.123	1.123
MAB vs Heuristic	2.370	2.180	0.986	1.186
MAB vs MCC	2.560	2.560	1.123	1.123
MCC vs MCC	2.560	2.560	1.123	1.123
MCC vs Heuristic	1.800	1.800	0.880	0.880
Heuristic vs Heuristic	1.800	1.800	0.880	0.880

TABLE 2.1: Resulting Prices and Profits

possible to set the exact collusive price ($p_i^C = 2.5$). Consequently, the algorithm chooses the price from the action set that is closest to the collusive price. Prices below the competitive level ($p_i^N = 2.0$) result from the design of the price heuristic, which states that the competitor's price is always undercut provided it is above marginal cost. Moreover, prices are bounded by the lower bound of the action set.

The first scenario confirms the results from other experiments that two independent Q-learning algorithms can learn to set collusive prices; however, the agents do not achieve a perfectly collusive outcome. In the second scenario, in which firm 2 employs an MAB, prices are lower and, as a result, the the firms' profits decrease compared to the first scenario. If both firms employ MAB algorithms, an even lower price and thus lower profits follow. As a result, if all firms employ reinforcement learning algorithms, choosing the Q-learning algorithm is a dominant strategy for both firms. The subsequent four scenarios include one firm employing a learning algorithm and the other firm using a simple pricing rule. In the fourth scenario, a Q-learning agent (firm 1) interacts with an opponent using the price heuristic (firm 2). It can be seen that the profit of firm 1 drops even lower, while firm 2 is able to achieve a profit that is larger than its share under perfect collusion. In this case, the agents set different supra-competitive prices, and firm 2 benefits from the resulting price level as well as from the fact that the price of the Q-learning agent is even higher, which provides firm 2 with a larger share of demand. In the fifth scenario, if firm 2 utilizes an MCC instead, both agents set a price that is slightly above the collusive level, and thus earn almost the same profit as under perfect collusion. The following two scenarios replicate the two proceeding scenarios but substitute the Q-learning agent for an MAB. The last three scenarios do not involve learning algorithms, so simulations are not necessary since prices and profits can be determined by economic reasoning. If both firms employ an MCC, perfectly collusive behavior occurs. This is in line with the existing literature on the collusive efficacy of competition clauses (Doyle, 1988; Logan and Lutter, 1989).[30] If one firm uses an MCC and the other implements the heuristic, the price of the firm using the MCC is undercut until marginal cost is met. Since action sets are bounded, firms set the lowest price in their action set. The same applies in the case where both firms use the heuristic.

As profits depend on the decision of whether learning algorithms or simple pricing rules are employed by the firm and its competitor, we can consider a firm's choice of a certain pricing algorithm or pricing rule to be a strategic decision in itself. Modeling this decision as a strategic game would entail that each firm knows which strategies are available for the firm and its competitor and which payoffs are associated with each strategy combination.[31]

The payoffs for this game are given by the results from the pricing algorithm experiments as averages over the last 1,000 periods over all runs and are shown in Table 2.2.

	Q	MAB	MCC	H
Q	1.074, 1.074	1.073, 1.015	1.123, 1.123	0.986, 1.186
MAB	1.015, 1.073	0.995, 0.995	1.123, 1.123	0.986, 1.186
MCC	1.123, 1.123	1.123, 1.123	1.123, 1.123	0.880, 0.880
H	1.186, 0.986	1.186, 0.986	0.880, 0.880	0.880, 0.880

Table 2.2: Payoff Matrix of the Strategic Game

The Nash equilibria in pure strategies are {(MCC,MCC), (Q,H), (H,Q), (MAB,H), (H,MAB)}. Of these equilibria, the equilibrium (MCC,MCC) gives the highest sum of payoffs. Moreover, it is the only symmetric equilibrium of the symmetric game. Thus, the equilibrium (MCC,MCC) could represent a so-called "focal point" (Schelling, 1960).

Moreover, our simulations demonstrate that when both firms use a learning algorithm, the outcome is not an equilibrium when alternative price setting rules are available. This holds true for the Q-learning strategy as well as for the even easier to implement MAB. Thus, the situation described by Calvano et al. (2020) and Klein (2021), wherein firms employ such an algorithm which then learns to set supra-competitive prices, represents a situation wherein both firms would

[30] In the theoretical papers of Doyle (1988) and Logan and Lutter (1989), it is claimed that all retailers have to adopt competition clauses in order to enforce collusion. In a more recent study, Trost (2021) questions this assumption and argues that it is in general not required that all retailers have to adopt competition clauses in order to enforce collusion.

[31] For games with imperfect information, finding a (Bayesian) Nash equilibrium is far more challenging.

have an incentive to deviate from their strategy and employ a simple pricing rule instead.

2.6 Conclusion

To summarize our findings, we can conclude that the assumptions made in simulation studies, which have demonstrated that learning algorithms are able to autonomously collude, are not applicable to economic reality. We showed this by first presenting several technical challenges that arise when algorithms are applied in the real world. For example, a significant increase in the dimensions of the Q-matrix comes with an increase in the number of agents, actions, or states, as well as a large time span needed to complete the learning process or deal with non-stationary environments.

In a second step, we took a closer look at the assumptions made in simulation studies where collusive behavior of algorithms was detected and tried to transfer these to economic reality. In particular, we considered the assumptions that firms offer only one homogeneous product at a time; that they charge a uniform price and thereby base their decisions only on their own costs, the price of the competitor(s), and their own price in the previous period; that they compete exclusively in prices; that they make use of the same or even identical type of algorithm, which are trained in the same scenario; that the environment is static and that the algorithms used are not changed externally over time.

We then further investigated the assumption that firms make use of the same type of algorithm and ran simulations where firms using a learning algorithm compete with firms either employing a learning algorithm as well or a simple pricing rule to set prices. Our simulations showed that when both firms use a learning algorithm, the outcome is not an equilibrium when alternative price setting rules are available. This result suggests that firms are more likely to use a simple pricing rule like price guarantees which are significantly cheaper to implement and also promise a higher payoff. Therefore, fears that learning algorithms may result in more opportunities for collusion appear unfounded, as even comparatively simple pricing rules seem to be more effective in producing cartel-like behavior.[32] Consequently, which antitrust regulations should be adopted in these markets depends on which pricing rules are available to the firms. At present,

[32] That price matching guarantees lead to higher prices in online markets has been shown by Zhuo (2017).

therefore, there seems to be no need for specific competition law regulations regarding algorithmic pricing. A closer examination of the anticompetitive effects of simple pricing rules would be more appropriate instead.

Although it is no longer assumed that firms use the same type of algorithm, the setting is still very stylized and it is debatable whether these supra-competitive profits and prices can persist in economic reality, where far more pricing options are available. For example, it can be observed that large online platforms, such as Amazon, use their own pricing software. Furthermore, a market for revenue management, repricing and price optimization has quickly emerged in recent years. The spectrum of solutions on offer ranges from simple adaptive approaches through modular systems in which the firm can determine the parameters, such as competitors' prices, the purchase price, the day of the week, the time of day, the stock level, the sales figures, and the weights with which these parameters are included in its own pricing, all the way to approaches from machine learning. The objective of these algorithms is not only direct profit maximization but also maximization of the chance of ending up in Amazon's Buy Box (which is more likely if a low price is charged), of appearing on the first page of search results, of increasing customer loyalty, or of tapping into a new market segment. Given the algorithms offered, the chances are good that different firms use different types of algorithms. However, it is unclear what the outcome might be. A problem could arise, however, if firms in the same market would coordinate on using an MCC or the algorithm of the same repricer, which could lead to a hub-and-spoke like cartel. This would shift the coordination problem on a higher level: firms do not coordinate on prices but on using the same algorithm or pricing rule. Since this behavior would be similar to explicit collusion this would not raise any novel concerns related to competition law.

To sum up, due to the above-mentioned aspects, autonomous collusion by learning algorithms does not currently seem to be a major competition concern. Nevertheless, this does not mean that no other concerns related to algorithms justify a close monitoring of what is going on. Structural market conditions are changing with the increasing use of algorithms, since market transparency has increased horizontally (between firms) as well as vertically (between firms and consumers). Moreover, algorithms are a new and efficient facilitative device that may be used by humans to enable and stabilize (explicit) collusion, as happened in the Topkins[33] and Trod[34] cases. In those cases, online firms selling posters on

[33] Department of Justice, Case 3:15-cr-00201-WHO, 6 April 2015, `https://www.justice.gov/atr/case/us-v-david-topkins`.

[34] Department of Justice, Case 3:15-cr-00419-WHO, 27 August 2015, `https://www.justice.gov/atr/case/us-v-daniel-william-aston-and-trod-limited`.

Amazon Marketplace agreed to use the same algorithm to fix the prices of their products. Another example of using price-setting algorithms to coordinate prices is the Eturas case, in which several travel agencies employed a booking system to engage in collusive behavior.[35] However, these cases were covered by existing laws.

Another area in which algorithms could be used in an anticompetitive manner is in vertical relations – specifically, enforcing resale price maintenance. This practice was already observed and sanctioned by the European Commission in 2018, when electronics manufacturers in Japan, Taiwan, and the Netherlands requested that online retailers who offered their products not sell them below a certain price.[36] The price-setting of the online retailers was tracked by sophisticated monitoring algorithms, and retailers who did not follow the instructions faced threats or sanctions, such as blocking of supplies. The commission emphasized that many retailers in online markets use pricing software that adapts their retail prices to those of competitors. They argued that horizontal pricing restrictions imposed by manufacturers therefore have a broader impact on overall online pricing for the respective products, since the prices of competing retailers are based on those of the restricted retailers. According to the European Commission, these interdependencies have kept the products' price levels high.

Additionally, the widely published statement by Competition Commissioner Vestager made it clear that firms' use of algorithms is a high priority for competition authorities and that a firm using automated systems "will be held responsible for what it does".[37] It remains to be seen whether this will include tacit collusion in the future. At the present, however, this can largely be ruled out.

[35] European Commission, Document 62014CJ0074, 21 January 2016, `https://eur-lex.europa.eu/legal-content/de/TXT/?uri=CELEX:62014CJ0074`.

[36] European Commission, Press release IP/18/4601, 24 July 2018, `https://ec.europa.eu/commission/presscorner/detail/en/IP_18_4601`.

[37] Speech by Commissioner Margrethe Vestager at the Bundeskartellamt 18th Conference on Competition, Berlin, 16 March 2017.

References

Acquisti, A. and H. R. Varian (2005). Conditioning prices on purchase history. *Marketing Science*, 24.3, pp. 367–381.

Assad, S., R. Clark, D. Ershov, and L. Xu (2020). Algorithmic pricing and competition: Empirical evidence from the German retail gasoline market. CESifo Working Paper No. 8521. DOI: 10.2139/ssrn.3682021.

Athey, S. and G. W. Imbens (2019). Machine learning methods that economists should know about. *Annual Review of Economicsvvvv*, 11, pp. 685–725.

Autorité de la concurrence and Bundeskartellamt (2019). Algorithms and competition. Report. URL: https://www.autoritedelaconcurrence.fr/sites/default/files/algorithms-and-competition.pdf (visited on 07/01/2022).

Ban, G.-Y. and N. B. Keskin (2020). Personalized dynamic pricing with machine learning: High dimensional features and heterogeneous elasticity. *Management Science*, 67.9, pp. 5549–5568.

Bertini, M. and O. Koenigsberg (2021). The Pitfalls of Pricing Algorithms. *Harvard Business Review*, 9.

Brown, Z. Y. and A. MacKay (2021). Competition in pricing algorithms. Working Paper No. 28860. National Bureau of Economic Research. DOI: 10.3386/w28860.

Calvano, E., G. Calzolari, V. Denicolò, and S. Pastorello (2019). Algorithmic pricing: What implications for competition policy? *Review of Industrial Organization*, 55.1, pp. 155–171.

— (2020). Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, 110.10, pp. 3267–3329.

— (2021). Algorithmic collusion with imperfect monitoring. *International Journal of Industrial Organization*, 79, p. 102712.

Cao, D., W. Hu, J. Zhao, G. Zhang, B. Zhang, Z. Liu, Z. Chen, and F. Blaabjerg (2020). Reinforcement learning and its applications in modern power and energy systems: A review. *Journal of Modern Power Systems and Clean Energy*, 8.6, pp. 1029–1042.

Changchien, S. W., C.-F. Lee, and Y.-J. Hsu (2004). On-line personalized sales promotion in electronic commerce. *Expert Systems with Applications*, 27.1, pp. 35–52.

Chen, X. and D. Simchi-Levi (2012). Pricing and inventory management. In: *The Oxford Handbook of Pricing Management*. Ed. by Ö. Özer and R. Phillips. Oxford: Oxford University Press, pp. 784–824.

Church, J. R. and R. Ware (2000). *Industrial Organization: A Strategic Approach.* USA: McGraw-Hill.

Competition and Markets Authority (2021). Algorithms: How they can reduce competition and harm consumers. URL: `https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers` (visited on 06/29/2022).

Compte, O., F. Jenny, and P. Rey (2002). Capacity constraints, mergers and collusion. *European Economic Review*, 46.1, pp. 1–29.

Dewan, R., B. Jing, and A. Seidmann (2000). Adoption of Internet-based product customization and pricing strategies. *Journal of Management Information Systems*, 17.2, pp. 9–28.

Doyle, C. (1988). Different selling strategies in Bertrand oligopoly. *Economics Letters*, 28.4, pp. 387–390.

Dubé, J.-P. and S. Misra (2017). Scalable price targeting. Working Paper No. 23775. National Bureau of Economic Research. DOI: `10.3386/w23775`.

Ezrachi, A. and M. E. Stucke (2015). Artificial intelligence and collusion: When computers inhibit competition. Legal Studies Research Paper Series No. 267. DOI: `10.2139/ssrn.2591874`.

— (2016). *Virtual Competition: the promise and perils of the algorithm-driven economy.* 2nd ed. Cambridge, Massachusetts: Harvard University Press.

Fagerstrøm, A. and G. Ghinea (2011). On the motivating impact of price and online recommendations at the point of online purchase. *International Journal of Information Management*, 31.2, pp. 103–110.

Federal Trade Commission (Apr. 8, 2020). Using artificial intelligence and algorithms. URL: `https://www.ftc.gov/business-guidance/blog/2020/04/using-artificial-intelligence-and-algorithms` (visited on 07/27/2022).

— (Apr. 19, 2021). Aiming for truth, fairness, and equity in your company's use of AI. URL: `https://www.ftc.gov/business-guidance/blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai` (visited on 07/27/2022).

Frass, A. G. and D. F. Greer (1977). Market structure and price collusion: An empirical analysis. *The Journal of Industrial Economics*, 26.1, pp. 21–44.

Gal, M. S. (2017). Algorithmic facilitated coordination: Market and legal solutions. *CPI Antitrust Chronicle*, pp. 1–7.

Gata, J. E. (2021). Collusion between algorithms: a literature review and limits to enforcement. *European Review of Business Economics*, 1.1, pp. 73–94.

Green, E. J. and R. H. Porter (1984). Noncooperative collusion under imperfect price information. *Econometrica*, 52.1, pp. 87–100.

Hansen, K., K. Misra, and M. Pai (2020). Algorithmic collusion: Supra-competitive prices via independent algorithms. CEPR Discussion Paper No. DP14372. URL: https://cepr.org/publications/dp14372.

Harrington, J. E. (1989). Collusion and predation under (almost) free entry. *International Journal of Industrial Organization*, 7.3, pp. 381–401.

— (2018). Developing competition law for collusion by autonomous artificial agents. *Journal of Competition Law & Economics*, 14.3, pp. 331–363.

Ittoo, A. and N. Petit (2017). Algorithmic pricing agents and tacit collusion: A technological perspective. In: *L'Intelligence Artificielle et le Droit*. Ed. by H. Jacquemin and A. de Streel. Bruxelles: Larcier, pp. 241–256.

Ivaldi, M., B. Jullien, P. Rey, P. Seabright, and J. Tirole (2003). The economics of tacit collusion. Report for DG Competition, European Commission. URL: https://publications.ut-capitole.fr/id/eprint/1200/1/tacit_collusion.pdf (visited on 07/27/2022).

Jun, M., Z. Yang, and D. Kim (2004). Customers' perceptions of online retailing service quality and their satisfaction. *International Journal of Quality & Reliability Management*, 21.8, pp. 817–840.

Kephart, J. O., J. E. Hanson, and A. R. Greenwald (2000). Dynamic pricing by software agents. *Computer Networks*, 32.6, pp. 731–752.

Klein, T. (2021). Autonomous algorithmic collusion: Q-learning under sequential pricing. *RAND Journal of Economics*, 52.3, pp. 538–599.

Logan, J. W. and R. W. Lutter (1989). Guaranteed lowest prices: do they facilitate collusion? *Economics Letters*, 31.2, pp. 189–192.

MacKay, A. and S. N. Weinstein (2022). Dynamic pricing algorithms, consumer harm, and regulatory response. *Washington University Law Review*, 100, pp. 111–174.

Mehra, S. K. (2016). Antitrust and the robo-seller: Competition in the time of algorithms. *Minnesota Law Review*, 100, pp. 1323–1375.

Office of Fair Trading (2012). Can 'Fair' Prices be Unfair? A Review of Price Relationship Agreements. OFT 1438. URL: https://www.learlab.com/wp-content/uploads/2016/04/Can-%E2%80%98Fair%E2%80%99-Prices-Be-Unfair_-A-Review-of-Price-Relationship-Agreements.pdf (visited on 09/29/2022).

Ordover, J. A. and G. Saloner (1989). Predation, monopolization, and antitrust. In: *Handbook of Industrial Organization*. Ed. by R. Schmalensee and R. D. Willig. Vol. 2. Amsterdam, The Netherlands: Elsevier, pp. 537–596.

Reinartz, W. (2002). Customizing prices in online markets. *Emerging Issues in Management*, 1, pp. 55–65.

Schelling, T. (1960). *The Strategy of Conflict*. Cambridge, Massachusetts: Harvard University Press.

Schlosser, R., C. Walther, M. Boissier, and M. Uflacker (2018). Data-driven inventory management and dynamic pricing competition on online marketplaces. In: *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI 2018)*. Ed. by J. Lang. AAAI Press/IJCAI, pp. 5856–5858.

Schwalbe, U. (2018). Algorithms, machine learning and collusion. *Journal of Competition Law & Economics*, 14.4, pp. 568–607.

Shiller, B. R. (2013). *First Degree Price Discrimination Using Big Data*. Waltham, Massachusetts: Brandeis University, Department of Economics and International Business School.

Slivkins, A. (2019). Introduction to multi-armed bandits. *Computing Research Repository*. DOI: 10.48550/arXiv.1904.07272.

Stigler, G. J. (1968). Price and non-price competition. *Journal of Political Economy*, 76.1, pp. 149–154.

Sutton, R. S. and A. G. Barto (2018). *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts: MIT Press.

Tirole, J. (1988). *The Theory of Industrial Organization*. Cambridge, Massachusetts: MIT Press.

Trost, M. M. (2021). The collusive efficacy of competition clauses in Bertrand markets with capacity-constrained retailers. SSRN Working Paper. DOI: 10.2139/ssrn.3866328.

Vasconcelos, H. (2005). Tacit collusion, cost asymmetries, and mergers. *RAND Journal of Economics*, 36.1, pp. 39–62.

Waltman, L. and U. Kaymak (2008). Q-learning agents in a Cournot oligopoly model. *Journal of Economic Dynamics and Control*, 32.10, pp. 3275–3293.

Watkins, C. and P. Dayan (1992). Q-learning. *Machine Learning*, 8.3, pp. 279–292.

Wertenbroch, K. and B. Skiera (2002). Measuring consumers' willingness to pay at the point of purchase. *Journal of Marketing Research*, 39.2, pp. 228–241.

Zhao, X., L. Wang, X. Guo, and R. Law (2015). The influence of online reviews to online hotel booking intentions. *International Journal of Contemporary Hospitality Management*, 27.6, pp. 1343–1364.

Zhuo, R. (2017). Do Low-Price Guarantees Guarantee Low Prices? Evidence from Competition between Amazon and Big-Box Stores. *Journal of Industrial Economics*, 65.4, pp. 719–738.

Chapter 3

Price Discrimination with Inequity-Averse Consumers: A Reinforcement Learning Approach

This chapter is the author's own work.

Abstract

With the advent of big data, unique opportunities arise for data collection and analysis and thus for personalized pricing. This study combines approaches from machine learning and behavioral economics to answer the question whether algorithms are able to exploit consumers' systematic deviations from rational behavior. In particular, this paper discusses the ability of reinforcement learning – a type of machine learning in which agents learn from interacting autonomously with their environment – to price discriminate based on a simulation of a price-setting algorithm. Thereby, personalized prices are set considering additional information about consumer sensitivities to analyze market outcomes for consumers who have a preference for fair, equitable outcomes. For this purpose, we compare a situation that does not consider consumers' sense of fairness to one in which we allow for inequity-averse consumers. We show that the algorithm learns to charge different, revenue-maximizing prices and simultaneously increase fairness in terms of a more homogeneous distribution of prices. We conclude that consumers' sense of fairness, which have prevented firms from engaging in price discrimination, can be incorporated into firms' pricing decisions with the help of learning algorithms, making differential pricing strategies more feasible.

3.1 Introduction

Engaging in price discrimination and determining the revenue-maximizing price of a product or service for a particular customer is challenging. It requires, among others, knowledge of the customer's willingness to pay and estimation of demands. With the advent of big data, unique opportunities arise for data collection and analysis and thus for personalized pricing. Concerns have been raised, besides the discussion about collusive behavior (Ezrachi and Stucke, 2015, 2016; Woodcock, 2017), that learning algorithms might engage in price discrimination (Office of Fair Trading, 2013; Reinartz, 2002) as these algorithms are able to use information about customers to segment them into ever-smaller groups based on certain characteristics related to their willingness to pay. Reports about customers who were charged different prices for the same product or service have been published in the *Wall Street Journal*, the *Washington Post*, and the German business news magazine *Wirtschaftswoche*.[1] Consumers tend to be highly sensitive to such attempts as they have a strong sense of fairness (Kahneman et al., 1986), which might explain why price discrimination is still relatively rare in economic reality (Competition and Markets Authority, 2018; Executive Office of the President of the United States, 2015; Odlyzko, 2009; Office of Fair Trading, 2013).

This begs the question of whether it is possible for price-setting algorithms to learn to engage in price discrimination on the basis of fairness to avoid upsetting customers but still maximize expected revenues by charging personalized prices. Therefore, we will discuss not only the potential of algorithms to conduct price discrimination but also the ability of self-learning pricing algorithms to consider and possibly exploit customers' deviations from rational behavior. If algorithms are able to consider customers' inequity aversion while setting differential prices, price discrimination is assumed to be more likely to occur. This is also in line with growing public policy concerns discussed, for example, by Bourreau and Streel (2018), the United Kingdom's Competition and Markets Authority (2018), and the Executive Office of the President of the United States (2015).

However, the effect of personalized pricing on consumer welfare is ambiguous. The Competition and Markets Authority (2018) stated that, in many cases, personalized pricing may be beneficial. Especially in markets with switching costs, the ability to offer targeted discounts to consumers might help potential entrants to compete with incumbents and could lead to expanding overall output in those

[1] On Orbitz, Mac Users Steered to Pricier Hotels, in: *Wall Street Journal*, August 23, 2012; On the Web, Price Tags Blur, *Washington Post*, September 27, 2000; Der Preis ist heiß, *Wirtschaftswoche*, March 02, 2017.

markets. According to the Competition and Markets Authority, however, personalized pricing may be harmful to consumers in some situations, particularly if there is a lack of competition in the market, if discrimination is particularly complex or opaque to consumers, or if consumers lose trust in the market and, as a consequence, withdraw their demand or refuse to continue to participate in it.

Effective price discrimination relies on the identification of different customer groups. For each group, certain assumptions about price sensitivities must be made. For instance, groups could be defined by customers' locations, communication channels, click behavior, or time spent on a certain website. The price sensitivities within each group should be taken into account by the algorithm in order to maximize expected revenues. In the two simulated scenarios studied here, we assume that customer groups can be identified and allow for differential pricing. An algorithm is simulated that sets personalized prices that, in a first scenario, must learn to set the revenue-maximizing prices to customers who have no fairness considerations. In a second scenario, we want to investigate whether the algorithm is able to learn the interdependencies between prices of different customer groups since the algorithm is now dealing with inequity-averse consumers who have a preference for fair outcomes. For simplification, it is assumed that fairness is maximized if everyone pays the same price. This equality-based fairness approach is widely used in the relevant literature, for example, by Fehr and Schmidt (1999) or Bolton and Ockenfels (2000). Under this assumption, the learning process of the algorithm should provide homogeneous prices among customer groups while simultaneously taking into account the price sensitivities within each group in order to maximize the expected revenue. Compared to the first scenario, an improvement in fairness can be observed in the situation wherein inequity-averse consumers are considered, while the algorithm maintains the goal of maximizing the firm's revenue.

The paper is organized as follows. Section 3.2 provides a brief introduction into the main concepts of price discrimination, fairness, and reinforcement learning. The methodology is then presented in Section 3.3, where the different groups of customers are introduced, and reinforcement learning is applied to the differential pricing problem. The simulation results are presented and discussed in Section 3.4. Robustness checks can be found in Section 3.5 followed by a conclusion in Section 3.6.

3.2 Price Discrimination and Fairness in Algorithmic Settings

In this section, the main economic classifications of price discrimination are revisited. Subsequently, real-world examples where some vendors have already engaged in price discrimination are described. We also introduce several studies that showed how artificial intelligence (AI) methods have contributed to facilitating price discrimination. As fairness aspects are considered, a method for measuring perceived fairness is presented, followed by a brief introduction to reinforcement learning, which is used as an algorithmic setting in this study.

3.2.1 Price Discrimination in Theory

In general, a firm engages in price discrimination when it charges different prices for two units of the same or similar products, wherein the price difference does not reflect any cost difference (Stigler, 1966). Classical economic literature distinguishes between three types of price discrimination (Pigou, 1920).[2] *First-degree price discrimination* occurs when the seller charges each customer the maximum price that he or she is willing to pay. Even if firms do not have sufficient information to assess each consumer's reservation price, they can still conduct imperfect price discrimination, known as *third-degree price discrimination* or group pricing. In this practice, sellers segment their customers into broad categories according to observable characteristics; these categories are charged different prices. Third-degree price discrimination is probably the most common form of price discrimination. If firms are able to use information technologies to collect and process a large amount of data, they can improve their knowledge of consumers' preferences. As a result, they might be able to refine the group segmentation, coming close to the ideal situation in which each group comprises a single consumer. Thus, first-degree price discrimination can be seen as an extreme form of group pricing.

First- and third-degree price discrimination rely on the existence of observable and verifiable indicators of consumers' willingness to pay. When it is not possible to identify consumer groups with similar levels of willingness to pay, the only opportunity for offering different prices to different consumers is to propose to all consumers the same menu of packages (i.e., some combination of price and product characteristics), among which consumers self-select. This practice is

[2] For a more detailed overview, we refer to Varian (1989) or Belleflamme and Peitz (2015).

known as *second-degree price discrimination*, nonlinear pricing, menu pricing, or versioning.

Successful price discrimination requires the discriminating firm to be able to segment their customers according to their different price elasticities of demand for goods or services.[3] In addition, the firm must either be able to prevent or exclude arbitrage due to certain product characteristics (e.g., if the product or service has to be consumed immediately). If these conditions are satisfied, the firm is able to increase its revenue by using discriminatory pricing strategies.[4] It is not possible, however, to charge every consumer a different price as there is a smallest monetary unit, and therefore, only a limited number of possible prices; consequently, there will always be some consumers who pay the same price as others.

In this paper, we focus on third-degree price discrimination. We define different customer groups, wherein each group has different sensitivities, i.e., different acceptance probabilities for the price bids of the firm. Among others, Ezrachi and Stucke (2016) argue that with an increasing use of big data, learning algorithms are able to differentiate and segment customers into ever-smaller reference groups who have similar price sensitivities and purchase behaviors and who share common biases and levels of willpower. According to the authors, pricing algorithms can use data on how other people in an individual's group react in order to predict the individual's reaction under similar circumstances. This method enables the algorithm to adjust prices for products and services according to the estimated willingness to pay. Consequently, the more data the algorithm obtains, the closer those personalized prices can be set to a customer's reservation price.

3.2.2 Real-World Price Discrimination

In recent years, we have witnessed a number of examples where price discrimination has occurred in real-world situations. The main reason is easier access to data and the possibility of personalizing offers and prices. Firms are able to create detailed profiles of their customers; this can be facilitated by customers either identifying themselves or by identifying them through static Internet addresses, credit card numbers, cookies or various caching methods (Cahn et al., 2016). Hypertext Transfer Protocols (HTTP) allow servers to set and read cookies that store unique identifiers or information about a transaction. These cookies persist even after the session ends, so the next time the user accesses the server with the

[3] Federal Trade Commission and U.S. Department of Justice, *Horizontal Merger Guidelines.*

[4] It should also be noted that a unit price can be a special case of price discrimination if every consumer's willingness to pay is exactly the same.

same account, the server can identify that user and retrieve the stored data, which can be matched with details of previous interactions (Acquisti and Varian, 2005). Information about identified regular customers, postal codes, or type of operating system may then be used to segment customers and apply personalization mechanisms such as targeted advertising, product and service recommendations, and personalized pricing.

As an increasing amount of commerce takes place on online platforms, there is also the advantage that sellers can more easily offer different prices, product combinations, and recommendations to different customers (or groups) or, alternatively, quote a flat price but offer customers individualized discounts and bonus schemes. In particular, such discounts are not limited to online marketplaces. Before they were taken over by the supermarket corporation Edeka, customers of the supermarket chain Kaiser's received individualized discounts depending on their purchase history. However, the results were not really customized.[5] Another use-case is the one of the Swedish furniture retailer IKEA, reported by the *Harvard Business Review*.[6] In 2020, the company launched a temporary initiative at its store in Dubai that allowed its customers to pay different prices for its products according to the time they spent driving to the store. Thereby, product prices were expressed in two components: a fixed shelf price and a variable component dependent on the travel time. When showing their Google Maps Timeline to the cashier, an algorithm would calculate the value of the drive by taking into account time spent, distance traveled, and the average hourly wage of a Dubai worker. The higher this value, the less the customer had to pay for the item. Similar approaches can also be found in online retailing. A study by ProPublica, a nonprofit organization in the United States, revealed that *The Princeton Review* based the pricing for their online tutorial courses on information about customers' ZIP codes.[7] This practice was feasible as charging different prices to customers in different geographic regions is regulated in Europe but not in the United States. Since ZIP codes in the United States can often serve as a proxy for the ethnicity of the majority of residents, this resulted in higher prices being charged for some ethnic groups than for others. For example, in some regions where the population is dominated by people of Asian descent, prices for the online tutorial courses were 10%–20% higher compared to prices offered to

[5] Der Preis ist heiß, *Wirtschaftswoche*, March 02, 2017.

[6] The Pitfalls of Pricing Algorithms, in: *Harvard Business Review*, September-October 2021.

[7] The Tiger Mom Tax: Asians Are Nearly Twice as Likely to Get a Higher Price from Princeton Review, `https://www.propublica.org/article/asians-nearly-twice-as-likely-to-get-higher-price-from-princeton-review`, September 1, 2015.

customers of the same descent who live in regions with other ZIP codes where another ethnic group was the majority. Another case of price discrimination based on ZIP codes was reported by the *Wall Street Journal*.[8] Staples, a US retailer of office supplies, has differentiated between high and low prices on its e-commerce website depending on the user's ZIP code. The *Wall Street Journal* reported that customers living close to a direct rival brick-and-mortar store were offered discounted prices. As a consequence, customers living in rural areas were charged higher prices than those living in urban areas.

Furthermore, not only personalized prices but also the personalization of product rankings displayed to consumers may occur. This practice is known as price steering, which occurs when two users make the same query and receive different search results or the same results but in a different order. There is evidence that the ranking of search results influences click-through rates (Ghose and Yang, 2009). Since products or search results that are displayed at the top of lists are frequently clicked by consumers, personal recommendations that are placed at the top can be an effective tool for price discrimination. The online travel agent Travelocity has taken advantage of this by showing users of a Safari web browser different search results than Chrome users are shown. In doing so, they have changed the order of the search results and shown users of Apple's operating system a price that was about 5% lower (Hannak et al., 2014). Interestingly, in addition to this example of Apple users being favored, there was another example of price steering where the opposite was true. This example involves the travel agent Orbitz, which placed expensive offers at higher ranks in the search results of Mac users compared to users of a Windows system because Orbitz determined that owners of Mac computers spend as much as 30% more a night on hotels.[9] After the case was made public, Orbitz stopped using personalization algorithms.

One case that also attracted wide attention and caused a storm of protest was Amazon's attempt at differential pricing, described in the *Washington Post* in 2000.[10] Customers discovered that the online retailer varied prices charged for DVDs depending on the particular customer's frequency of purchases. For example, one user noticed that, after deleting cookies that identified him as a regular customer, the price of a DVD in his Amazon shopping cart dropped from $26.24 to $22.74. The online retailer denied that this practice was a discriminatory pricing strategy, instead describing it as a random "price test." Moreover, the company refunded all customers who had paid the higher price (Cavallo, 2018).

[8] Websites Vary Prices, Deals Based on Users' Information, in: *Wall Street Journal*, December 24, 2012.

[9] On Orbitz, Mac Users Steered to Pricier Hotels, in: *Wall Street Journal*, August 23, 2012.

[10] On the Web, Price Tags Blur, in: *Washington Post*, September 27, 2000

After the public outcry, Amazon stopped charging different prices for its DVDs and since then has served as an example of the reaction of customers to detected price discrimination.

To summarize, the advent of the Internet and electronic commerce has provided the opportunity for sellers to move from a fixed-price strategy for their goods and services toward a dynamic pricing model or even engage in differential pricing. Reasons for this development are that transaction costs are reduced since buyers and sellers do not have to be physically present in the same place at the same time to participate in the market and realize a transaction. In addition, the opportunity of selling products online has dramatically increased not only the number of potential customers but also the number of competitors, resulting in a higher degree of uncertainty regarding other sellers' prices and demand volatility. As a consequence, using a single fixed price in volatile online markets may be inefficient and ineffective for retailers. Instead, retailers may choose to vary their prices, thereby considering two dimensions of price variation. First, prices can vary over time, i.e., prices may be set differently at different times of the day, week, or year. Second, the retailer may charge different customers and/or groups of customers different prices, which is the aforementioned price discrimination. Retailers now have to decide along these two dimensions when setting their prices. In the simulations provided in this study, we only focus on the latter dimension.

3.2.3 Algorithmic Price Discrimination

Nearly 20 years ago, a paper was published by Raju et al. (2003), who solved dynamic pricing problems using reinforcement learning techniques such as Q-learning and actor-critic algorithms. In particular, the authors discussed how sellers can use these methods of automated pricing agents, which are also known as price bots, to determine revenue-maximizing prices. Since then, however, companies' technical capabilities have expanded even further. The scientific studies presented below have examined the ability of pricing algorithms to provide methods for calculating the best (i.e., profit-maximizing) prices for individual consumers or groups of consumers. These studies attempted to acquire a more precise estimate of more narrowly defined consumer groups' willingness to pay. To calculate such estimates, algorithms use not only demographic data but also information about consumers' browsing behaviors, purchasing histories, and preferences derived from their online activities. Analyzing this information may enable firms to obtain a more precise estimate of those consumers' willingness to pay and to better estimate purchase probabilities through machine learning or econometric models.

Shiller (2013) proposed a model predicting the probability that a consumer subscribes to the streaming service provider Netflix. The probability of subscribing was derived by running a probit regression model. The author used a data set including both demographic and web behavioral variables. The simulations showed that the increase in profits due to personalizing prices was relatively small when based on demographic variables alone (0.8%). Conversely, there was a 12.2% increase in profits when web-browsing behavior, such as frequency of website visits and time spent browsing, were used to predict the consumers' willingness to pay, and prices were charged accordingly. In this latter case, the price range offered to different consumers was significantly large, with some consumers even having to pay almost double the price paid by others.

Another model for implementing algorithmic price discrimination was developed by Chen et al. (2022). In their model, an algorithm making joint pricing and assortment decisions based his action selection on a particular set of customer features. After an action was chosen, the algorithm observed whether the product was purchased by the customer or not. By repeating this process, the algorithm learned to select the action that yields the highest revenue with regard to the specific set of customer features. The interdependencies between customer features, action selection and the resulting outcome, i.e., the purchasing decision, were learned using a logit regression model, whereby the probability of a customer choosing a product was dependent on the selected action and the customer's attributes. The expected revenue obtained by the seller was calculated using the estimated probabilities. By using such a pricing algorithm, the authors showed that the seller was able to derive the customers' reservation value for the products offered and set prices accordingly to maximize his profit. Furthermore, the results of experimental simulations were transferred to the real world using data from a European airline carrier on sales of a priority-seating option. The results have indicated that the model proposed by Chen et al. (2022) for determining customers' willingness to pay and then setting prices accordingly resulted in higher revenue compared to a uniform pricing policy.

In a further study, Ban and Keskin (2020) based their personalized pricing model on similar assumptions. The algorithm again learned the effect of customer characteristics on demand for the product offered and used this information to set revenue-maximizing prices. The interaction effects between demand and pricing policy were modeled by a lasso regression. The authors were able to show that their experimentation-based personalized pricing strategy is superior to other approaches such as myopic pricing and segment-then-optimize policies, which are comparable to group pricing based on clusters of customers.

Another study dealing with algorithmic price discrimination was conducted by Dubé and Misra (2021). The authors assumed that customers' individual price sensitivity can be characterized by a vector of observable customer attributes. With the application of machine learning tools, sellers should be able to learn statistically about the demands of heterogeneous customers in order to set personalized prices. For this purpose, the authors used a lasso regression model to identify those observable customer features that have a significant impact on demand. Afterward, they approximated the supplier's demand uncertainty using a logistic regression model. The model was then applied by evaluating business-to-business price experiments for potential new customers of a recruiting firm. In a first experiment, the algorithm was trained on experimental data in order to learn the relationship between price sensitivity and the observed features of customers. In a second experiment, the trained algorithm was applied to a new set of customers, and the pricing recommendations of the algorithms were evaluated against the firm's status-quo pricing. The model showed that personalized pricing increased a firm's expected profits by 19% relative to an optimized uniform pricing strategy designed by the authors and by 86% relative to the firm's status-quo pricing. The authors also considered the impact of pricing strategies on consumer surplus. In doing so, they found that personalized pricing would reduce total consumer surplus. However, at an individual level, the majority of customers would benefit from being charged prices lower than the uniform price.

3.2.4 Fairness

When it comes to differential pricing, the fairness debate is not far behind. Consumers' notion of fairness have been a discussion point in economics since the works of Rabin (1993), Fehr and Schmidt (1999), and Bolton and Ockenfels (2000), if not earlier. A number of experiments have demonstrated that economic agents care not only about their own payoffs but also about the payoffs of others, and they act accordingly. Fehr and Schmidt (1999) modeled fairness as inequity aversion, which means that economic agents resist inequitable outcomes and are willing to forgo part of their payoff to achieve a more equitable outcome. Thus, fairness or unfairness with regard to prices is evaluated by a comparison to prices offered to other consumers under similar circumstances. The approach of Fehr and Schmidt (1999) also includes the aspect of self-centered inequity aversion, which means that consumers suffer more from inequality that is to their disadvantage (they have to pay a higher price) than from inequality that is to their advantage (they have to pay a lower price). A way to measure the fairness of a given distribution is provided by Jain et al. (1984), who took into consideration consumers' preferences for equality, which favors similarly distributed, i.e., homogeneous, prices. The proposed fairness index is represented by the following equation:

$$f(x) = \frac{\left[\sum x_i\right]^2}{n \sum x_i^2} \in [0, 1], \tag{3.1}$$

where f measures the fairness if resources are allocated to n individuals such that the i^{th} individual receives an allocation x_i. If customer groups rather than individuals are considered, n can be interpreted as the number of customer groups. The variable x_i can represent either the price that is charged or the quantity that is sold to a particular customer group i. The index measures the equality of the allocation x: If all customers receive the same amount or pay the same price, i.e., all x_i's are equal, then the index is 1, and the policy is 100% fair. If disparity increases, the outcome is perceived as unfair; a policy that favors only a few selected customers has a fairness index near 0. The proposed index is dimensionless and independent of scale; it is bounded between 0 and 1; and it is continuous, so that any slight change in x_i changes the index.

Previous attempts have been made to combine the aforementioned aspects of personalized pricing and fairness considerations when solving dynamic pricing problems by using reinforcement learning techniques. Examples for these are the works by Maestre et al. (2018) and Cohen et al. (2021). The present work also contributes to this strand of literature. Maestre et al. (2018) considered the topic of dynamic pricing with demand learning and integrated the fairness aspect into their model using Jain's index as a metric. The authors demonstrated that reinforcement learning algorithms have the ability to adapt pricing policies that consider fairness but, at the same time, maintain optimization of revenue. Since the authors considered fairness part of the revenue function of firms, fairness would be taken into consideration only if firms explicitly wanted to do so. In contrast, our approach considers fairness part of the demand side as we assume a lower propensity to buy if customers note unfair price-setting behavior by the respective firm. A further difference is the algorithm that is employed. Maestre et al. (2018) applied neural networks to Q-learning, also known as deep Q-learning, while we show that even simple Q-learning algorithms are able to incorporate fairness considerations with a manageable amount of computing capacity.

Cohen et al. (2021) studied dynamic pricing with unknown demand under price fairness constraints. Similar to the work conducted by Raju et al. (2003), they considered prices as fair if they were similar for different customer groups (group fairness) and if prices were stable over time for each customer group (time fairness). By using an infrequently changed upper-confidence-bound algorithm, the authors showed that imposing group fairness does not affect the demand learning problem in contrast to imposing time fairness. By contrast, results have revealed that imposing time fairness, unlike imposing group fairness, does not affect the optimal revenue.

3.2.5 Reinforcement Learning

As this study examines price discrimination and fairness considerations through learning algorithms, we introduce reinforcement learning as a basic machine learning concept. In this method, the algorithm learns how to map situations to actions in order to maximize a numerical reward signal. The algorithm (or agent) is not explicitly told what actions to choose but instead has to learn which actions yield the highest reward through trial. Reinforcement learning uses the formal framework of Markov decision processes to define the interaction between a learning agent and its environment in terms of states, actions, and rewards. In the following paragraphs, we give a short summary of the most important features of this type of machine learning, which are discussed in further detail by Sutton and Barto (2018).

One of the challenges that arises in reinforcement learning is the trade-off between the strategies of exploration and exploitation. To obtain a high reward, an agent with reinforcement learning exploits what it has learned so far by preferring actions it has already tried and that have led to a high reward in the past. However, to discover such actions, it has to try (or explore) actions that have not been chosen before. Therefore, the algorithm should use a dynamic action selection policy that balances exploitation, that is choosing the optimal action as currently perceived, and exploration, that is choosing another action to improve action selections in the future.

Independent Q-learning (Watkins and Dayan, 1992) is a simple but well-established reinforcement learning algorithm. By interacting with its environment, the algorithm learns to maximize a reward according to the Q-function $Q(s, a)$ that matches the optimal long-run value of choosing any action $a \in \mathcal{A}$ when faced with any given state $s \in \mathcal{S}$. During this interaction, the algorithm uses the above-mentioned dynamic action selection policy, which balances actions exploiting what has been previously learned with those exploring what has not been tried before. The Q-function can be represented as a $|\mathcal{S}| \times |\mathcal{A}|$ matrix. If the Q-matrix is known, the algorithm can easily choose the optimal action for any given state. However, as this matrix is unknown, the Q-learning algorithm first has to estimate the values of the Q-matrix through an iterative procedure without knowing the underlying model. Starting from an arbitrary initial matrix Q_0, the algorithm chooses action a_t in state s_t, observes reward r_t and subsequent state s_{t+1}, and updates the corresponding cell of the matrix $Q(s_t, a_t)$ according to the following recursive relationship:

$$Q(s_t, a_t) \leftarrow (1-\alpha)Q(s_t, a_t) + \alpha\left(r_t + \delta \max_a Q(s_{t+1}, a)\right), \tag{3.2}$$

where the updated value $Q(s_t, a_t)$ is a convex combination of the previous value $Q(s_t, a_t)$ and the reward obtained after performing action a_t in state s_t plus the discounted value of the state that is reached in the following period. For all other cells of the matrix, the Q-value does not change. The parameter $\alpha \in (0, 1]$ is the learning rate that regulates how quickly new information replaces old information, and $\delta \in [0, 1)$ is the discount factor. Action a denotes the optimal strategy (i.e., the action leading to the highest reward) until this time step.

To balance exploration and exploitation, the Q-learning algorithm adopts a probabilistic action selection policy according to equation (3.3).

$$a_t = \begin{cases} [a_{min}, a_{max}] & \text{with probability } \varepsilon_t \\ \arg\max_a Q(s_t, a) & \text{with probability } 1-\varepsilon_t \end{cases} \tag{3.3}$$

Using what is called a ε-greedy strategy, the algorithm follows a random action (exploration) within a given interval $[a_{min}, a_{max}]$ with $\varepsilon_t \in [0, 1]$ probability and exploitative action with $1-\varepsilon_t$ probability. If several actions yield the same highest Q-value under the exploitation approach, the algorithm randomizes across these actions.[11]

The probability of exploration is determined by:

$$\varepsilon_t = \varepsilon_0(1-\theta)^t,$$

where $\varepsilon_0 \in [0, 1]$ is the initial exploration probability and $\theta \in [0, 1]$ is a decay parameter. Whenever $\theta > 0$, the decay in exploration ensures convergence to a deterministic strategy.

Although other more complex independent-learning algorithms are currently being applied to diverse strategic settings, this paper focuses on a simple Q-learning algorithm. In contrast to more sophisticated algorithms, this simple algorithm can be fully described by just two parameters: the learning rate α and the decay parameter θ. For more complex algorithms such as deep learning

[11] This approach is based on the textbook of Sutton and Barto (2018) and is also used by Calvano et al. (2020) and Klein (2021).

algorithms, a number of additional parameters must be specified. Apart from learning and experimentation parameters, the functional form of the Q-function, the number of estimation layers, and the structure of the neural network in each layer has to be determined. A disadvantage of the Q-learning algorithm is the slowness of learning, which increases with the complexity of the environment. However, since the environment in this paper is very simple, Q-learning provides an adequate approach for this problem.

3.3 Methodology

Customer groups g_i with $i = 1, 2, \ldots, n$ are defined in advance. This can be interpreted as an algorithm that divides customers into groups based on collected and analyzed data. The characteristics used to classify customers must have an effect on their willingness to pay. For the sake of simplicity, customer groups in this paper are exogenous. Therefore, the processes of defining customer groups, allocating customers, and setting prices are sequential.

Furthermore, we assume that customers are equally uniformly distributed in the interval $[0, 1]$; thus, we assume a unit mass of customers and unit demand. We define two customer groups $i = 1, 2$ which can be expanded in later simulations. Every customer belongs to exactly one group (either g_1 or g_2), and each customer accepts the price charged by the algorithm with probability $\phi \in [0, 1]$. This acceptance probability is given by the following function:

$$\phi_i(a) = \left[1 + e^{-(v_i + w_i \cdot a)}\right]^{-1}, \tag{3.4}$$

where a is the action (i.e., the price) chosen by the algorithm, and v_i and w_i are parameters defining the sensitivity of each customer group, defined in Table 3.1.

Group	v	w
1	18.229	−23.690
2	4.4757	−15.526

TABLE 3.1: Parameters of the Probability Functions

The larger the parameter v is, the higher the acceptance probability of the customer group. Parameter w determines how strong and in which direction (positive or negative) the customer reacts to a change in the price. If $w > 0$, the customers' acceptance rate increases with the price. This happens when customers judge quality by price, for example. If $w < 0$, the acceptance rate decreases when the price increases.

We provide a scenario simulating two different types of customer behavior. Customers in g_1 accept much higher prices than customers in g_2. Both groups show a decrease in the acceptance probability if the price increases. The acceptance probability functions of the two customer groups for the price interval $[0, 1]$ are shown in Figure 3.1.

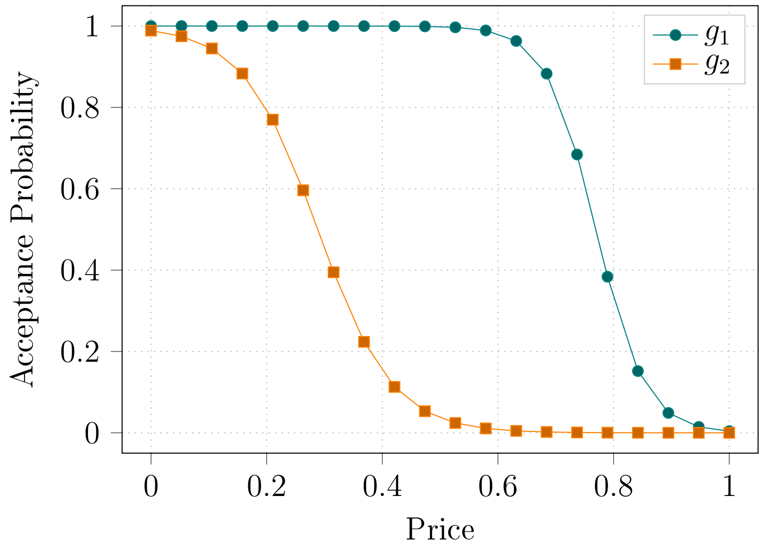

FIGURE 3.1: Acceptance Probability Functions

The action space is defined as an interval for choosing action $a \in [a_{min}, a_{max}]$, i.e., the price. Since Q-learning requires a finite action space, action a is a discrete variable scaled between 0 and 1 with k equally-sized intervals; thus, actions are taken from a discrete set $\mathcal{A} = \{0, \frac{1}{k}, \frac{2}{k}, \ldots, 1\}$. Note that, as long as the action set $\mathcal{A}$ is constrained between a_{min} and a_{max}, the algorithm will never set "non-sense" prices.

The state space of group i is given by all prices that can be set for group j and vice versa. Consequently, the state space is given by $\mathcal{S} = \{0, \frac{1}{k}, \frac{2}{k}, \ldots, 1\}$.

Whether a price distribution is considered fair is measured by the following fairness index:

$$f(a) = \frac{\left[\sum_i a_{max} - a_i\right]^2}{n \sum_i (a_{max} - a_i)^2} \in [0, 1] \,, \tag{3.5}$$

where a_{max} is the maximum price that can be set (i.e., the maximum price of the given action set), a_i is the price set for group g_i, and n is the number of groups. This index is based on the index of Jain et al. (1984) shown in equation (3.1) and was modified as proposed by Maestre et al. (2018). Using the original index, an allocation x is perceived as fair, and thus increases the index, if all x_i's are large and homogeneous distributed. However, in a differential pricing context, higher prices are not perceived as fair. Therefore, the original index was modified as presented in equation (3.5) to guarantee that lower, homogeneously-distributed prices lead to a higher fairness index.

In order to consider the customers' sense of fairness, the fairness parameter shown in equation (3.5) is integrated into the acceptance probability function such that a low fairness index decreases the acceptance probability and vice versa. The modified acceptance probability is given as:

$$\phi(a) = \left[1 + e^{-(v_i + w_i \cdot a) + \beta(1 - f(a))}\right]^{-1}, \tag{3.6}$$

where β denotes a parameter for considering fairness ($\beta = 1$) or not considering fairness ($\beta = 0$). The modified acceptance probabilities are shown in Figure 3.2 for different values of the fairness parameter f. Note that customers' aversion to higher prices is captured in both the acceptance probability function (as the acceptance probability decreases with a higher price) and the fairness index (f is larger for lower, homogeneously-distributed prices). To determine how fairness develops over time, it is necessary to capture this feature as a separate aspect in the form of a fairness index. Since the fairness aspect is modeled as a part of the acceptance probability function, the algorithm has to learn how important fairness is for customers in order to be able to set the revenue-maximizing prices.

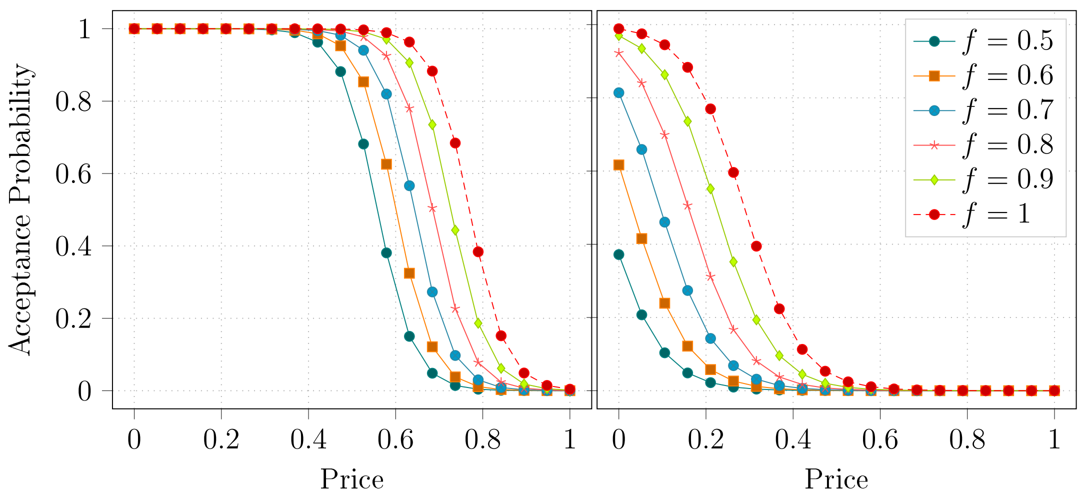

FIGURE 3.2: Acceptance Probability Functions of Customer Groups g_1 (left) and g_2 (right) for Different Fairness Levels

As customers are equally uniformly distributed within each group g_i, the reward r_i – i.e., the expected revenue generated by the sale to group i – is defined in equation (3.7) as the price multiplied by the acceptance probability from equation (3.6), where $i = 1, 2$ denotes the respective customer group:

$$r_i = a_i \cdot \phi_i \, . \tag{3.7}$$

The Q-learning is simulated as follows. In $t = 1$, the algorithm sets a price for each group. For all subsequent periods, the algorithm sets prices alternately, wherein the last price set for one group serves as the current state for the other group. Additionally, this price serves as the comparative price for the fairness index. Depending on the selected parameter values, the reward is either independent of the state ($\beta = 0$, if inequity aversion is not considered in the model) or dependent on the current state ($\beta = 1$, if customers are inequity-averse). The algorithm learns to set the optimal price for each group by updating its Q-function for customer group i according to the following recursive relationship:

$$Q_i(a_j, a_i) \leftarrow (1-\alpha)Q_i(a_j, a_i) + \alpha \left(r_i + \delta r'_j + \delta^2 \max_a Q_i(a'_j, a) \right), \tag{3.8}$$

where action a_i denotes the price set for group i and a_j denotes the price set for group j, which determines the state for the Q-function of customer group i. The updated value $Q_i(a_j, a_i)$ is a combination of the previous value $Q_i(a_j, a_i)$, the reward obtained for group i, the discounted reward r'_j for group j that will be achieved in the following period, and the discounted value of the new state a'_j that will be reached in the following period.

To constrain the number of possible actions, we use the discrete action set $\mathcal{A} = \{0, \frac{1}{k}, \frac{2}{k}, \ldots, 1\}$ with $k = 10$. Consequently, there are a total of eleven possible prices. We consider two customer groups, making $i = 1, 2$. We take an initial exploration probability $\varepsilon_0 = 1$, decay parameter $\theta = 6.908 \cdot 10^{-5}$, learning rate $\alpha = 0.1$, and discount factor $\delta = 0.9$. To assess the performance of the algorithm, the statistics are computed over 100 runs. In each run, $100,000$ price choices (periods) are simulated. Over these $100,000$ periods, the exploration probability drops to below 0.1%. To compensate for outliers, we average over the 100 simulated runs for each period to see how an average market price and reward develop over time.

The pseudocode for the Q-learning algorithm is provided below.

Pseudocode Q-learning
Set demand and learning parameters
Initialize Q_1 and Q_2 as empty matrices
Initialize $a_{1,1}$ and $a_{2,1}$ randomly
Initialize $t = 2$, $i = 1$ and $j = 2$
Loop over each period
| Set action $a_{i,t}$ according to (3.3)
| Calculate reward according to (3.7)
| Update $Q_i(a_j, a_i)$ according to (3.8)
| Update $t \leftarrow t + 1$
Until $t = T$ (specified number of periods)

3.4 Results and Discussion

Two different scenarios are conducted in order to compare the outcomes of these two different approaches. In the first scenario, the algorithm learns to set different prices for each customer group. We do not include inequity aversion in our model in this case, so the parameter β is set equal to zero. Therefore, the acceptance probability simplifies to equation (3.4). The reward is calculated according to equation (3.7). Since the fairness index has been dropped, the reward of a chosen action does not depend on the current state (i.e., the price charged to the other group).[12] In the second scenario, inequity aversion is considered ($\beta = 1$); thus, prices a_i and a_j depend on each other, due to the fairness index. Customers are assumed to be inequity-averse, i.e., they have a strong preference for equitable, homogeneous prices. As a result, the acceptance probability decreases with a decreasing value of the fairness parameter. The reward is calculated the same way as in scenario I, with the only difference being that the acceptance probability now depends on the fairness parameter and is calculated according to equation (3.6). Consequently, the algorithm learns to increase the price within each group and thereby maximizing the expected revenue (as observed in the first scenario) while reducing price differences between groups, in order to maximize revenue and fairness at the same time. It is assumed that maintaining a balance of price distribution among the different groups of customers leads to fairer prices, as these prices are chosen by considering customers' preference for equality.

[12] Note that this simplified setting is a one-stationary problem, which could be solved more efficiently by a multi-armed bandit algorithm; however, for a better comparison of both scenarios, the Q-learning algorithm is also applied in this case.

To examine whether this statement is true, we display the price histories for both scenarios in Figures 3.3 and 3.4, whereby prices are calculated as the averages over all runs. The corresponding reward histories for both scenarios are shown in Appendix 3.A in Figures 3.A.1 and 3.A.2.

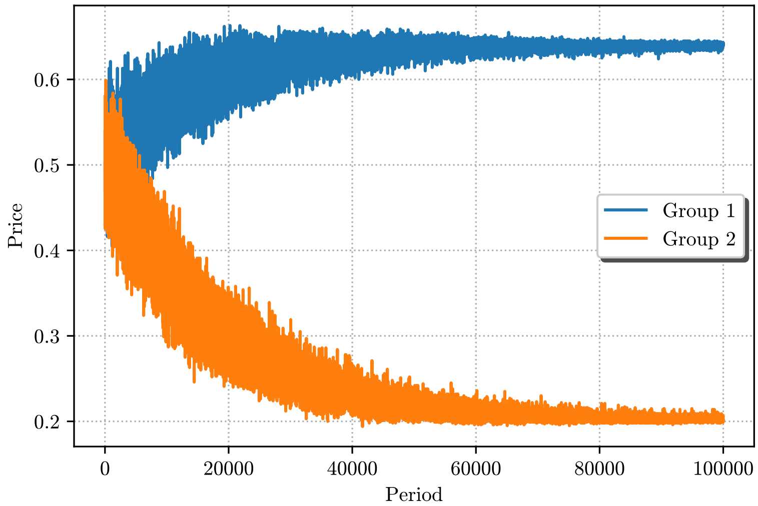

FIGURE 3.3: Price Histories Scenario I

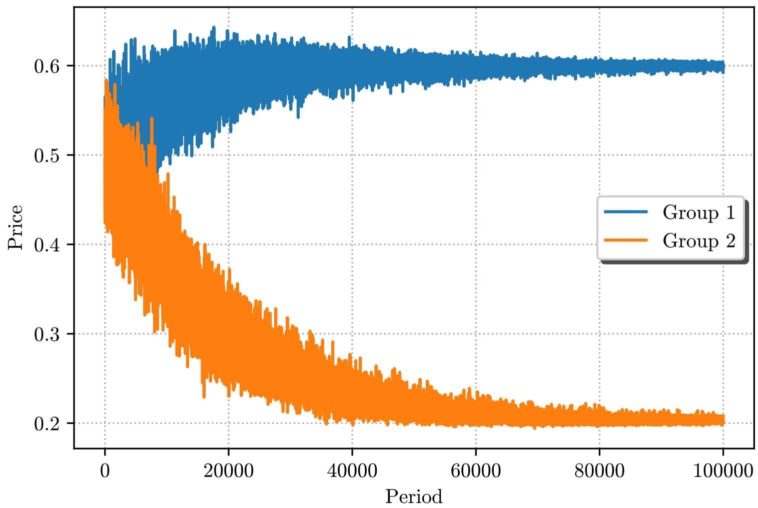

FIGURE 3.4: Price Histories Scenario II

It can be observed that the algorithm starts on average at a common price level for both groups in the first periods, but then learns very quickly that a price differentiation between customer groups leads to a higher reward. After about 60,000 periods, prices converge to the final level. In the second scenario, the algorithm adjusts the price for group 1 back down slightly to achieve a more homogeneous distribution of prices. As a consequence, rewards in both scenarios increase continuously and stabilize at the levels of 0.747 and 0.744, respectively.

Table 3.2 shows the pricing strategies the algorithm has learned, as captured by the average over the final 1,000 periods of all runs. In these periods, the algorithm barely explores but instead exploits what it has learned thus far. Moreover, the resulting fairness index and the reward generated at these prices are displayed. In the first scenario, the Q-learning agent is able to conduct price discrimination and does not consider inequity aversion. Accordingly, the value of the fairness parameter is lower than 1, since customers are charged different prices. If inequity aversion is introduced in the second scenario, the price for customer group 2 remains nearly unchanged, while the price charged to group 1 decreases. As a result, prices are more homogeneously distributed (however, there is no unit price), since the price-setting policy balances between maximizing revenue and providing a fair market outcome. This is further illustrated by a higher value of the fairness parameter in scenario II compared to scenario I. The higher reward is obtained in scenario I, wherein prices are set on a discriminatory basis without considering inequity aversion.

	$\bar{a}_1$	$\bar{a}_2$	$\bar{r}$	$\bar{f}$
Scenario I	0.640	0.201	0.747	0.872
Scenario II	0.600	0.201	0.744	0.900

TABLE 3.2: Average Market Outcomes of the Final 1,000 Periods

To assess the performance of the algorithm, we compare the results of the algorithmic learning procedure with the analytical solution of the profit maximization problem displayed in Table 3.3. Since we assume zero marginal cost, profit equals revenue in this case, which is equal to the reward that the algorithm tries to maximize. If inequity aversion is not considered in the model, prices for both customer groups are higher compared to scenario II wherein inequity aversion is considered. Thus, the firm is able to extract more consumer surplus from both groups in this case.

	a_1^*	a_2^*	r^*
Scenario I	0.656	0.228	0.778
Scenario II	0.649	0.225	0.767

TABLE 3.3: Analytical Solution

The comparison of simulated and analytical market outcomes displayed in Tables 3.2 and 3.3 shows that the algorithm charges prices slightly below the optimal (i.e., the profit-maximizing) levels. This may be either because we restrict the action set to numbers with one decimal place (in the case of $k = 10$), leading to the fact that the simulation results are not able to converge to the exact numbers of the analytical solution, or because the algorithm does not have sufficient time to learn in the provided setting. Nevertheless, the algorithm is able to take into account customers' fairness considerations and realize a revenue that is only slightly below the calculated value of the analytical solution of the profit maximization problem, despite this more difficult relationship.

Certainly, the simulation setting is very stylized, however, the assumptions made above do reflect economic reality at least to some extent. For example, in a perfectly competitive market, price discrimination is not feasible, as firms undercut each other until they end up in a (unit) price that equals marginal cost. However, prices might be differentiated due to search costs, lock-in effects, transportation costs, or asymmetric information. As a result, firms compete with each other, but real markets do not seem to be perfectly competitive, and differential prices are likely to occur. Much empirical work tests for the presence of price discrimination in imperfectly competitive environments, e.g. Shepard (1991) in the gasoline market, Goldberg (1995) in the market for European automobiles, and Leslie (2004) in the Broadway theater market.[13]

Another assumption that applies in reality is that of a limited action space: since firms have a price limit under which they would never sell the product (typically under average or marginal cost) the price interval has a lower bound, and the upper bound of the price interval is potential customers' maximum willingness to pay. Moreover, the lower bound is ensured by law: Despite the willingness to obtain negative profits in particular situations, retailers in the European Union are not allowed to sell below cost (called predatory pricing) according to the Treaty on the Functioning of the European Union (TFEU), Art. 102. The reason given is that such behavior, exercised over a longer period of time, can drive smaller competitors out of the market. With regard to the upper bound of the price interval, it should be mentioned that each customer's actual willingness to pay cannot be determined by firms, as they only obtain imperfect information about each customer. For this reason, personalized prices in real markets are generally not identical to perfect price discrimination.

Additionally, the implications of personalized pricing for consumer and total welfare are ambiguous, as welfare could move in many directions relative to the benchmark of a unified market. Bergemann et al. (2015) showed that while obtaining additional information about customers can never hurt the seller, it

[13] For a survey of price discrimination in imperfectly competitive markets, see Armstrong (2006).

can increase both total and consumer surplus, decrease both, or increase one and decrease the other. In the study of Dubé and Misra (2021) mentioned in Section 3.2.3, the authors conducted an experiment with an online recruiting company, comparing the existing uniform price, an optimized uniform price, and targeted prices. They found that customer surplus declines slightly with personalized pricing relative to uniform pricing. However, over 60% of customers benefit from personalized prices since they had to pay a lower price than the optimal uniform price. Nevertheless, the distribution of welfare in cases of price discrimination must be assessed on a case-by-case basis, as no general conclusions can be drawn.

3.5 Robustness Checks

To check robustness of our results, we consider various changes in the parameters that are used for defining the algorithm and the environment. We run simulations with changes in (1) parameter k to allow for a larger action space, (2) learning rate α, (3) discount parameter δ, (4) decay parameter θ, and (5) price sensitivities v_i and w_i for customer groups $i = 1, 2$.

Choosing the discretization parameter k that determines the number of actions a within the action set $\mathcal{A}$ is a trade-off between, on the one hand, making the acceptance probability functions of the two customer groups more substantial and, on the other hand, negatively affecting the learning process since the Q-matrix increases significantly with the number of possible actions (and states), resulting in a much longer time needed to learn the optimal pricing strategy. Experimenting with various values for the parameter k, we find $k = 10$ providing a good balance between expanding and reducing the number of possible actions and, therefore, being appropriate for our setting. For example, for the value $k = 100$ we observe significantly higher average prices charged to both groups, however, the reward is significantly lower. This is valid in both scenarios. This leads us to conclude that the algorithm has not yet finished its learning process after the $100,000$ periods. Please note that the optimal value of k changes if simulation settings, such as the number of periods per run, changes.

The learning parameter α, which determines how quickly new information replaces old information, may range from 0 to 1. High values of α indicate extensive experimentation, as the algorithm forgets too quickly what it has learned in the past. Thus, values of α close to 1 may disrupt the learning process. To be effective, the learning process has to be persistent, which requires the learning rate to be very small, i.e., close to zero.[14] Accordingly, in our baseline simulation,

[14] In computer science literature, a value of $\alpha = 0.1$ is common.

we set $\alpha = 0.1$, which is also in line with the findings of Calvano et al. (2020). In contrast, in his study, Klein (2021) used a parameter value of $\alpha = 0.3$. Changing the value of α to this number as a further robustness check does not effect our results significantly.

The discount factor of $\delta = 0.9$ was chosen reasonably close to one (i.e., the future is not discounted so much) as the price setting intervals and, therefore, the periods, are generally small. In case of very short time periods, the actual discount factor would be even closer to 1. However, choosing a discount factor very close to 1 in the algorithmic learning environment, sufficient learning may not be possible because Q-values estimated in the past will get too much weight. When observing this problem, it may be required to set a lower value of δ. This issue is also discussed in more detail by Klein (2021).

Changes in the decay parameter θ lead to changes in the exploration probability since θ determines the speed of convergence to a deterministic strategy, i.e., how fast the probability of exploration decreases. Thereby, the higher the value of θ, the faster the algorithm converges to a deterministic strategy. For example, if we set $\theta = 9.210 \cdot 10^{-5}$ as suggested by Klein (2021), the exploration probability decreases to 0.01% at the end of the run. Thus, there is a trade-off between a higher probability that the algorithm will explore (and not exploit the previously learned best strategy) in the final periods and the risk that the algorithm converges to a deterministic strategy although the learning process might not have been finished at that point and, as a consequence, converged to an ineffective strategy. In our simulation, we choose a value of $\theta = 6.9075 \cdot 10^{-5}$ resulting in the exploration probability gradually decreasing from 100% in the first period to 3.16% halfway through the run and reaching 0.1% at the end of the run. What we observe in our robustness checks is that for lower values of θ, price variation increases since the exploration probability is quite high even in the final periods leading to prices being set randomly very high or very low in these periods. However, the average results over 100 runs do not vary significantly.

If we model the consumers' acceptance probability functions with different sensitivities, this has an impact on the difference in prices charged for the two customer groups as well as on the extent of the price adjustment in the second scenario: The more the acceptance probabilities of both groups diverge, the larger the price difference between the two groups in the first scenario. In the second scenario, prices are adjusted to a greater extent if acceptance probabilities are more diverged. If the difference in willingness to pay is sufficiently large, it is profit-maximizing for the firm to charge the same price to both groups of customers in scenario II in order to obtain the maximum revenue from group 1 having the higher willingness to pay, which is the highest for a fairness index of 1. In this case, the acceptance probability of customer group 2 is close to zero. The exact numbers of a robustness check with similar willingness to pay and a robustness check with a very high willingness to pay for group 1 and a very low willingness to pay for group 2 are provided in Appendix 3.B. It can be shown that the algorithm is also able to learn the respective profit-maximizing prices for different values of sensitivity parameters, showing that our findings are robust to such changes.

3.6 Conclusion

Using reinforcement learning, this paper demonstrates how a firm maximizes its revenue while taking into consideration customers' inequity aversion. If we assume that customers' acceptance probability decreases when they realize that they are being charged different prices, companies have to sacrifice part of their revenue compared to a scenario with customers who do not consider prices charged to others and equalize prices at least to some extent to ensure improved fairness between customer groups.

In the present simulation, we chose two customer groups with similar demand characteristics. Revenue maximization by the algorithm becomes incomparably more complex as more groups with more diverse demand functions are in the market. Considering additional customer groups would be a further attempt to make the environment more realistic, but solving this problem would require much more computing capacity.

Referring to the use of alternative algorithms, a comparison with other reinforcement learning methods could be useful for gaining additional insights. The complexity of the algorithm might be further increased by including the processes of defining and refining customer groups as well as setting prices simultaneously. Moreover, using different algorithms might allow for the investigation of multi-agent learning by simulating a competitive environment instead of a monopolistic structure. At this time, the outcomes of settings with multiple agents using different algorithms have hardly been studied.

Further extensions of this study might include a redefinition of the fairness index. For example, Jain's index does not account for self-centered inequity aversion, which occurs when economic agents do not care about inequity that is present among other economic agents and are interested only in the equity of their own payoff relative to the payoffs of others. Moreover, according to Fehr and Schmidt (1999), the judgment and feelings associated with advantaged and disadvantaged price inequality are different, because consumers suffer more from inequality that is to their disadvantage. These approaches, however, are not considered in the current index and are of particular interest if more than two groups of customers are considered.

A shortcoming of the employed algorithm is the relatively long time that is needed to converge to the optimal price. If the algorithm were employed in the real world, it would not have 100,000 trials in which to identify the optimal prices, as frequent price changes might lead to a loss of consumer confidence. A possible solution could be to train algorithms offline before they are deployed in the market to avoid upsetting customers by explorative pricing, which occurs more often in the first periods, and to shorten the learning process.

Apart from the aforementioned aspects, the applicability of personalized pricing in the real world depends on many factors, which may include both technical and legal restrictions on data collection and use as well as on technical limitations of the algorithms themselves. From the simulations, we can conclude that consumer characteristics such as their sense of fairness, which have prevented firms from engaging in price discrimination, can be incorporated into firms' pricing decisions with the help of learning algorithms, making differential pricing strategies more feasible. Policy debates should take these aspects into consideration when evaluating the benefits and risks associated with the use of pricing algorithms.

References

Acquisti, A. and H. R. Varian (2005). Conditioning prices on purchase history. *Marketing Science*, 24.3, pp. 367–381.

Armstrong, M. (2006). Recent developments in the economics of price discrimination. In: *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*. Ed. by R. Blundell, W. K. Newey, and T. Persson. 2nd ed. Econometric Society Monographs. Cambridge, United Kingdom: Cambridge University Press, pp. 97–141.

Ban, G.-Y. and N. B. Keskin (2020). Personalized dynamic pricing with machine learning: High dimensional features and heterogeneous elasticity. *Management Science*, 67.9, pp. 5549–5568.

Belleflamme, P. and M. Peitz (2015). *Industrial Organization: Markets and Strategies*. Cambridge, United Kingdom: Cambridge University Press.

Bergemann, D., B. Brooks, and S. Morris (2015). The limits of price discrimination. *American Economic Review*, 15.3, pp. 921–957.

Bolton, G. and A. Ockenfels (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90.1, pp. 166–193.

Bourreau, M. and A. de Streel (2018). The regulation of personalised pricing in the digital era. OECD Background Note. DOI: `10.2139/ssrn.3312158`.

Cahn, A., S. Alfeld, P. Barford, and S. Muthukrishnan (2016). An empirical study of web cookies. In: *Proceedings of the 25th International Conference on World Wide Web*. Association for Computing Machinery, pp. 891–901.

Calvano, E., G. Calzolari, V. Denicolò, and S. Pastorello (2020). Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, 110.10, pp. 3267–3329.

Cavallo, A. (2018). More Amazon effects: Online competition and pricing behaviors. Working Paper No. 25138. National Bureau of Economic Research. DOI: `10.3386/w25138`.

Chen, X., Z. Owen, C. Pixton, and D. Simchi-Levi (2022). A statistical learning approach to personalization in revenue management. *Management Science*, 68.3, pp. 1923–1937.

Cohen, M. C., S. Miao, and Y. Wang (2021). Dynamic pricing with fairness constraints. SSRN Working Paper. DOI: `10.2139/ssrn.3930622`.

Competition and Markets Authority (2018). Pricing algorithms: Economic working paper on the use of algorithms to facilitate collusion and personalised pricing. URL: `https://assets.publishing.service.gov.uk/government/`

uploads/system/uploads/attachment_data/file/746353/Algorithms_econ_report.pdf (visited on 06/29/2022).

Dubé, J.-P. and S. Misra (2021). Personalized pricing and customer welfare. Chicago Booth School of Business Working Paper. DOI: 10.2139/ssrn.2992257.

Executive Office of the President of the United States (2015). Big data and differential pricing. URL: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf (visited on 01/24/2020).

Ezrachi, A. and M. E. Stucke (2015). Artificial intelligence and collusion: When computers inhibit competition. Legal Studies Research Paper Series No. 267. DOI: 10.2139/ssrn.2591874.

— (2016). *Virtual Competition: the promise and perils of the algorithm-driven economy*. 2nd ed. Cambridge, Massachusetts: Harvard University Press.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, pp. 817–868.

Ghose, A. and S. Yang (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science*, 55.10, pp. 1605–1622.

Goldberg, P. K. (1995). Product differentiation and oligopoly in international markets: The case of the US automobile industry. *Econometrica*, 63.4, pp. 891–951.

Hannak, A., G. Soeller, D. Lazer, A. Mislove, and C. Wilson (2014). Measuring price discrimination and steering on e-commerce web sites. In: *Proceedings of the 2014 Conference on Internet Measurement Conference*, pp. 305–318.

Jain, R. K., D.-M. W. Chiu, and W. R. Hawe (1984). A quantitative measure of fairness and discrimination. Eastern Research Laboratory, Digital Equipment Corporation, 21. Hudson, Massachusetts.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, 76.4, pp. 728–741.

Klein, T. (2021). Autonomous algorithmic collusion: Q-learning under sequential pricing. *RAND Journal of Economics*, 52.3, pp. 538–599.

Leslie, P. (2004). Price discrimination in Broadway theater. *RAND Journal of Economics*, 35.3, pp. 520–541.

Maestre, R., J. Duque, R. Alberto, and J. Arevalo (2018). Reinforcement learning for fair dynamic pricing. In: *Intelligent Systems and Applications. Proceedings* *of SAI Intelligent Systems Conference*. Ed. by K. Arai, S. Kapoor, and R. Bhatia. Vol. 1. Springer International Publishing, pp. 120–135.

Odlyzko, A. (2009). Network neutrality, search neutrality, and the never-ending conflict between efficiency and fairness in markets. *Review of Network Economics*, 8.1, pp. 40–60.

Office of Fair Trading (2013). The economics of online personalised pricing. OFT 1488. URL: `https://webarchive.nationalarchives.gov.uk/ukgwa/20140402154756/http:/oft.gov.uk/shared_oft/research/oft1488.pdf` (visited on 06/29/2022).

Pigou, A. C. (1920). *The Economics of Welfare*. London: Macmillan.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review*, pp. 1281–1302.

Raju, C., Y. Narahari, and K. Ravikumar (2003). Reinforcement learning applications in dynamic pricing of retail markets. In: *IEEE International Conference on E-Commerce*. Ed. by IEEE, pp. 339–346.

Reinartz, W. (2002). Customizing prices in online markets. *Emerging Issues in Management*, 1, pp. 55–65.

Shepard, A. (1991). Price discrimination and retail configuration. *Journal of Political Economy*, 99.1, pp. 30–53.

Shiller, B. R. (2013). *First Degree Price Discrimination Using Big Data*. Waltham, Massachusetts: Brandeis University, Department of Economics and International Business School.

Stigler, G. J. (1966). *The Theory of Price*. New York: Macmillan.

Sutton, R. S. and A. G. Barto (2018). *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts: MIT Press.

Varian, H. R. (1989). Price discrimination. In: *Handbook of Industrial Organization*. Ed. by R. Schmalensee and R. D. Willig. Vol. 1. Amsterdam, The Netherlands: Elsevier, pp. 597–654.

Watkins, C. and P. Dayan (1992). Q-learning. *Machine Learning*, 8.3, pp. 279–292.

Woodcock, R. (2017). The bargaining robot. *CPI Antitrust Chronicle*, pp. 1–6.

3.A Appendix: Reward Histories

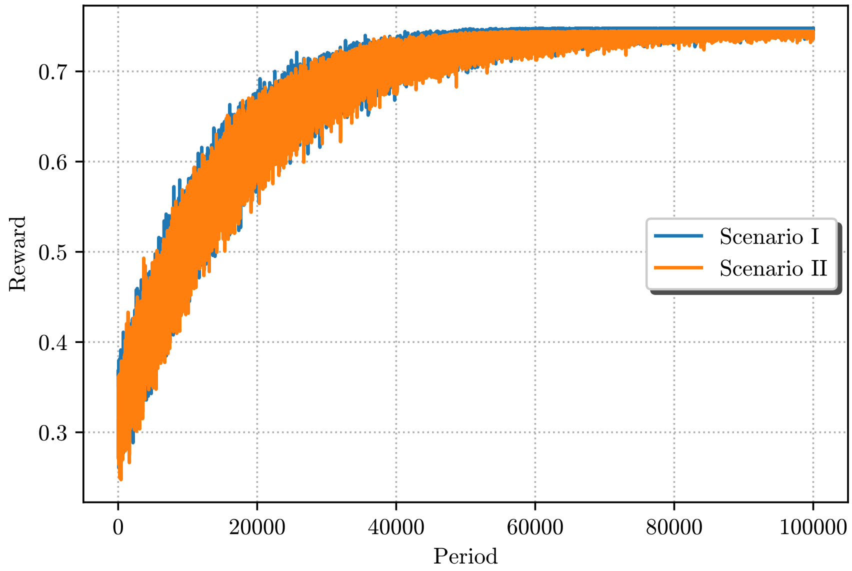

Figure 3.A.1: Reward Histories

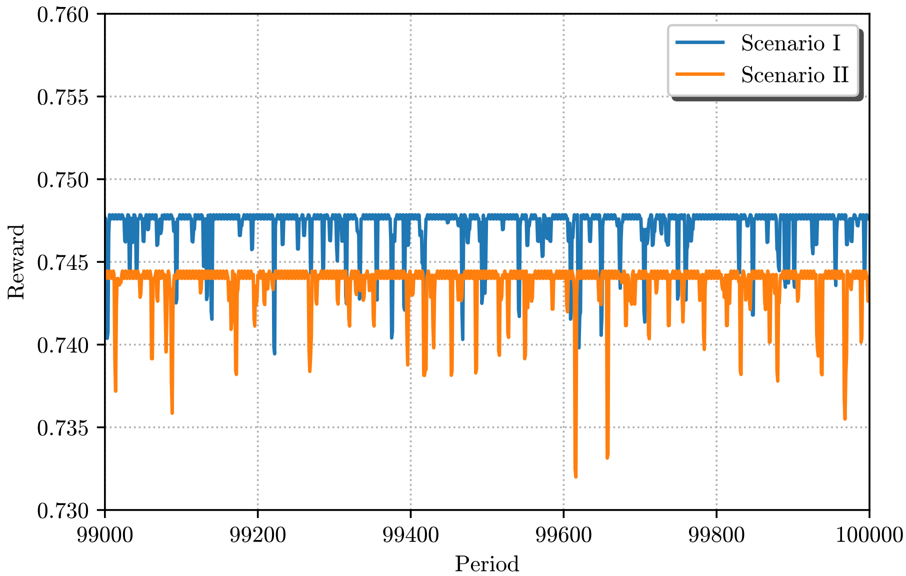

Figure 3.A.2: Reward Histories of the Final 1,000 Periods

3.B Appendix: Robustness Checks

Robustness checks were conducted for different values of the sensitivity parameters. Below are the values representing two customer groups with very distinct willingness to pay for a first robustness check (RC1) and two customer groups with similar willingness to pay for a second robustness check (RC2).

	Group	v	w
RC1	1	20	-23.690
	2	3.5	-15.526
RC2	1	8	-16
	2	5.5	-15.526

The respective acceptance probability functions for both robustness checks are shown in Figure 3.B.1.

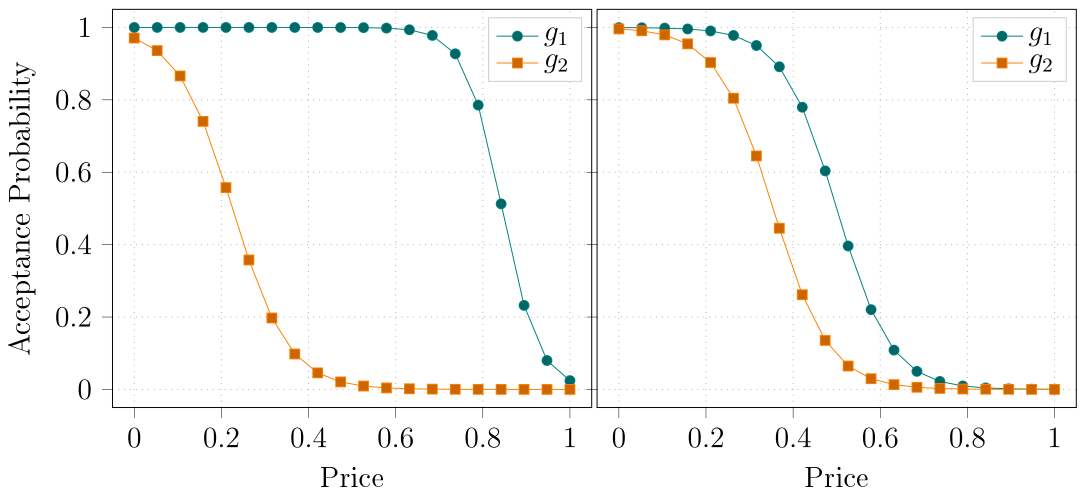

FIGURE 3.B.1: Acceptance Probability Functions for Distinct (left) and Similar (right) Sensitivities

Table 3.B.1 provides the analytical solution as well as simulation results for both robustness checks.

Scenario		a_1^*	a_2^*	r^*	$\bar{a}_1$	$\bar{a}_2$	$\bar{r}$	$\bar{f}$
RC1	I	0.727	0.185	0.805	0.700	0.190	0.794	0.826
	II	0.726	0.698	0.685	0.700	0.700	0.678	1.000
RC2	I	0.395	0.277	0.546	0.400	0.300	0.542	0.994
	II	0.394	0.278	0.545	0.399	0.299	0.541	0.994

TABLE 3.B.1: Analytical Solution and Simulation Results for RC1 and RC2

Chapter 4

The Effects of Movement Restrictions on Consumer Prices During the COVID-19 Pandemic

This chapter is joint work with Marcel Gehrung.

Abstract

In this study, we present empirical evidence for the impact of government measures that were imposed to contain the spread of the coronavirus and, as a consequence of their enforcement, reduced mobility on consumer prices during the COVID-19 pandemic. For this, we use the Oxford COVID-19 Government Response Tracker (OxCGRT) to examine the effect of the stringency of governmental measures and compare this result to the effect of actual mobility by using mobility data sourced from Google on the overall consumer price index as well as on the categories of food and housing and utilities. Since the main categories of the consumer price index are a highly aggregated measure, we take a more detailed look at sub-categories of the food sector, which involve nonperishable staple foods that can be stored as well as high-value products that should be consumed immediately. Our sample includes 32 European countries plus the United States for the period from January 2020 to May 2021. To summarize our findings, results of fixed-effects panel regressions show that more stringent measures lead to a significant increase in consumer prices. This finding is supported by the result of the regression analysis using actual mobility data as the independent variable.

4.1 Introduction

By the end of 2021, the cumulative number of confirmed COVID-19 infections had reached more than 288.6 million worldwide, and the death toll associated with the coronavirus had increased to more than 5.4 million cases. The rapid increase in the number of COVID-19 infections prompted governments in affected countries to impose measures designed to contain the spread of the coronavirus, including border closures that severely restricted mobility between countries, so-called stay-at-home restrictions, and workplace closures. These restrictions were expected to affect both demand and supply. On the one hand, workers and goods could cross national borders only under more restrictive conditions, resulting in an abrupt worker shortage and decreased supplies of certain goods. On the other hand, consumers adjusted their spending behavior in response to the pandemic due to shifts in preferences, expected income or health risks, or higher economic uncertainty. After the relaxation of several lockdown measures, a strong economic recovery and continuing supply chain disruptions have caused the price of goods and services to rise since the start of 2021 (Attinasi et al., 2021; Helper and Soltas, 2021).

In this study, we present empirical evidence for the impact of pandemic-related government measures and, as a consequence of their enforcement, reduced mobility on consumer prices. We use data from the European Union's Harmonized Index of Consumer Prices (HICP) for all member states of the European Union as well as for Switzerland, the United Kingdom, Iceland, Norway, Serbia and the United States for the period January 2020 to May 2021. Thus, we cover the very beginning of the pandemic, the major lockdowns during winter 2020, and the time when the first restrictions were eased in spring 2021. Moreover, we use the Oxford COVID-19 Government Response Tracker (OxCGRT) to determine the stringency of governmental measures.

Since the stringency of government-imposed restrictions does not necessarily reflect the real-world behavior of consumers, we further examine the effects of actual mobility by using mobility data sourced from Google on the consumer price index in an additional specification.

By means of fixed-effects panel regressions, we then attempt to answer the following research questions: (1) Is there a significant impact of the stringency of pandemic-related measures that were imposed by governments on consumer prices? (2) Are there different effects of government measures that aim to reduce population's mobility and the actual observed mobility on consumer prices?

Our findings show that the stringency of measures imposed by governments has a positive and significant impact on the overall consumer price index as well as on the sub-index of the food category, which means that more stringent measures induce higher consumer prices in these categories. Regressions with actual mobility data instead of the stringency of government measures support these results. To the best of our knowledge, this is the first paper to consider a larger time span of the pandemic and to provide supporting results with regard to the effect of the stringency of measures with actual mobility data.

The paper is structured as follows: Section 4.2 summarizes the existing literature on consumer spending during the COVID-19 pandemic, as well as the impact of pandemic-related lockdowns on the supply-side of markets. Section 4.3 then presents our data followed by the econometric methodology described in Section 4.4. Section 4.5 provides our results with further robustness checks in Section 4.6, and Section 4.7 gives a conclusion as well as several policy implications of our results.

4.2 Literature Review

During the pandemic, numerous measures were adopted in the affected countries to contain the spread of the coronavirus. In addition to closing borders, which severely restricted mobility between countries, stay-at-home restrictions and workplace closures were imposed. These measures were expected to affect both demand and supply.

In its technical report, the OECD (2020a) estimated the potential direct impact of lockdowns, taking into account sectoral output as well as consumption patterns across countries. Looking at different output categories, the authors conclude that activities involving travel and direct contact between consumers and service providers were adversely affected by movement restrictions and social distancing. Due to these restrictions, most retailers, restaurants, and cinemas were closed and the loss of sales could not be fully compensated by take-away or online offers. Moreover, nonessential construction work was delayed, either because of containment policies such as closed borders, which directly affect labor availability, or because of temporary reductions in investments. In manufacturing sectors, the impact of lockdown measures seemed to be smaller since these sectors are less labor intensive. Nevertheless, complete shutdowns or limited production of certain producers of transport equipment were observed due to difficulties obtaining necessary inputs from suppliers in other countries.

Additionally, the authors of the OECD report looked at detailed categories of consumer spending to estimate the effect of shutdowns in different sectors.

Similar to the results for changes in output, shop closures and travel restrictions naturally led to less spending or no spending, for example, on clothing and package holidays. Moreover, spending involving direct contact between consumers and service providers was likely to be postponed. Although expenditures for local travel, restaurants, and recreational services continued to a certain extent, there was a decline overall. Spending on essential items, however, remained almost unchanged. The European Central Bank drew the same conclusions (Lane, 2021).

As the report of the OECD (2020a) has shown, the impact of pandemic-related government measures on different economic sectors varies. Therefore, it is reasonable to focus on a sector where restrictions are expected to have the most severe impact: In addition to the main categories of the consumer price index, which are highly aggregated, we take a more detailed look at the food sector in our study. This sector is of particular interest because it includes different types of products such as meat, fruits, vegetables, and dairy products, as well as nonperishable goods that are differentiated in the calculation of the index. Food products can be divided into different categories according to their production, delivery, and storage options. The first category includes staple foods that can be stored for a longer period, such as various grains, beans, or oil seeds. The second category includes high-value products that should be consumed immediately, such as fruits, vegetables, fish, and meat. According to the report of the Food and Agriculture Organization of the United Nations (2020), constrained mobility between regions and countries has a negative impact on the distribution of staple foods, while the largely mechanized processing of these products is expected to be less affected. By contrast, in the production of high-value and fresh products, the higher demand for labor input leads to a more sensitive reaction to restrictions on workers, which was also identified by S. Aday and M. S. Aday (2020).

Tauber and van Zandweghe (2021) divided goods into durable and non-durable products. In their study, they observed an increase in consumer spending for durable goods on the US market. The increase was explained by consumer demand that shifted from services toward durable goods due to lockdowns, social distancing, and panic-buying, on the one hand, and an increase in disposable income resulting from fiscal policy measures designed to stimulate consumption expenditures, including those on durable goods, on the other hand.

Additional studies that examined the impacts of COVID-19 on spending behavior in Denmark (Andersen et al., 2020), France (Landais et al., 2020), Spain (Carvalho et al., 2020), the United Kingdom (Chronopoulos et al., 2020), the United States (Baker et al., 2020) and China (Chen et al., 2021) used bank transaction data and observed a decline in aggregate consumer spending during the first wave of the pandemic. On the US market, Baker et al. (2020) found significant changes in consumer spending across a broad variety of product categories, with only food delivery and grocery spending as major exceptions, which differed mainly by levels of liquidity and family structure. Andersen et al. (2020) found that in Denmark the decline in consumer spending varied across product categories, where individual exposure to health risks and supply restrictions was associated with much larger spending cuts than exposure to income risk and unemployment. However, these demand shocks were temporary, and consumer spending recovered quickly as the number of infections declined and restrictions eased.

Restricting mobility during a lockdown might also have the potential to impair economic activity, especially in the food market. For example, instead of visiting their usual retailers, consumers may shop at grocery stores close to their place of residence instead. Contributing to this topic, Bounie et al. (2020) observed that households concentrated their purchases on a smaller number of food retailers. Visiting a smaller number of retailers nearby reduces consumers' ability to choose from a sufficiently large portfolio of competing offers, thereby limiting substitution opportunities and, thus, decreasing the price elasticity of demand. Moreover, restaurants, cafés or snack bars which are engaged in food services or the sale of processed foods can be considered as substitutes for purchasing food at grocery stores and preparing meals at home. During lockdowns, these businesses needed to temporarily shut down, resulting in a reduced number of food suppliers and, as a result, less competition. Together with the aforementioned reduced price elasticity of demand, the temporary increase of market power of retailers that maintain food provision may have led to raising retail prices. The underlying micro-economic effects are further discussed by Ihle et al. (2020) who also provided a framework to quantify them.

The study by Akter (2020) further examined the effect of government-induced stay-at-home restrictions on food prices in the European Union in the first wave of the pandemic. Results suggested that meat, fish and seafood, and vegetables witnessed the most significant price surges, whereas prices of bread and cereals, fruits, milk, cheese and eggs as well as oils and fats were not significantly affected. However, the author considered only the period from January to March 2020.

As the consumer price index is the dependent variable in our regression analysis, we would like to draw attention to several problems that might occur when using this variable. As previously stated, consumers changed their spending patterns during the pandemic. These changes in consumption and, thereby, the composition of expenditure are taken into account by reweighting the price index each January. It follows, then, that the 2021 HICP, which considers the different expenditure composition of the previous year, assigns more weight to sectors that experienced a surge in expenditure in 2020 and, as a result, are associated with a higher pricing pressure compared with sectors that suffered a drop in expenditure and are associated with a lower pricing pressure. Gonçalves et al. (2021) showed that this reweighting of the price index alone accounted for 0.3 percentage points of the increase in the HICP in January 2021.

Related to this, Cavallo (2020) empirically studied the impact that changes in consumer expenditure patterns have had on the measurement of consumer price index inflation in the United States.[1] The author computed alternative "Covid Basket" price indices and showed that low-income households experienced higher inflation during the crisis than high-income households. However, we rely on the data for the traditional basket of goods used to calculate the consumer price index.

In addition, Blundell et al. (2020) stated that some cost increases could not be recorded. The authors argued that supply disruptions and stockpiling forced consumers to switch to smaller and usually more expensive package sizes (as measured by unit cost) or to different brands than they would normally buy. Consumers may also have had to shop at other stores where prices of the purchased goods were higher. Because the consumer price index captures the prices of fixed items at fixed locations, these additional costs were not reflected in the overall price increase.

With regard to the consumer price index, another problem arises during the collection of consumer prices in a pandemic as the collection in bricks-and-mortar stores stopped when they were closed due to government-related measures. In addition, sampling in supermarkets and drugstores was largely discontinued in order to protect price collectors. This resulted in the requirement of an estimation in areas where the collection of actual prices was substantially reduced.[2]

Additionally, other external factors such as the change in crude oil price as well as the implementation of temporary value added tax (VAT) reductions in several countries contributed to a change in consumer prices (Lane, 2021). We will consider these external factors in greater detail in our robustness checks.

[1] A similar attempt to construct experimental indices based on real-time consumption patterns was made by Kouvavas, Trezzi, Eiglsperger, et al. (2020) for the euro area.

[2] For a detailed discussion of this issue, see Kouvavas, Trezzi, Goldhammer, et al. (2020).

4.3 Data

Our sample includes all member states of the European Union as well as Switzerland, the United Kingdom, Iceland, Norway, Serbia, and the United States; thus, we consider a total of 33 countries in our sample. We collected data for the period January 2020 to May 2021 to grasp the full development and spread of the pandemic and its impacts from its beginning until the easing of restrictions in spring 2021.

Data for the consumer prices stems from the European Statistical Office (Eurostat).[3] We use monthly data of the HICP for the above-mentioned countries. To calculate the consumer price indices, the national statistical institutes in the selected countries collect the prices of over 700 representative items in different regions and types of shops across the respective country. The list of items is updated annually. Price collectors survey the prices of examples of these items each month to see how they are changing, holding constant features such as brand, make, and package size. Average price changes for each item are then weighted according to their importance in households' budgets in a baseline year.

The consumer price index measures how much the cost of purchasing a typical basket of goods and services has changed over time, leading to a reasonable idea of how price increases are affecting households, or at least a "typical" one. In our analysis, we use the HICP, where the term "harmonized" indicates that all the countries follow the same methodology to determine the index. This ensures that the data for one country is comparable to the data for another. To disentangle the effects on different product areas, we utilize the classification of HICP components according to the Classification of individual consumption by purpose (COICOP). In addition to the overall index, we consider the COICOP headings of (1) *food*[4] and (2) *housing, water, electricity, gas, and other fuels*[5], since these two categories include the most essential products for everyday life and account for the largest shares of expenditures covered by the overall index.[6] Furthermore, we look more closely at the food classification, which is further divided into the categories (1.1)

[3] Data are available at `https://ec.europa.eu/eurostat/databrowser/product/view/PRC_HICP_AIND`. For the United Kingdom, data are provided by the Office of National Statistics at `https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceinflation`.

[4] Contains food products purchased for consumption at home. Note that *food* is a sub-category of the classification *food and non-alcoholic beverages*.

[5] Contains rentals for housing, maintenance and repair of the dwelling, water supply, electricity, gas and other fuels. In the following, we refer to this category as *housing and utilities*.

[6] In 2020, the average expenditure weight of the food category for our sampled countries was at 15.94%. On average, 13.27% of total expenditure covered by the index was spent on housing and utilities in the same year. In 2021, average expenditure weights of the categories of food and housing and utilities were at 18.04% and 14.40%, respectively.

bread and cereals; (1.2) *meat*; (1.3) *fish and seafood*; (1.4) *milk, cheese and eggs*; (1.5) *oils and fats*; (1.6) *fruits*; (1.7) *vegetables*; and (1.8) *sugar, jam, honey, chocolate and confectionery.* Except for the United States, all countries in our sample report prices in these sub-categories.

To explain changes in supply and demand induced by pandemic policies, we use the stringency index (SI) of the OxCGRT introduced by Hale et al. (2021).[7] The OxCGRT tracks governmental COVID-19 responses in real-time and aggregates them in several policy domains. The stringency values are provided as a composite index ranging from 0 (low stringency) to 100 (high stringency), and are calculated based on nine indicators including (1) school closures, (2) workplace closures, (3) canceled public events, (4) restrictions on gathering size, (5) public transport closures, (6) stay-at-home requirements, (7) restrictions on internal movement, (8) restrictions on international travel, and (9) public information campaigns. Each of these indicators is measured on a daily basis by an ordinal scale and takes the geographical scope into account. The codebook used by Hale et al. (2021) can be found in the Appendix 4.A. For our analysis, we calculated average monthly values of each indicator for each country. The stringency index $SI_{i,t}$ is then calculated by averaging the values of the $k = 9$ indicators according to the equation

$$SI_{i,t} = \frac{1}{k} \sum_{j=1}^{k} \left(100 \cdot \frac{v_{i,j,t} - 0.5(F_j - f_{i,j,t})}{m_j} \right),$$

where j is the indicator, t is the month, i represents the country, $v_{i,j,t}$ is the recorded policy value, F_j is a binary variable that is 1 if the indicator has a flag variable or 0 if not, $f_{i,j,t}$ is the recorded binary flag indicating the geographical scope of the respective indicator[8], and m_j is the maximum value of the respective indicator. This calculation normalizes the different ordinal scales to produce an index score between 0 and 100.

The various measures taken by the governments in our sample aimed to restrict the mobility of their respective citizens and, thereby, to reduce close contacts and the resulting spread of coronavirus infections. For an understanding of the effects of reduced mobility on consumer behavior and prices, we include mobility data by

[7] Data are available at `https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker`.

[8] For school closures, workplace closures, canceled public events, restrictions on gathering size, public transport closures, stay-at-home requirements and public information campaigns, this flag variable is equal to 1 meaning that measures apply across the country. For the indicators of restrictions on internal movement and restrictions on international travel, this flag variable is equal to 0 meaning that measures apply only to a sub-region of a country, or a specific sector.

Google in our data set. Reduced mobility should coincide with disrupted supply chains as workers have to stay home instead of going to their factories or offices. Furthermore, we also expect changes in demand since customers might choose to do less grocery shopping to protect themselves from infection and might reduce consumption to a necessary minimum or shift consumption to a later time in line with a precocious saving behavior.

Google's data sets use aggregated and anonymized data from its app Google Maps showing how visits and length of stays at different geographic regions and categories of locations[9] change compared to a baseline. Changes for each day are compared to a baseline value for that day of the week, which is the median value for the corresponding day of the week during the five-week period from January 3 to February 6, 2020.[10] For Cyprus and Iceland, there is missing mobility data. In our analysis, we consider only the category of workplaces. There are a number of reasons for choosing this category of mobility data since a change in mobility at workplaces induces changes on both the demand and the supply side. Reduced mobility at workplaces suggests that people are more likely to stay at home and consume other goods and services there, such as food or energy. Moreover, reduced mobility at workplaces could also indicate that employees are not only working from home but that their working hours were reduced or they have even lost their jobs. The latter would imply a further shift in consumption. The supply side is also affected by workplace closures, which are associated with reduced mobility at workplaces; this is especially true for work areas where employees' presence is mandatory, such as on production lines.

The extent to which these changes have actually occurred cannot be quantified on the basis of Google's mobility data. However, in combination with a comparison of the impact of the stringency of government measures, this provides an initial idea about the extent to which the measures have been implemented specifically at the workplace and the kind of effect(s) these types of policy measures have had on consumer prices.

We also include indicators of pandemic severity as additional control variables in our data set. These are measured by the number of new cases and deaths due to COVID-19 per million people per month for each country.[11] COVID-19 case and death numbers were $log(x+1)$-transformed to reduce skewness in their

[9] Categories are Residential, Parks, Grocery and Pharmacy Stores, Retail and Recreation, Workplaces, and Transit Stations.

[10] Due to this baseline date, data on mobility changes will not be available before February 15, 2020. Data are available at `https://www.google.com/covid19/mobility/`.

[11] Data on confirmed cases and deaths stem from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. The data are available at `https://coronavirus.jhu.edu/`.

distributions. To control for external effects of oil-price fluctuations that occurred during the pandemic (Le et al., 2021; OECD, 2020b), we added the crude oil price (in US dollars) as a further control variable in our data set.[12]

4.4 Methodology

With the data described above, we run panel regressions with country-fixed effects. We use serial correlation-robust standard errors to account for heteroskedasticity and autocorrelation. Before that, we check whether the coefficients of a fixed-effects model differ significantly from a random-effects approach. The test by Hausman (1978) suggests that a random-effects model in general would be enough. However, it is common in policy evaluation and applied microeconometrics (see, e.g., Angrist and Pischke (2008) or Khandker et al. (2009)) to focus on fixed-effects estimates to account for latent time-invariant characteristics in the observed sample. Furthermore, random effects require much stronger assumptions for consistency, which, in practice, are difficult to fulfill. Particularly, the consistency of random-effects estimates requires that unobserved individual effects that are constant across time (captured by α_i in our model) are not correlated with any of the observable covariates. In our particular case, this is not likely to be reasonable as unobserved heterogeneity might occur in terms of a country's (health) infrastructure, main industry type (capital- vs. labor-intensive industries), trust in the policy regime, and other factors that are constant over time. Such factors are likely to be correlated with the ability to deal with the pandemic, the policy trend (a more relaxed or a stricter approach to fighting the pandemic), or the willingness of a country's inhabitants to comply with measures, which, in turn, affects our observable covariates.

Additionally, Wooldridge (2015) stated that, in some applications of panel data methods, a sample cannot be treated as a random sample from a large population, especially when the unit of observation is a large geographic unit such as states or provinces. Then, it often makes sense to think of each α_i as a separate intercept to estimate for each cross-sectional unit and use a fixed-effects approach. Wooldridge (2015) also noted that a fixed-effects approach is almost always much more convincing than a random-effects approach for policy analysis using aggregated data. Nevertheless, in addition, we run panel regressions with random effects where the results do not differ qualitatively. The results are reported in Appendix 4.B.

[12] Data are available at `https://databank.worldbank.org/home`.

Another approach often used in policy analysis is the additional inclusion of time-fixed effects, which allows the elimination of bias from unobservables that change over time but are constant over entities. As a result, any variable that varies only across time and not across units will be collinear with the dummy variables, and its effect cannot be estimated. Given that, due to the nature of the pandemic, there were parallel trends in COVID-19 case numbers and, as a result, also in the stringency of restrictions across the countries in our sample. Including time-fixed effects in our regressions led to very large standard errors, indicating only minor differential variation in the explanatory variable across time by country, which is further exacerbated by the small sample size. Thus, in our main regressions, we only use cross-sectional fixed effects.

We then distinguish between two types of regressions to examine the effect of government measures as well as actual consumer behavior.

The equation for regressions of type one is

$$\begin{aligned}\mathbf{HICP_{i,t}} &= \beta_0 + \beta_1 SI_{i,t} + \beta_2 ln(NewCases + 1)_{i,t} \\ &\quad + \beta_3 ln(NewDeaths + 1)_{i,t} + \beta_4 oilprice_{i,t} + \alpha_i + u_{i,t}\ ,\end{aligned}$$

where the vector $HICP_{i,t}$ represents a set of harmonized consumer price indices. In a first specification we use the overall HICP, the index for food, and the one for housing and utilities. After that, we differentiate between specific sub-indices for different food categories, namely bread and cereals, meat, fish and seafood, dairy products, oils and fats, vegetables, and fruits. Meanwhile, $SI_{i,t}$ indicates the stringency of government measures to reduce mobility in the countries in our sample, $ln(NewCases+1)_{i,t}$ gives the logarithm of new case numbers for COVID-19 infections per million people, and $ln(NewDeaths + 1)_{i,t}$ gives the logarithm of the number of new deaths per million people resulting from COVID-19. In addition, $oilprice_{i,t}$ denotes the crude oil price, α_i represents country fixed-effects, and $u_{i,t}$ the error term. These regressions shed light on the effects of the stringency of government-imposed measures on consumer prices.

The equation for the second type of regressions is then

$$\begin{aligned}\mathbf{HICP_{i,t}} &= \beta_0 + \beta_1 MobilityData_{i,t} + \beta_2 ln(NewCases + 1)_{i,t} \\ &\quad + \beta_3 ln(NewDeaths + 1)_{i,t} + \beta_4 oilprice_{i,t} + \alpha_i + u_{i,t}\ ,\end{aligned}$$

where the vector $HICP_{i,t}$ again represents a set of harmonized consumer price indices. As before, we first use the overall HICP classifications described above and then different sub-indices for different food categories. Instead of the stringency of restrictions imposed by a government, we use $MobilityData_{i,t}$ as a variable consisting of mobility data at workplaces provided by Google. Again, $ln(NewCases+1)_{i,t}$ gives the logarithm of new case numbers for COVID-19 infections per million people, $ln(NewDeaths+1)_{i,t}$ gives the logarithm of the number of new deaths per million people resulting from COVID-19, $oilprice_{i,t}$ denotes the crude oil price, α_i represents country fixed-effects, and $u_{i,t}$ the error term. These regressions examine the effects of reduced actual mobility on consumer prices.

4.5 Results

4.5.1 Descriptive Statistics

Table 4.1 presents descriptive statistics for our sampled countries from January 2020 to May 2021. Because the United States does not report detailed food classifications, the number of observations differs between the overall HICP and this sub-category. The summary statistics illustrate, once again, why the food sector in particular is of great interest for our study: This category shows major fluctuations in the HICP, which are driven primarily by the product categories of fruits and vegetables.

TABLE 4.1: Summary Statistics

Variables	Obs.	Mean	SD	Min	Max
HICP All	561	106.84	3.92	98.41	118.40
HICP Food	544	109.76	6.12	91.50	127.14
Bread, Cereals	544	107.72	6.72	90.20	125.85
Meat	544	109.76	6.87	90.20	124.88
Fish, Seafood	544	112.18	7.55	92.90	131.34
Milk, Cheese, Eggs	544	107.20	6.42	93.10	129.52
Oils, Fats	544	112.71	9.38	90.44	135.39
Fruits	544	115.14	11.79	95.90	172.35
Vegetables	544	115.14	13.03	88.09	172.82
Sugar, Jam, a.o.	544	104.32	6.51	78.70	117.69
HICP Housing and Utilities	561	107.86	6.27	92.64	137.50
SI	561	53.23	22.78	0	98.45
Mobility Workplaces	496	-25.24	11.54	-68.60	5.87
New Cases per Million	561	135,452.80	182,040.70	0	935,091.20
New Deaths per Million	561	2,531.38	3,626.40	0	21,354.43
Crude Oil Price	561	47.20	12.78	21.04	66.40

The price index for housing and utilities also shows fluctuations. Within this category, rentals account for the largest share of expenditures and thus have the greatest impact on the index value. Since only a small fraction of the rental payments covered by the HICP – typically those referring to new rental agreements for non-social housing – might be expected to directly adapt to market forces, index fluctuations during the observed period may be explained primarily by changes in prices for energy supply. However, with respect to our research question, prices for energy supply are assumed to be quite robust to pandemic-related policy measures due to the energy market's essential and highly automated character making it less prone to disruptions in the workforce or supply chain.

The reason for the negative mean value of the mobility at workplaces is that Google displays mobility changes by comparing changes in mobility for each day to a baseline value for that day of the week. Therefore, a negative sign indicates a reduction in mobility.

As an illustrative example, Figure 4.1 presents the trend in aggregate HICP in Germany for all categories as well as for the specific categories of food and housing and utilities. The trend of the HICP for these categories is almost parallel, although a stronger fluctuation of the index was observed for food items. We can already see from this figure that the price index rose in those months when there were stricter restrictions due to increases in the numbers of COVID-19 cases, i.e., in the second quarter of 2020 and in the first and second quarters of 2021.

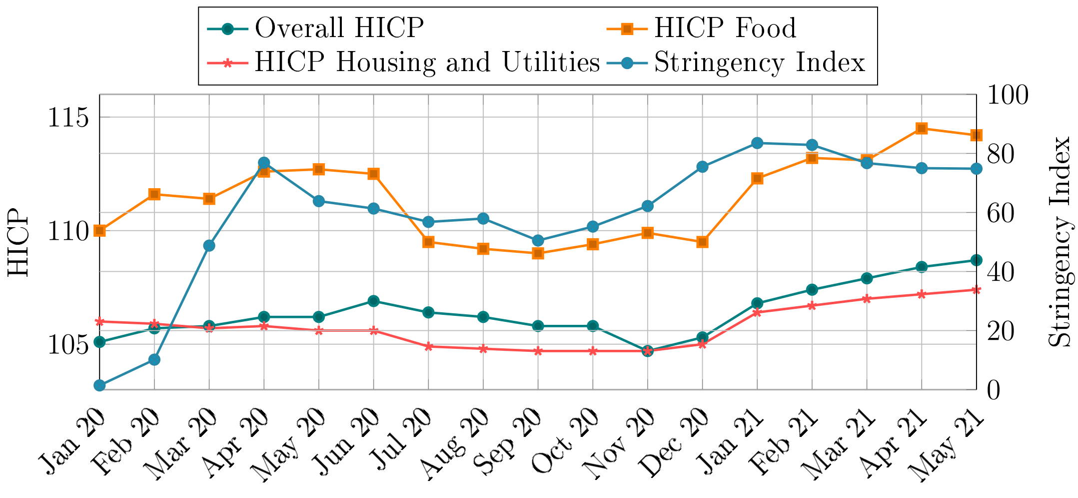

Figure 4.1: HICP and Stringency Index in Germany

In contrast to the HICP for food, the German price index for housing and utilities remained stable and only minor fluctuations were observed, at least until the end of 2020. The increase in the price index at the beginning of 2021 could be explained by the reweighting of the index in January 2021. We consider this in our robustness checks.

4.5.2 Regressions Type 1

Table 4.2 presents a panel fixed-effects regression model for the overall HICP and the sub-indices for food and housing and utilities according to the COICOP. For each sub-index, coefficients of the stringency of government measures to reduce mobility and curb the spread of the virus are estimated from a model with a set of control variables including indicators for the severity of the pandemic and the crude oil price. Please note that the interpretation of coefficients is complex since the HICP still includes many different goods even in its sub-categories and, therefore, is a highly aggregated measure.

Table 4.2: Regressions Type 1: Main Categories

HICP	All	Food	Housing and Utilities
SI	0.017***	0.031***	0.020
	(0.005)	(0.008)	(0.017)
Control Variables	Yes	Yes	Yes
Constant	103.379***	109.273***	102.137***
	(0.466)	(0.642)	(1.436)
Obs.	561	544	561
Overall R^2	0.325	0.087	0.276

This table reports results of fixed-effects panel regressions with heteroskedasticity and autocorrelation consistent (HAC) standard errors. The HICP and its specifications is the dependent variable. The independent variable is the SI indicating the stringency of measures imposed by the government due to the COVID-19 pandemic. Controls include the crude oil price and variables describing the severeness of the pandemic. These include the logarithm of new reported cases of COVID-19 infections per million people and the logarithm of new reported deaths due to COVID-19 per million people. Standard errors are in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

The coefficient of SI is positive and significant at the 1% level for the overall HICP and the HICP for food items. This implies that prices of items collected for these indices increase parallel to the strictness of government-imposed measures. On average, each one-point increase in the stringency index increases the overall HICP by 0.017 points. Consequently, a 58.82 increase in the stringency index is associated with a one-point increase in the overall HICP when all other variables in the model are held constant. For the food category, a smaller increase in the stringency index of 32.26 is associated with a one-point increase in the HICP for food, assuming all other variables are held constant.

The coefficient of SI for housing and utilities is also positive but not significant. This result supports our hypothesis that prices of goods and services covered by this category are robust to the stringency of pandemic-related government measures.

Investigating the relationship between the stringency of government measures and prices of sub-categories of the HICP for food items, we observe that coefficients of SI are positive and significant at the 1% level for meat; at the 5% level for fish and seafood, milk, cheese, and eggs, vegetables and fruits; and at the 10% level for oils and fats. The coefficients of SI for bread and cereals as well as for sugar, jam, and others are not significant. As the regression results shown in Table 4.3 indicate, it appears that the prices of perishable, labor-intensive products are more sensitive to restrictions than those of durable, capital-intensive products. Naturally, fresh products have a shorter shelf life and have to be processed and transported quickly using just-in-time delivery. Thus, shortages of workers and disrupted supply chains and transportation routes due to movement and travel restrictions significantly and strongly affect the prices of these goods (Attinasi et al., 2021; Helper and Soltas, 2021).

TABLE 4.3: Regressions Type 1: Food Categories

HICP Food	Bread, Cereals	Meat	Fish, Seafood	Milk, Cheese, Eggs	Oils, Fats	Vegetables	Fruits	Sugar, Jam, and others
SI	0.011	0.024***	0.019**	0.017**	0.026*	0.060**	0.127**	-0.004
	(0.007)	(0.008)	(0.008)	(0.007)	(0.014)	(0.026)	(0.047)	(0.010)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	105.452***	110.256***	111.682***	105.947***	107.935***	116.910***	116.494***	103.217***
	(0.792)	(0.718)	(0.897)	(0.638)	(1.836)	(1.804)	(2.325)	(0.911)
Obs.	544	544	544	544	544	544	544	544
Overall R^2	0.191	0.128	0.018	0.080	0.093	0.106	0.118	0.045

This table reports results of fixed-effects panel regressions with HAC standard errors. The HICP for food and its specifications is the dependent variable. The independent variable is the SI indicating the stringency of measures imposed by the government due to the COVID-19 pandemic. Controls include the crude oil price and variables describing the severeness of the pandemic. These include the logarithm of new reported cases of COVID-19 infections per million people and the logarithm of new reported deaths due to COVID-19 per million people. Standard errors are in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

4.5.3 Regressions Type 2

With the second type of regressions, we want to determine whether there is a difference between governments' instructions advising people to restrict their mobility and their actual mobility. Again, Table 4.4 shows the results of a panel fixed-effects regression model for the overall HICP and the sub-indices for food and housing and utilities. For each sub-index, coefficients of Google's data on mobility at workplaces are estimated from a model with a set of control variables including indicators for the severity of the pandemic and the crude oil price.

TABLE 4.4: Regressions Type 2: Main Categories

HICP	All	Food	Housing and Utilities
Mobility Workplaces	-0.014**	-0.016**	-0.019
	(0.007)	(0.007)	(0.012)
Control Variables	Yes	Yes	Yes
Constant	104.018***	110.604***	101.510***
	(0.480)	(0.422)	(1.401)
Obs.	496	480	496
Overall R^2	0.307	0.054	0.298

This table reports results of fixed-effects panel regressions with HAC standard errors. The HICP and its specifications is the dependent variable. The independent variable is mobility data from Google, describing the percentage change from the baseline in the area of workplaces. Controls include the crude oil price, and variables describing the severeness of the pandemic. These include the logarithm of new reported cases of COVID-19 infections per million people and the logarithm of new reported deaths due to COVID-19 per million people. Standard errors are in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

While our previous regression results suggest a significant effect of the stringency of measures on overall HICP and the HICP for food, the results of regressions considering actual mobility suggest a significant impact of actual mobility on the same categories. The coefficients of mobility at workplaces are negative and significant for overall HICP and HICP for food at the 5% level. The negative sign of our coefficients indicates that less mobility (which might be induced by stricter mobility restrictions or their expectation) is associated with higher prices. On average, a one percentage point decrease in mobility is associated with a 0.014 point increase in the overall HICP. In other words, a 71.43 percentage point reduction in mobility increases the overall HICP by one point while the other variables in the model are held constant. In the food category, a 62.5 percentage point decrease in mobility results in a one-point increase in the HICP for food, assuming all other variables are held constant.

This is in line with the regression results for the coefficients of SI in Table 4.2, where we observed a positive relationship between SI and HICP, i.e., stricter government measures designed to reduce mobility are associated with higher prices.

As for the first regression type, actual mobility data also show no significant effect on the HICP for the category of housing and utilities.

As shown in Table 4.3, the coefficients of mobility at workplaces are negative and significant for fruits at the 1% level; negative and significant at the 5% level for milk, cheese, and eggs and oils and fats; negative and significant at the 10% level for fish and seafood; and positive and significant at the 5% level for vegetables. For the sub-categories of meat, bread and cereals, as well as sugar, jam, and others, coefficients are not significant, meaning that actual mobility does not have a significant effect on prices of these items. Again, the negative and significant coefficients are in line with our previous findings. The positive sign for the sub-index of vegetables is somewhat counterintuitive.

TABLE 4.5: Regressions Type 2: Food Categories

HICP Food	Bread, Cereals	Meat	Fish, Seafood	Milk, Cheese, Eggs	Oils, Fats	Vegetables	Fruits	Sugar, Jam, and others
Mobility Workplaces	-0.009	-0.006	-0.025*	-0.012**	-0.049**	0.095**	-0.136***	0.004
	(0.006)	(0.007)	(0.014)	(0.005)	(0.019)	(0.041)	(0.032)	(0.008)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Constant	106.208***	111.673***	112.252***	106.387***	108.039***	123.747***	118.034***	103.610***
	(0.684)	(0.722)	(0.912)	(0.417)	(2.107)	(1.987)	(1.455)	(0.777)
Obs.	480	480	480	480	480	480	480	480
Overall R^2	0.158	0.091	0.022	0.021	0.101	0.089	0.155	0.021

This table reports results of fixed-effects panel regressions with HAC standard errors. The HICP for food and its specifications is the dependent variable. The independent variable is mobility data from Google, describing the percentage change from the baseline in the area of workplaces. Controls include the crude oil price, and variables describing the severeness of the pandemic. These include the logarithm of new reported cases of COVID-19 infections per million people and the logarithm of new reported deaths due to COVID-19 per million people. Standard errors are in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

4.6 Robustness Checks

To further determine the validity of our results, we employ different robustness checks. First, we use different specifications of the OxCGRT to explain changes in the categories of the HICP. Our main results focus on the stringency of a composite index of various government measures. With this robustness check, we want to examine whether our results change if we only focus on the stringency of particular measures. More specifically, we run regressions with the indicators of stay-at-home restrictions, internal movement restrictions, and workplace closings as our independent variable. The regression results show that the sign of coefficients is the same as in our main regressions, however, some coefficients are not significant anymore. This can be explained by the fact that the individual indicators have a very coarse scale, whereas the composite stringency index allows for more variation as it covers more indicators and thus areas which are affected by the restrictions.

With regard to current numbers related to the COVID-19 pandemic, we included the number of new cases per million people, the number of new deaths per million people, and intensive care unit (ICU) and hospital admissions per million people simultaneously in our regressions as a second robustness check. This comes with the risk of multicollinearity, but our results are robust and do not change qualitatively to the regressions where only new cases and deaths are included.

A further issue we take into consideration for a robustness check is the effect of a reduced VAT in Germany.[13] In addition, several other countries induced sector-specific VAT cuts; for example, the British government allowed VAT-registered businesses to apply a temporary 5% reduced VAT rate to certain supplies relating to hospitality and hotel and holiday accommodations purchased between July 15, 2020 and March 31, 2021. To check our results for robustness to the VAT cuts, we run our regression analyses with a sub-sample excluding Germany, which was the only country in our sample that introduced a cut in the VAT across all products present in the index categories we consider here. Other countries that implemented VAT reductions focused, instead, on areas that we do not consider specifically in our sub-indices (e.g., hospitality or tourism). Since products from these categories are present only in the overall HICP at a low weighting because they were hardly consumed in relation to other products, their effect on the overall HICP can be ignored. This thesis is also supported by the results of our robustness checks. Neither signs of coefficients nor significance change when

[13] The standard VAT rate was reduced from 19% to 16% from July 1 to December 31, 2020. The reduced VAT rate of 7% was cut to 5%.

running the regression with the sub-sample excluding Germany. Thus, our results are robust to the implemented VAT cuts.

In order to capture changes in consumption and, thereby, the composition of expenditure, HICP weights are updated annually. Since the COVID-19 pandemic led to large shifts in consumption patterns, this resulted in considerable adjustments in weights of the 2021 HICP. Gonçalves et al. (2021) showed that this reweighting accounted for 0.3 percentage points of the increase in HICP in January 2021. To rule out an effect on our regression results, we run our regression analyses with a subset of our data including only the months from January 2020 to December 2020, thereby finding the qualitatively same results as for our main regressions.

4.7 Conclusion

The COVID-19 pandemic continues to affect many aspects of daily life worldwide, and while every consumer directly experienced the effects of government measures to curb its spread, they might also have been influenced indirectly by the virus through higher consumer prices.

To summarize the results of our analysis, with regard to our first research question, we find that the stringency of government measures to restrict mobility in order to contain the spread of the virus significantly increased consumer prices. This is true for the overall HICP as well as for the HICP for food. When we look further at the food sector, we find price increases due to more strict lockdown measures for the specific indices of meat, fish and seafood, dairy, oils, vegetables, and fruits. One reason might be that suppliers have to deal with disrupted supply chains, which is especially a problem for products that are freshly processed and delivered and require labor-intensive production. In contrast, products with a long shelf life are not affected by short-term restrictions.

For our second type of regressions, we find a negative and significant impact of actual mobility on overall HICP as well as on the HICP for food, meaning that less mobility is associated with a higher price index, which supports the results of our first regression model. Again, those results also hold for the granular sub-categories of the food index.

The HICP for housing and utilities is not significantly affected neither by the stringency of government measures nor by actual mobility. As mentioned earlier, prices for rentals do not seem to be affected by the pandemic in the short run, which leads us to conclude that changes in the HICP for this category may be explained by price changes in energy supply. However, energy supply is one of the most essential business sectors in an economy and, thus, disruptions due to missing workers were kept to an absolute minimum. Additionally, the energy supply is highly automated as pipelines and refineries are set up for long-term production, making this sector robust to supply-chain disruptions.[14]

Our findings, then, also have implications for economic policy. The first and most important goal of policymakers at the beginning of the pandemic had to be the protection of citizens and their health. Thus, strict measures to curb the spread of the virus and to reduce mobility were necessary at that point. However, those measures came with a variety of drawbacks. Besides the psychological strain lockdowns put on people, the economic consequences cannot be neglected. As we show, the severity of government measures significantly pushes prices upwards. In a pandemic, with lower incomes or short-time work, this hits consumers hard, especially households with lower incomes. Thus, with fewer deadly variants of the virus and high numbers of vaccinated citizens, governments must now re-evaluate their stay-at-home measures and keep an eye on their contribution to price increases.

[14] The current war in Ukraine poses a whole different shock to the energy supply, however, did not affect prices during the observed period.

References

Aday, S. and M. S. Aday (2020). Impact of COVID-19 on the food supply chain. *Food Quality and Safety*, 4.4, pp. 167–180.

Akter, S. (2020). The impact of COVID-19 related 'stay-at-home' restrictions on food prices in Europe: findings from a preliminary analysis. *Food Security*, 12.4, pp. 719–725.

Andersen, A. L., E. T. Hansen, N. Johannesen, and A. Sheridan (2020). Consumer responses to the COVID-19 crisis: Evidence from bank account transaction data. *Covid Economics*, 7, pp. 88–114.

Angrist, J. D. and J. Pischke (2008). *Mostly Harmless Econometrics*. Princeton and Oxford: Princeton University Press.

Attinasi, M. G., M. Balatti, M. Mancini, and L. Metelli (2021). Supply chain disruptions and the effects on the global economy. In: *Economic, Financial and Monetary Developments*. Ed. by European Central Bank. Economic Bulletin, 8/2021, pp. 51–57.

Baker, S. R., R. A. Farrokhnia, S. Meyer, M. Pagel, and C. Yannelis (2020). How does household spending respond to an epidemic? Consumption during the 2020 COVID-19 pandemic. *The Review of Asset Pricing Studies*, 10.4, pp. 834–862.

Blundell, R., P. Levell, and M. O'Connell (2020). Could COVID–19 infect the Consumer Prices Index? *Fiscal Studies*, 41.2, pp. 357–361.

Bounie, D., Y. Camara, and J. W. Galbraith (2020). Consumers' mobility, expenditure and online-offline substitution response to COVID-19: Evidence from french transaction data. SSRN Working Paper. DOI: `10.2139/ssrn.3588373`.

Carvalho, B., S. Peralta, and J. Pereira dos Santos (2020). What and how did people buy during the Great Lockdown? Evidence from electronic payments. *Covid Economics*, 28, pp. 119–158.

Cavallo, A. (2020). Inflation with covid consumption baskets. Working Paper No. 27352. National Bureau of Economic Research. DOI: `10.3386/w27352`.

Chen, H., W. Qian, and Q. Wen (2021). The impact of the COVID-19 pandemic on consumption: Learning from high-frequency transaction data. AEA Papers and Proceedings. DOI: `10.2139/ssrn.3568574`.

Chronopoulos, D. K., M. Lukas, and J. O. S. Wilson (2020). Consumer spending responses to the COVID-19 pandemic: An assessment of Great Britain. *Covid Economics*, 34, pp. 145–186.

Food and Agriculture Organization of the United Nations (2020). COVID-19 and the risk to food supply chains: How to respond? URL: `http://www.fao.org/3/ca8388en/CA8388EN.pdf` (visited on 08/24/2020).

Gonçalves, E., L. Henkel, O. Kouvavas, M. Porqueddu, and R. Trezzi (2021). 2021 HICP weights and their implications for the measurement of inflation. In: *Economic and Monetary Developments*. Ed. by European Central Bank. Economic Bulletin, 2/2021, pp. 75–79.

Hale, T., N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Philips, S. Webster, E. Cameron-Blake, L. Hallas, S. Majumdar, and H. Tetlow (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behavior*, 5.4, pp. 529–538.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46.6, pp. 1251–1271.

Helper, S. and E. Soltas (2021). Why the Pandemic Has Disrupted Supply Chains. The White House. Washington, D.C. URL: `https://www.whitehouse.gov/cea/written-materials/2021/06/17/why-the-pandemic-has-disrupted-supply-chains/` (visited on 07/24/2022).

Ihle, R., O. D. Rubin, Z. Bar-Nahum, and R. Jongeneel (2020). Imperfect food markets in times of crisis: economic consequences of supply chain disruptions and fragmentation for local market power and urban vulnerability. *Food Security*, 12.4, pp. 727–734.

Khandker, S. R., G. B. Koolwal, and H. A. Samad (2009). *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington, D.C.: World Bank Publications.

Kouvavas, O., R. Trezzi, M. Eiglsperger, B. Goldhammer, and E. Gonçalves (2020). Consumption patterns and inflation measurement issues during the COVID-19 pandemic. In: *Update on Economic and Monetary Developments*. Ed. by European Central Bank. Economic Bulletin, 7/2020, pp. 34–41.

Kouvavas, O., R. Trezzi, B. Goldhammer, and J. Nordeman (2020). Inflation measurement in times of economic distress. In: *Update on Economic and Monetary Developments*. Ed. by European Central Bank. Economic Bulletin, 3/2020, pp. 39–43.

Landais, C., D. Bounie, Y. Camara, E. Fize, J. Galbraith, C. Lavest, T. Pazem, and B. Savatier (2020). Consumption dynamics in the COVID crisis: Real time insights from French transaction and bank data. CEPR Discussion Paper DP15474. URL: `https://cepr.org/publications/dp15474`.

Lane, P. R. (2021). Inflation dynamics during a pandemic. European Central Bank. URL: https://www.ecb.europa.eu/press/blog/date/2021/html/ecb.blog210401~6407b23d87.en.html (visited on 07/01/2022).

Le, T.-H., A. T. Le, and H.-C. Le (2021). The historic oil price fluctuation during the Covid-19 pandemic: What are the causes? *Research in International Business and Finance*, 58, p. 101489.

OECD (2020a). Evaluating the impact of COVID-19 containment measures on activity and spending. OECD Economic Outlook, Volume 2020 Issue 1. OECD Publishing, Paris. DOI: 10.1787/fe40a82a-en.

— (2020b). The impact of Coronavirus (COVID-19) and the global oil price shock on the fiscal position of oil-exporting developing countries. URL: https://www.oecd.org/coronavirus/policy-responses/the-impact-of-coronavirus-covid-19-and-the-global-oil-price-shock-on-the-fiscal-position-of-oil-exporting-developing-countries-8bafbd95/ (visited on 07/01/2022).

Tauber, K. and W. van Zandweghe (2021). Why has durable goods spending been so strong during the COVID-19 pandemic? *Economic Commentary*, 2021-16.

Wooldridge, J. M. (2015). *Introductory Econometrics: A modern approach.* Boston, USA: Cengage learning.

4.A Appendix: Codebook of Indicators

Table 4.A.1: Codebook of Indicators of Government Measures (Hale et al., 2021)

ID	Name	Description	Measurement	Coding instructions
C1	School closing	Record closings of schools and universities	Ordinal scale; binary for geographic scope	0 – No measures 1 – Recommend closing 2 – Require closing (only some levels or categories) 3 – Require closing all levels 0 – Targeted 1 – General
C2	Workplace closing	Record closings of workplaces	Ordinal scale; binary for geographic scope	0 – No measures 1 – Recommend closing (or work from home) 2 – Require closing (or work from home) for some sectors or categories of workers 3 – Require closing (or work from home) all-but-essential workplaces (e.g. grocery stores, doctors) 0 – Targeted 1 – General
C3	Cancel public events	Record canceling public events	Ordinal scale; binary for geographic scope	0 – No measures 1 – Recommend canceling 2 – Require canceling 0 – Targeted 1 – General

ID	Name	Description	Measurement	Coding instructions
C4	Restrictions on gatherings	Record the cut-off size for bans on gatherings	Ordinal scale; binary for geographic scope	0 – No restrictions 1 – Restrictions on very large gatherings (> 1000 people) 2 – Restrictions on gatherings between 101-1000 people 3 – Restrictions on gatherings between 11-100 people 4 – Restrictions on gatherings of 10 people or less 0 – Targeted 1 – General
C5	Close public transport	Record closing of public transport	Ordinal scale; binary on geographic scope	0 – No measures 1 – Recommend closing (or significantly reduce volume/ route/means of transport available) 2 – Require closing (or prohibit most citizens from using it) 0 – Targeted 1 – General
C6	Stay at home requirements	Record orders to "shelter-in-place" and otherwise confine to home	Ordinal scale; binary on geographic scope	0 – No measures 1 – Recommend not leaving house 2 – Require not leaving house with exceptions for daily exercise, grocery shopping, and "essential" trips 3 – Require not leaving house with minimal exceptions (e.g. only one person can leave at a time) 0 – Targeted 1 – General

ID	Name	Description	Measurement	Coding instructions
C7	Restrictions on internal movement	Record restrictions on internal movement	Ordinal scale; binary on geographic scope	0 – No measures 1 – Recommend not to travel between regions/cities 2 – Internal movement restrictions in place 0 – Targeted 1 – General
C8	International travel controls	Record restrictions on international travel	Ordinal scale	0 – No measures 1 – Screening 2 – Quarantine arrivals from high-risk regions 3 – Ban on arrivals from some regions 4 – Ban on all regions or total border closure
H1	Public info campaigns	Record presence of public info campaigns	Ordinal scale; binary on geographic scope	0 – No COVID-19 public information campaign 1 – Public officials urging caution about COVID-19 2 – Coordinated public information campaign (e.g. across traditional and social media) 0 – Targeted 1 – General

4.B Appendix: Results of the Random-Effects Model

TABLE 4.B.1: Random-Effects Regressions Type 1: Main Categories

HICP	All	Food	Housing and Utilities
SI	0.016***	0.030***	0.020
	(0.005)	(0.008)	(0.017)
Control Variables	Yes	Yes	Yes
Constant	103.388***	109.284***	102.133***
	(0.464)	(0.927)	(1.291)
Obs.	561	544	561

This table reports results of random effects panel regressions with HAC standard errors. The HICP and its specifications is the dependent variable. The independent variable is the SI indicating the stringency of measures imposed by the government due to the Covid-19 pandemic. Controls include the crude oil price and variables describing the severeness of the pandemic. These include the logarithm of new reported cases of Covid-19 infections per million and the logarithm of new reported deaths due to Covid-19 per million. Standard errors are in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

TABLE 4.B.2: Random-Effects Regressions Type 1: Food Categories

HICP Food	Bread, Cereals	Meat	Fish, Seafood	Milk, Cheese, Eggs	Oils, Fats	Vegetables	Fruits	Sugar, Jam, and others
SI	0.010	0.024***	0.018**	0.017**	0.025*	0.059**	0.124***	-0.005
	(0.007)	(0.008)	(0.008)	(0.007)	(0.014)	(0.026)	(0.047)	(0.010)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Constant	105.462***	110.267***	111.689***	105.949***	107.945***	116.986***	116.618***	103.225***
	(0.922)	(1.416)	(1.583)	(0.856)	(2.245)	(2.295)	(2.868)	(1.068)
Obs.	544	544	544	544	544	544	544	544

This table reports results of random effects panel regressions with HAC standard errors. The HICP for food and its specifications is the dependent variable. The independent variable is the SI indicating the stringency of measures imposed by the government due to the Covid-19 pandemic. Controls include the crude oil price and variables describing the severeness of the pandemic. These include the logarithm of new reported cases of Covid-19 infections per million and the logarithm of new reported deaths due to Covid-19 per million. Standard errors are in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

Table 4.B.3: Random-Effects Regressions Type 2: Main Categories

HICP	All	Food	Housing and Utilities
Mobility Workplaces	-0.014**	-0.015**	-0.019
	(0.007)	(0.007)	(0.012)
Control Variables	Yes	Yes	Yes
Constant	104.029***	110.636***	101.510***
	(0.464)	(0.978)	(1.287)
Obs.	496	480	496

This table reports results of random effects panel regressions with HAC standard errors. The HICP and its specifications is the dependent variable. The independent variable is mobility data from Google, describing the percentage change from the baseline in the area of workplaces. Controls include the crude oil price, and variables describing the severeness of the pandemic. These include the logarithm of new reported cases of Covid-19 infections per million and the logarithm of new reported deaths due to Covid-19 per million. Standard errors are in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

Table 4.B.4: Random-Effects Regressions Type 2: Food Categories

HICP Food	Bread, Cereals	Meat	Fish, Seafood	Milk, Cheese, Eggs	Oils, Fats	Vegetables	Fruits	Sugar, Jam, and others
Mobility Workplaces	-0.009	-0.005	-0.025*	-0.012**	-0.048**	0.101**	-0.124***	0.005
	(0.006)	(0.007)	(0.014)	(0.005)	(0.019)	(0.040)	(0.032)	(0.008)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Constant	106.222***	111.703***	112.258***	106.400***	108.051***	124.011***	118.471***	103.629***
	(0.859)	(1.566)	(1.646)	(0.926)	(2.593)	(2.603)	(2.686)	(1.102)
Obs.	480	480	480	480	480	480	480	480

This table reports results of random effects panel regressions with HAC standard errors. The HICP for food and its specifications is the dependent variable. The independent variable is mobility data from Google, describing the percentage change from the baseline in the area of workplaces. Controls include the crude oil price, and variables describing the severeness of the pandemic. These include the logarithm of new reported cases of Covid-19 infections per million and the logarithm of new reported deaths due to Covid-19 per million. Standard errors are in parentheses. *** denotes significance at the 1% level; ** denotes significance at the 5% level; * denotes significance at the 10% level.

Chapter 5

General Conclusions

This dissertation has dealt with questions around the implementation of pricing algorithms that use a reinforcement learning approach from two different perspectives. The first addresses algorithmic collusion. This topic was discussed in Chapter 2 of this dissertation in order to come to a better assessment of the extent to which there is indeed a risk of collusion by learning algorithms. The occurrence of tacit collusion was identified as a concern by Ezrachi and Stucke (2016) because, according to the authors, the existing regulations in the United States are not sufficient to prohibit tacit collusion where no intent to collude can be proven. Other scholars have called for improved enforcement practices, such as monitoring of algorithms by competition authorities (Harrington, 2018) or allowing firms to correct pricing decisions of their algorithms if they implement collusive strategies, which requires them to continuously monitor their algorithms (Massarotto, 2019).

In the European Union, the situation is somewhat different. In his study, Blockx (2017) pointed out that EU competition law allows tacit collusion to be captured. However, as stated in Chapter 2, collusive behavior of algorithms is (at least from the current point of view) not very likely to occur in reality. The reasons for this are technical challenges of reinforcement learning algorithms that are not solved at that point, such as dealing with non-stationary environments and a large number of possible actions and competitors, both requiring immense processing and time capacities to complete the learning process. Moreover, assumptions made in experimental studies were shown to be quite unrealistic and do not reflect economic reality. Consequently, autonomous collusion by learning algorithms does not currently seem to be a major competition concern.

Nevertheless, technological developments are not standing still, of course, and the application of other pricing algorithms and new technologies (such as blockchain technology) will certainly keep competition authorities busy, also with regard to collusive behavior (Cong and He, 2019; Schrepel, 2019). Thus, from a legal point of view, raising awareness for technological improvements and developments is still very important and one of the major challenges in this area.

Another ongoing debate is that of personalized pricing and the potential of learning algorithms to engage in discriminatory pricing strategies when more realistic assumptions about consumers are made. In particular, inequity aversion of consumers was considered in this thesis. As shown in Chapter 3, learning algorithms are indeed capable of recognizing consumers' fairness preferences and still setting revenue-maximizing, differential prices while taking the inequity aversion into consideration. This leads to the question of whether this result calls for political action. From a welfare economic perspective, the effect of price discrimination is ambiguous. Armstrong (2006), Stole (2007) and Varian (1989) provide a good overview on the outcomes and welfare effects of price discrimination in various settings. Generally, for price discrimination to improve overall welfare, a substantial increase in output is required by offering products to consumers who were previously unserved due to a low willingness to pay. This may lead to an improvement of both consumer and producer welfare. In cases where price discrimination does not lead to an increased supply, it often shifts welfare from consumers to producers or even reduces total welfare when the loss of consumer surplus is larger than the gain on the producers' side. In any case, consumers who are willing to pay higher prices are most probably worse off if sellers engage in price discrimination (Zuiderveen Borgesius and Poort, 2017).

For this reason, it is also worth taking a look at the current legal situation with regard to price discrimination and, a topic that is also involved, privacy issues of consumers. In general, personalized pricing strategies do not entirely fit the scenarios where discrimination is accepted under Article 102 of the Treaty on the Functioning of the European Union (TFEU); however, Article 102 TFEU is sufficiently flexible to capture the currently rare forms of price discrimination (Graef, 2017). As algorithmic personalized pricing is only possible when firms have enough information that is usually obtained by analyzing data, the European General Data Protection Regulation (GDPR) may be applied. This data protection law applies if personal data are processed. Zuiderveen Borgesius and Poort (2017) argue that online shops, which use cookies for identification, tracking, and categorization of customers, indeed use personalized pricing strategies and, thus, the GDPR can be applied to these cases. Since this regulation requires firms to be transparent about the purpose of the personal data processing, they must inform their customers and request their consent for processing the data if they engage in personalized pricing. However, when further regulating the collection of personal data one must take into consideration that this is only advantageous under certain conditions, e.g., when consumers are not very privacy-sensitive and the market is characterized by limited competition and firms engaging in price discrimination (Koh et al., 2017).

Another challenge faced by many governments in the short and mid run is combating inflation. This is currently reaching peak levels worldwide due to the war in Ukraine. Furthermore, measures introduced by governments during the ongoing pandemic also resulted in price increases. In Chapter 4, we examined the impact of the stringency of government measures that were imposed during the COVID-19 pandemic on consumer prices. Results of the fixed-effects regression models have demonstrated that governmental measures affecting population mobility had a significant impact on consumer prices mainly in the food sector. Within this category of the HICP, stricter measures led to higher prices. As a consequence, such measures should be weighed carefully, especially since it will probably not be the last time that the use of such measures will be considered as there may be further waves of the pandemic. The particular reason for the increase in prices was beyond the scope of this dissertation, but it is reasonable to assume that this is due to rising costs on the supply side. From a welfare economics perspective, producer welfare has at best remained the same (with the exception of a few profiteers of the pandemic), while consumer welfare has in any case declined due to the increase in prices. Thus, it is likely that an overall welfare loss occurred. Since it cannot be assumed that governments will be able to completely compensate for the rising prices (nor should they), it remains to be seen whether suitable measures will be found that do not drive prices up even further.

To conclude, this dissertation addressed the topic of algorithmic pricing, thereby showing that autonomous collusion by learning agents does not currently pose a major competition concern, in contrast to personalized pricing strategies, which have been shown to be more feasible with pricing algorithms using a reinforcement learning approach. Additionally, the thesis provides insight into the effects of pandemic-related policy measures by showing that more stringent government measures led to an aggravation of the situation during the COVID-19 pandemic by contributing to price increases in certain product categories.

References

Armstrong, M. (2006). Recent developments in the economics of price discrimination. In: *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*. Ed. by R. Blundell, W. K. Newey, and T. Persson. 2nd ed. Econometric Society Monographs. Cambridge, United Kingdom: Cambridge University Press, pp. 97–141.

Blockx, J. (2017). Antitrust in digital markets in the EU: Policing price bots. *Radboud Economic Law Conference*, 9.

Cong, L. W. and Z. He (2019). Blockchain disruption and smart contracts. *The Review of Financial Studies*, 32.5, pp. 1754–1797.

Ezrachi, A. and M. E. Stucke (2016). *Virtual Competition: the promise and perils of the algorithm-driven economy*. 2nd ed. Cambridge, Massachusetts: Harvard University Press.

Graef, I. (2017). Algorithms and fairness: What role for competition law in targeting price discrimination towards ends consumers. *Columbia Journal of European Law*, 24, pp. 541–558.

Harrington, J. E. (2018). Developing competition law for collusion by autonomous artificial agents. *Journal of Competition Law & Economics*, 14.3, pp. 331–363.

Koh, B., S. Raghunathan, and B. R. Nault (2017). Is voluntary profiling welfare enhancing? *Management Information Systems Quarterly*, 41.1, pp. 23–41.

Massarotto, G. (2019). From digital to blockchain markets: What role for antitrust and regulation. SSRN Working Paper. DOI: 10.2139/ssrn.3323420.

Schrepel, T. (2019). Collusion by blockchain and smart contracts. *Harvard Journal of Law & Technology*, 33.1, pp. 117–166.

Stole, L. A. (2007). Price discrimination and competition. In: *Handbook of Industrial Organization*. Ed. by M. Armstrong and R. H. Porter. Vol. 3. Amsterdam, The Netherlands: Elsevier, pp. 2221–2299.

Varian, H. R. (1989). Price discrimination. In: *Handbook of Industrial Organization*. Ed. by R. Schmalensee and R. D. Willig. Vol. 1. Amsterdam, The Netherlands: Elsevier, pp. 597–654.

Zuiderveen Borgesius, F. and J. Poort (2017). Online price discrimination and EU data privacy law. *Journal of Consumer Policy*, 40, pp. 347–366.