Revenue Management with Repeated Competitive Interactions

A Dissertation
Presented to the School of Business and Economics
of
Freie Universität Berlin
in Candidacy for the Degree of
doctor rerum politicarum
(Dr. rer. pol.)

by
Dipl.-Math. Benedikt Zimmermann

Berlin, 2013

Referees:

1. Prof. Dr. Natalia Kliewer
2. Prof. Dr. Catherine Cleophas

Date of Disputation: 10 January 2014

Declaration of Authorship

Except where reference is made in the text of this thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis presented by me for another degree or diploma. No other person's work has been used without due acknowledgement in the main text of the thesis. This thesis has not been submitted for the award of any other degree or diploma in any other tertiary institution.

Signed: ____________________

Date: ____________________

Abstract

While revenue management (RM) frequently claims to provide a competitive advantage, its long-term competitive effects have hardly been explored. Current research on RM under competition has started to examine competitive interactions in RM with the help of game theory. However, single-stage game analyses predominate, although for most practical applications of RM, a multi-stage view would be more appropriate. In fact, a single-stage view of RM under competition is dangerous, since—similarly to the Iterated Prisoner's Dilemma (IPD)—the single-stage optimum can prove disastrous in the repeated game. This thesis provides a formal model of two competing service providers implementing RM in an environment with price-elastic demand who face each other repeatedly. Based on approximations to the solutions of the single-stage game, we present a heuristic to transfer strategies from the IPD to the repeated RM game. In this thesis, we use both computational results derived using a stochastic simulation system and mathematical analyses of a simplified version of the game to analyze repeated-game strategies. We investigate the influence of capacity restrictions, forecasting techniques and observation errors on the performance of multi-stage strategies. Our results stress the relevance of a multi-stage view of revenue management games. Furthermore, our results show the strong impact of observation errors and the resulting importance of robust strategies.

Zusammenfassung

Obwohl Revenue Management (RM) häufig mit einem Wettbewerbsvorteil verbunden wird, wurden die langfristigen Effekte von RM im Wettbewerb bisher kaum untersucht. In der aktuellen Forschung wird zwar Spieltheorie genutzt, um im RM die Wechselwirkungen von Wettbewerbern zu untersuchen. Allerdings wird dabei fast ausschließlich ein einmaliges Spiel betrachtet, obwohl in den meisten praktischen Anwendungen eine Darstellung als wiederholtes Spiel angebracht wäre. Ähnlich wie beim Gefangenendilemma (IPD) ist eine Betrachtung als einmaliges Spiel sogar gefährlich, da die Wiederholung der optimalen Strategien des einmaligen Spiels im wiederholten Spiel zu sehr schlechten Ergebnissen führen kann. In dieser Arbeit stellen wir ein Modell für zwei RM-nutzende Dienstleister auf, die sich wiederholt im Wettbewerb um preissensitive Kunden gegenüber stehen. Aufbauend auf Annäherungen an die Lösungen des einmaligen Spiels präsentieren wir eine Heuristik, mit der Strategien vom IPD auf das wiederholte RM-Spiel übertragen werden können. Wir verwenden sowohl Simulationen mit einer stochastischen Simulationsumgebung als auch mathematische Analysen eines vereinfachten Spiels, um Strategien des wiederholten Spiels zu analysieren. Dabei untersuchen wir den Einfluss von Kapazitätsbeschränkungen, Prognosemethoden und fehlerhaften Beobachtungen auf das Abschneiden der Strategien. Unsere Ergebnisse unterstreichen die Bedeutung einer Betrachtung von Wettbewerb zwischen RM-Betreibern als wiederholtes Spiel. Weiterhin verdeutlichen unsere Resultate die starken Auswirkungen von Beobachtungsfehlern und daraus resultierend die Wichtigkeit robuster Strategien.

Contents

1 Motivation **11**

2 Literature Review: RM and Game Theory **15**
2.1 Revenue Management in a Monopoly . . . 15
2.1.1 Forecast . . . 16
2.1.2 Optimization . . . 19
2.1.3 Inventory . . . 23
2.1.4 Summary and Implications . . . 23
2.2 Game Theory . . . 24
2.2.1 Basics . . . 24
2.2.2 The Prisoner's Dilemma . . . 26
2.2.3 Summary and Implications . . . 31
2.3 Revenue Management under Competition . . . 31
2.3.1 Amendments to the Monopoly Model . . . 32
2.3.2 Game Theory in Revenue Management . . . 34
2.3.3 Summary and Implications . . . 37
2.4 Simulation . . . 37
2.4.1 Simulations in Revenue Management . . . 38
2.4.2 Simulations in Game Theory . . . 39
2.4.3 Summary and Implications . . . 39

3 Research Gap: Repeated RM Competition **40**

4 RM Competition as a Repeated Game **44**
4.1 Modeling Revenue Management under Competition . . . 46
4.1.1 Demand . . . 46
4.1.2 Forecast . . . 47
4.1.3 Optimization of Single Stage . . . 53
4.2 Multi-stage Strategies . . . 55
4.2.1 Heuristic to Adapt Strategies from the IPD to the Repeated RM Game . . . 56
4.2.2 Strategies . . . 58
4.2.3 Properties of Successful Multi-stage Strategies . . . 60
4.3 Simulation Environment . . . 61

4.4 Simulation Experiments . . . 66
4.4.1 Psychic Forecast . . . 68
4.4.2 Standard Forecasts . . . 77
4.5 Summary . . . 86

5 The Repeated RM Game as a Markov Process **89**
5.1 Mathematical Basics . . . 92
5.1.1 Stochastic Processes . . . 92
5.1.2 Tensor Products . . . 97
5.2 Construction of the Markov Chain of the RM game . . . 100
5.2.1 Prerequisites . . . 100
5.2.2 General Markov Strategies . . . 101
5.2.3 Reactive Strategies . . . 102
5.2.4 Observation Errors . . . 103
5.2.5 Reputation . . . 104
5.3 Payoff Structure in the Markov RM game . . . 106
5.3.1 Assumptions . . . 107
5.3.2 Payoff Calculation . . . 108
5.3.3 Examples . . . 108
5.4 Limiting Behavior of Repeated-Game-Strategies . . . 111
5.4.1 Strategies . . . 111
5.4.2 Stationary Measures . . . 114
5.4.3 Erroneous Threshold Products . . . 118
5.4.4 Example . . . 120
5.5 Evolution of Markov Strategies . . . 132
5.5.1 Experimental Setup . . . 133
5.5.2 Simulation Results . . . 134
5.6 Zero-Determinant Strategies . . . 137
5.6.1 Prerequisites . . . 139
5.6.2 Existence of Extortionate Strategies . . . 143
5.6.3 Example . . . 150
5.7 Dynamic Pricing . . . 153
5.7.1 Construction of the Markov Process . . . 154
5.7.2 Strategies . . . 157
5.8 Summary . . . 160

6 Conclusion **163**
6.1 Summary . . . 163
6.2 Practical Implications . . . 166
6.3 Limitations and Future Research . . . 167

References **169**

Symbols **184**

Acronyms **189**

Glossary **191**

List of Tables

4.1 Prices used in the simulation 67

5.1 Prices used by Isler and Imhof (2008) 110
5.2 Prices used in the simulation in Section 4.4 111

List of Figures

4.1 Booking class graph 51
4.2 Normal form of the prisoner's dilemma 56
4.3 Evolution of reputation 58
4.4 Flowchart of the revenue management process in REMATE 63
4.5 Revenue in a monopoly 67
4.6 ALLD and ALLC 69
4.7 TFT for $\varepsilon = 0$ 70
4.8 TFT for $\varepsilon = 0.1$ 70
4.9 GTFT for range of generosity probabilities, $\varepsilon \in [0, 0.2]$ and CAP=30 . . . 71
4.10 GTFT for $\varepsilon = 0.1$ 72
4.11 CTFT for $\varepsilon = 0.1$ 73
4.12 PAVLOV for $\varepsilon = 0.1$ 74
4.13 Robust strategies against TFT for $\varepsilon = 0.1$ 74
4.14 MATCH for $\varepsilon = 0.1$ 75
4.15 UNDER for $\varepsilon = 0.1$ 76
4.16 UNDER vs. various repeated game strategies for $\varepsilon = 0.1$ 77
4.17 Independent demand forecast d^I 79
4.18 Hybrid demand forecast d^H 80
4.19 Dependent demand forecast d^K 81
4.20 TFT using dependent demand forecast d^K with $\varepsilon = 0.1$ 82
4.21 GTFT using dependent demand forecast d^K with $\varepsilon = 0.1$ 83
4.22 CTFT using dependent demand forecast d^K with $\varepsilon = 0.1$ 83
4.23 PAVLOV using dependent demand forecast d^K with $\varepsilon = 0.1$ 84
4.24 Standard forecast vs.psychic forecast 84
4.25 Standard forecast vs.irrational strategies with $\varepsilon = 0.1$ 85

5.1 Evolutionary success of GTFT for different γ and $\varepsilon = 0.1$ 135
5.2 Average evolution with low temptation and $\varepsilon = 0.1$ 136
5.3 Average evolution per outcome with low temptation and $\gamma = 0.2, \varepsilon = 0.1$ 136
5.4 Average evolution with high temptation and $\varepsilon = 0.1$ 137
5.5 Average evolution per outcome with high temptation and $\gamma = 0.2, \varepsilon = 0.1$ 138

1 Motivation

Finding the right price for a perishable product is no easy task. A seller has to gather enough information to balance the gain earned from a sale against the possible revenue from selling the same resource at a different occasion, at the risk of not selling the product at all before it turns worthless. As additional difficulties, a firm often has to face capacity constraints and uncertain demand, when determining the optimal selling price. Revenue Management (RM) or yield management is the field encompassing the techniques dedicated to solving this problem. Put more eloquently for the case of the airline industry, airline RM has been described as in *The Art of Managing Yield* (1987) "selling the right seats to the right customers at the right prices".

While research on RM had started earlier (e.g. Littlewood, 1972), it was the deregulation of the U.S. airline market in 1978 that enabled the transfer from theory to practice (Smith, Leimkuhler, & Darrow, 1992). Since then, RM has been applied to industries as diverse as hotels (Choi & Mattila, 2004), rental cars (Geraghty & Johnson, 1997), retail (Vinod, 2005), cruise lines (Ladany & Arbel, 1991) and advertising (Kimms & Müller-Bungart, 2007).

Given the broad distribution of RM in practice, it comes as no surprise that research on the topic has evolved to solve increasingly realistic and complex problems. Early RM research modeled customers as independent demand in a monopoly with a single resource (Littlewood, 1972). In the last decades, researchers have contributed new methods accounting for network effects (Bertsimas & Popescu, 2003), dependent demand (Weatherford & Ratliff, 2010) and new fare structures (Fiig, Isler, Hopperstad, & Belobaba, 2009).

Despite the progress of RM in many areas, "RM competition is not well understood and practically all known implementations of RM software and most published models of RM do not explicitly model competition" (Martínez-de Albéniz & Talluri, 2011). Nevertheless, the influence of competition on demand is so important that it cannot be ignored in practice. Even without a system yielding optimal control in the presence of competition, service providers react to their competitor's prices in one way or another. As a consequence, a large part of the providers' price dispersion is due to competition (Hayes & Ross, 1998).

Such a behavior leads to lower prices (Kwoka & Shumilkina, 2010; Stavins, 2001) and lower occupancies (Kalnins, Froeb, & Tschantz, 2010) under competition, which is a very costly combination. In fact, the airline industry in the U.S. alone faces losses of several hundred million dollars every year due to fare wars of competing carriers (Morrison, Winston, Bailey, & Carlton, 1996). Additionally, the rise of the internet has led to a higher transparency of prices in recent years. Nowadays, the possibility to compare prices within seconds has effectively eliminated the search costs for the customer. As a result, prices in RM industries have dropped, especially in competitive markets (Orlov, 2011).

Owing to the importance of competition in Revenue Management, recent years have seen researchers increase efforts to include "competitive awareness" into RM systems (Ratliff & Vinod, 2005). However, the presence of competition leads to challenges in both forecasting and optimization, the cornerstones of any RM system.

The forecast provides the basis of a RM system. Even in a monopoly, producing accurate and reliable information is not a simple task, since the behavior of each potential customer can only be partially observed. If a customer buys a ticket for a given price, the selling firm does not know whether he would have paid for a more expensive alternative. The situation gets even worse for potential customers that end up not purchasing a ticket at all, or that have bought a competitor's product. All this missing information has to be reconstructed from past observations with the help of assumptions to the demand model. However, this usually means a loss in quality of the forecast. Attempts to account for competitor influences in the forecast present a way to improve forecast quality in a competitive situation. However, this has to be balanced against the considerable increase in complexity that this approach may bring. Modern forecasting systems have raised their level of complexity quickly in order to cope with the challenges of the monopoly forecasting problem. Simply adding more explanatory variables to such models may hurt the performance more than it helps (Bartke, Cleophas, & Zimmermann, 2013). On the other hand, completely ignoring competitive influences means ignoring the basis of potential customers' decision.

The optimization part of a typical RM system is closely related to optimal control theory. It is commonly assumed that it is possible to control the demand flow by varying the set of classes and prices in the market for any one firm. This holds no longer true in the presence of competitors. With competitors offering comparable products, a provider treating the demand flow as a result of its price-setting actions risks losing customers to the competition. Instead of forcing customers to pay higher prices, it may end up forcing them away. In fact the interaction between competing firms may result in the optimization problem being closer related to game theory than optimal control theory.

Game theory has been used to analyze competitive interactions between firms for a very long time (Bertrand, 1883; Cournot, 1838). Unfortunately, simple models of price or quantity competition fail to model the RM problem sufficiently well, since in typical RM applications, a firm has to face capacity constraints, uncertain demand spread over time and more (Kalnins et al., 2010). Due to this shortcoming of traditional game theory, recent years have seen an increased interest in research in game theory applied to revenue management (e.g. Bertsimas & Perakis, 2006; Kwon, Friesz, Mookherjee, Yao, & Feng, 2009; Lin & Sibdari, 2009; Martínez-de Albéniz & Talluri, 2011; Mookherjee & Friesz, 2008; Netessine & Shumsky, 2005; Perakis & Sood, 2006; Simon, 2007; Zhao, 2003).

While modeling RM under competition as a game, it should be noted that in practice, RM is used to yield optimal pricing and capacity decisions for reoccurring events, e.g. flights in the airline industry or overnight stays in the hotel industry. Thus, RM under competition should be modeled as a repeated game with an infinite horizon (Isler & Imhof, 2008). However, due to the complexity of the RM problem under competition, the only analyses of similar repeated games treat simple price or quantity competition, while the research on the RM problem focuses almost exclusively on the single-stage game. This is consistent with the theory of RM in a monopoly, which deals with finding the best solution to a single sales period for one or more products. In this case, treating more than one sales period is unnecessary, since a repetition of the optimal solution yields the best strategy for any repetitions of such events. However, with the addition of competitive effects, a long-term aspect of the problem emerges. When two or more service providers face each other repeatedly for the same problem, they may learn from previous encounters and adapt their strategy accordingly. In particular, threats can enforce collusion in repeated games and thus help avoid price wars (Chamberlin, 1929).

While overt collusion is usually illegal, tacit collusion represents simply a set of strategies that avoids aggression out of fear of retaliation (Feuerstein, 2005). In practice, some weakened forms of tacit collusion can be found as well as forms of non-cooperative behavior (Fischer & Kamerschen, 2003). Researchers stress that the competitive behavior of service providers tends to be “far from collusive” (Brueckner & Spiller, 1994). It remains unclear whether this is due to an imperfect implementation of purely non-cooperative behavior or due to practitioners being alert to the danger of price wars.

Unfortunately for firms trying to prevent fare wars, successful strategies in the repeated RM game have hardly been researched at all. Inspiration for such strategies can be drawn from the similarity of the repeated RM game to the much simpler Iterated Prisoner’s Dilemma (IPD) (Isler & Imhof, 2008). In the IPD, each player faces a dilemma during each stage. He can either follow the cooperative solution and risk being exploited, or he can choose the non-cooperative option and risk a very low payoff if his competitor does the same. Due to its simplicity and applicability to many practical more complex problems, the IPD has seen a host of research examining the success of various strategies (e.g. Axelrod, 1984; Imhof, Fudenberg, & Nowak, 2007; Nowak & Sigmund, 1993; Press & Dyson, 2012).

However, concentrating only on game-theoretic methods for the analysis of RM competition provides pitfalls. A typical game theoretic treatment of the problem relies on the assumption that each carrier acts rationally by following game-theoretic reasoning. With current RM systems using monopoly methods, such assumptions seem hardly realistic. In this thesis, we will analyze the long-term behavior of standard RM strategies as well as of strategies inspired by game theory in a competitive environment. We will develop a heuristic to adapt strategies from the IPD to the RM game, which will be evaluated via simulations in a realistic setting and via thorough mathematical analysis in a simplified case.

First, we will use Chapter 2 to give an overview of existing research on RM, game theory and simulations and put it in context to our problem. Based on our analysis of the literature, we will identify research gaps and outline our course of action in Chapter 3. In Chapter 4, we will use a simulation approach to the problem, while in Chapter 5, we will perform mathematical analysis on a simplification of the repeated RM game. Finally, in Chapter 6 we will present a conclusion to our work, where we highlight our findings as well as the limitations of this thesis and outline future research.

2 Literature Review: RM and Game Theory

In this chapter, we will give an overview of the literature in various fields that help analyzing the revenue management problem under competition. In Section 2.1, we review the large body of research covering the classic case of Revenue Management (RM) in a monopoly. In Section 2.2, we take a look at the game theoretic approaches to competitive situations. Next, we present research combining both aspects in Section 2.3. Finally, in Section 2.4, we document research on the use of simulations to answer similar questions.

2.1 Revenue Management in a Monopoly

RM or yield management deals with the problem of selling perishable goods with negligible unit costs under capacity restrictions so that revenue is maximized. As an example, the airline industry, which has pioneered RM, faces relatively fixed costs in the short term, so that maximizing revenue is sufficient for maximizing profit. Here, a typical RM decision consists of whether to accept a booking request in a cheap booking class or not. This has to be weighed against the possibility of selling this seat for a higher price to a high-value customer arriving later in the booking horizon. On the other hand, when the flight leaves, the value of any unsold seats perishes completely.

A simple form of RM is treated in the so-called newsvendor or newsboy problem. In the newsvendor problem, a service provider tries to find the optimal capacity allocation to a perishable good with fixed costs, which is sold at a fixed price over a single time period to an uncertain amount of customers. This problem has occupied researchers for over a century. Already in 1888, Edgeworth analyzed how much money a bank should store in order to fulfill stochastic withdrawal requests. This is a classic inventory problem, which we will see quite often in slight variations in the following sections. If the decision maker allocates too much capacity for the withdrawal requests, he suffers losses in revenue, since he could not use all of his capacity to earn revenue. However, if he stores insufficient capacity, he suffers a loss as well, in this case from having to raise money in short notice to satisfy the withdrawal requests. Several decades later, Arrow, Harris, and Marschak (1951) gave a treatment of the newsvendor problem that inspired a lot of related research. The interested reader can find a summary of the literature on the newsvendor problem in the book of Silver, Pyke, Peterson, et al. (1998). The very simple form of RM in the newsboy problem lends itself well to a thorough analysis and allows for many extensions as outlined by Khouja (1999). However, its simplicity has limited the range of applications of its solution concepts to real life. Especially the restriction to a single time period prevents a transfer of solutions of the newsvendor problem to industries such as the airline, hotel or car rental businesses.

Since the 1950s, airlines have tried to find ways to deal with the uncertainty of demand in the form of cancellations and no-shows by developing overbooking policies (Beckmann, 1958). In contrast to this, the so-called seat inventory control problem of optimizing the mix of booking classes, so that revenue is maximized, has only seen attention from researchers and practitioners since the 1970s. For this problem, decision makers have to balance their prices so that they sell a fixed capacity of their resource in a way that they neither ignore low-value demand nor accidentally shut out late-arriving high-value customers. As Gallego and van Ryzin (1994) put it, "yield management is an attempt to 'synthesize' a range of optimal prices from a small, static set of prices in response to a shifting demand function". In the following, we will concentrate on the seat inventory control problem when discussing revenue management.

As a field of research of its own, RM has been growing ever since Littlewood (1972) published the first RM optimization technique for the seat inventory control problem. It has also seen tremendous use in practice since British Overseas Airways Corporation implemented Littlewood's techniques to manage their mix of full and discounted fares. This has only increased after American Airlines used RM successfully in order to fight off competition after the deregulation of the U.S. airline market (*The Art of Managing Yield*, 1987). Back then, American Airlines famously introduced new discounted fares managed by inventory management techniques, so that they could match the market's low prices without diluting their whole inventory.

There are several excellent articles providing an overview over the state of the art of revenue management. McGill and van Ryzin (1999) examine the evolution of RM from its very beginnings until the end of the millennium. For a fully comprehensive account of RM methods, we refer the reader to the book of Talluri and van Ryzin (2004b). Dana (2008) and Chiang, Chen, and Xu (2007) give an account of more recent research opportunities, while Pölt (2011) and Cross, Higbie, and Cross (2011) focus on the challenges imposed on practitioners by recent changes of the market environment.

2.1.1 Forecast

In revenue management, information is the key to success. Great effort is put into forecasting demand to obtain an accurate estimate of the market situation. An overview of common forecasting techniques used in RM can be found in the review articles of Weatherford and Kimes (2003) and Cleophas, Frank, and Kliewer (2009a).

The main problem in forecasting demand is caused by the censoring of observations via the firm's availability control, so that customer behavior can hardly ever be observed in an unbiased way. Unconstraining is the process of reconstructing the true demand process from the censored observations (for an overview see Guo, Xiao, & Li, 2012). This is similar to problems in signal processing (e.g. Foxlin, 1996; Grewal, Henderson, & Miyasako, 1991), where a standard solution approach consists of applying the Kalman filter invented by Kalman et al. (1960). As a consequence, Bartke (2013) adapted the Kalman filter to the unconstraining problem in RM.

The unconstraining process depends heavily on the customer model, which represents the core of every forecasting system. Customer models used in early RM systems relied on a number of assumptions, many of which have been weakened in recent years. As a consequence, unconstraining has become ever more challenging and even evaluating the forecast performance has become a difficult task (Cleophas, 2009).

Since an oversimplification of the customer model may significantly decrease forecast accuracy, it might be tempting to model customer behavior as accurately as possible. However, Bartke et al. (2013) showed that a sophisticated customer model can lead to a high degree of complexity. This can lead to instability of estimation results as well as to decreases in performance due to usability problems.

It is also important to keep in mind that a misspecified demand model may lead to the calculation of incorrect optima or even to a chaotic behavior already for very simple models (Bischi, Chiarella, & Kopel, 2004). In order to avoid misspecification, it may be beneficial to use less assumptions on the underlying functional relationship between product properties and demand. Instead of modeling demand as a parametric function of product properties, where unconstraining consists of estimating the parameters, a non-parametric demand model may be used. In fact, the non-parametric approach is applied frequently in the industry (Weatherford & Kimes, 2003). However, non-parametric models should be used with caution. Compared to the correctly specified parametric model, a non-parametric demand model represents a loss of information and thus results in a loss of revenue (Besbes & Zeevi, 2009).

In Littlewood's (1972) model, customers were supposed to

1. request a ticket on a flight instead of a pair of Origin and Destination (O&D),
2. arrive ordered by their willingness-to-pay with low-value customers first and high-value customers last,
3. arrive sequentially without simultaneous arrival of groups,
4. never cancel their ticket and always show up for the flight,
5. demand a single product independently of the availability of substitutes and
6. request a ticket exactly once, independently of offers at other times in the booking horizon.

Assumptions 1 – 4 are constraints of the early optimization techniques in revenue management, rather than conditions met in the real world or limitations of the available forecasts. As soon as optimization techniques improved, these requirements could be relaxed or dropped (McGill & van Ryzin, 1999).

In contrast to the above, assumptions 5 and 6 of statistically independent demand per booking class and time were fulfilled in the early days of revenue management. This has changed recently, which has prompted significant changes for RM forecasting techniques. In the following, we will outline challenges and solution concepts related to the loss of assumptions 5 and 6.

Dependent demand In previous years, service providers ensured the validity of the assumption of customers' demand targeting a single class by attaching a set of restrictions to each booking class. For example in the airline industry, in order to keep high-value demand targeting expensive classes, cheaper booking classes were typically offered only in combination with a longer duration of stay, thus making it inconvenient for business travelers. Using appropriate restrictions, airlines were able to separate high-value demand from low-value demand. In recent years, the rise of low-cost carriers triggered a change in the industry. The success of low-cost carriers using simple fare structure with hardly any restrictions caused a decrease in use of restrictions in the whole airline industry. Therefore, demand cannot be separated as efficiently anymore (Zeni, 2007).

Ignoring the substitution effects of demand of similar classes leads to an effect called Spiral Down. If customers can substitute high-value classes for low-value classes, the airline will observe only low-value sales whenever both options are available to the customer. Failing to account for potential sales in the high-value classes leads to a lower forecast and consequentially lower protected seats for the high-value classes. Thus, in the next departure, both classes are more likely to be available at the same time and the cycle starts over anew. This feedback loop can ultimately lead to drastically decreasing revenues (Cooper, de Mello, & Kleywegt, 2006).

In so-called Q-Forecasting, customers are assumed to always buy the cheapest class available, if they can afford it (Belobaba & Hopperstad, 2004; Cléaz-Savoyen, 2005). In this model the willingness to pay is exponentially distributed, so that demand for every price level can be deduced from observing the purchases in the cheapest available class. In the hybrid forecasting technique, the forecast consists of a linear combination of the Q-Forecasting customer model and the independent demand model (E. A. Boyd & Kallesen, 2004). This approach has been extensively studied (Reyes, 2006; Weatherford & Ratliff, 2010) and is also widely used in practice (Wishlinski, 2006). In a recent contribution using the hybrid demand model, Bartke (2013) calculated theoretical boundaries for the quality of estimates and adapted variations of the Kalman filter to provide a superior estimation technique for hybrid demand.

Some modern RM forecasting systems adopt a more customer-centric approach, assuming that each customer will consider a set of appropriate classes and then choose the most attractive option (Akçay, Natarajan, & Xu, 2010; Cirillo & Hetrakul, 2011). This process can be modeled by discrete choice models. A comprehensive overview of general discrete choice models is given in the book of Train (2003), while Shen and Su (2007) focus on the application of these models in revenue management. Another approach to modeling and estimating such complex customer behavior has been put forward by Winter (2010), who represented a firm's product in an ordered graph, on which demand flowed depending on the availability situation of the firm's products.

Strategic customers The task of modeling customers' decision process as longer-term process, during which they might compare prices and wait for bargains, has been tackled first by Zhou, Fan, and Cho (2005) and C. Anderson and Wilson (2003). An overview on the advances of modeling these so-called strategic customer behavior can be found in Shen and Su's (2007) overview on customer modeling. However, while more recent articles on the subject refine the customer modeling and explore the service provider's optimal policy, they do not treat the problem of estimating and forecasting demand. The inclusion of strategic customers further expands the choice set for each customer, which may complicate the estimation procedure significantly.

2.1.2 Optimization

In RM, the optimization step is used to find the revenue-maximizing strategy to sell the remaining resources given the forecasted demand. We distinguish two different approaches, depending on the strategic variable chosen. Quantity-based RM uses a set of booking classes with fixed prices for each travel opportunity. The result of the RM process is an availability situation, i.e. a subset of booking classes offered to customers. In this case, the optimization consists mainly of the decision of how many seats of each class to offer in a particular situation. Instead of choosing quantity as a control variable and keeping prices fixed, it is also possible to keep quantity fixed and optimize prices. This is called price-based Revenue Management or Dynamic Pricing. We will first focus on quantity-based RM, before we will turn toward Dynamic Pricing.

Quantity-based revenue management

The first optimization technique in RM was published by Littlewood (1972). In his seminal paper, Littlewood described a simple and intuitive rule to decide whether to accept a booking or not for the two fare class problem. A request for the lower class is accepted if the revenue earned by this sale is greater or equal than the expected revenue for this seat if it is protected for the higher class. This new method was necessitated by a change of fare structures and enabled by increasing data quantity and quality due to the introduction of computer reservation systems, two factors that have proved decisive for a great part of innovations in Revenue Management. Given that Littlewood's rule had been implemented by British Overseas Airways Corporation (BOAC), this was not only the first theoretical work in RM, but also its first practical application.

As we have mentioned in the forecasting section, Littlewood had to impose several constraints on the demand process in order to find an optimal solution. For the convenience of the reader, we reproduce these assumptions here. Customers were supposed to

1. request a ticket on a flight instead of a pair of O&D,
2. arrive ordered by their willingness-to-pay with low-value customers first and high-value customers last,
3. arrive sequentially without simultaneous arrival of groups,
4. never cancel their ticket and always show up for the flight,
5. demand a single product independently of the availability of substitutes and
6. request a ticket exactly once, independently of offers at other times in the booking horizon.

Using the same assumptions, Littlewood's rule was later applied to more than two classes by Belobaba's (1987; 1989) Expected Marginal Seat Revenue (EMSR) heuristics, the two best-known versions being called Expected Marginal Seat Revenue—Version a (EMSRa) and Expected Marginal Seat Revenue—Version b (EMSRb). In the first version, EMSRa, Littlewood's rule is applied pairwise for all adjacent classes. Thus, this results in the amount of seats to protect for the higher class of the pair, if there were no other classes. The final protection levels are then calculated by summing over the pairwise protection levels. EMSRa has been shown to be overly conservative, leading to too many protected seats for high-value classes (Talluri & van Ryzin, 2004a, pp.47–48). In order to mitigate this, EMSRb aggregates demand and prices of lower classes beforehand in order to create a virtual class representing demand for all classes lower than a given class. Littlewood's rule is then applied for each class compared with the representation of all lower classes. This procedure results directly in protection levels that can be used for the particular flight.

Optimal policies for the single-leg problem have since been developed independently by Curry (1990), S. L. Brumelle and McGill (1993) and Wollmer (1992). These controls often differ significantly from the EMSR heuristics, but the resulting revenue has been shown to be very similar for typical airline demand distributions (Wollmer, 1992). Thus, the EMSR heuristics have spurred immense interest among researchers (e.g. Gallego, Li, & Ratliff, 2009; Walczak, Mardan, & Kallesen, 2009) as well as practitioners (Netessine & Shumsky, 2002; Weatherford, 2004) because of their appealing simplicity.

Network effects The above approaches have only treated the single-leg problem, in which every customer requests a single resource, e.g. a seat on a flight in the airline case. This is an oversimplification, since an airline usually has to manage a network of multiple flights. Since customers traveling on connecting flights may use more than one resource per request, control strategies on different flights may affect each other. Rejecting a request for a connection of multiple flights on one flight results in losing the request on all other flights as well. Thus, the optimal control of a network of flights does not consist of independent single-flight solutions. In fact, treating the network problem as a collection of single-resource problems may result in significant loss in revenue (Williamson, 1992). The first formulation of the RM problem for multiple resources was given by Glover, Glover, Lorenzo, and McMillan (1982). In the following, there have been many proposed solutions, for which we refer the reader to the overview given by McGill and van Ryzin (1999) and Talluri and van Ryzin (2004b, pp.81–128). Unfortunately, it is well known that "in the network case, exact optimization is for all practical purposes impossible" (Talluri & van Ryzin, 2004b, p.83), so that researchers and practitioners have to use simplifying assumptions and develop heuristics to make a solution feasible. In particular, two heuristics to the network problem have proved very successful in practice, because they enabled the use of the established single-leg techniques. Both Displacement Adjusted Virtual Nesting (DAVN) (Bertsimas & De Boer, 2005) and Dynamic Programming Decomposition (Talluri & van Ryzin, 2004a, p.107) solve a linear program to obtain static valuations of each leg in the network. These valuations are then used to solve the interdependence of the different flights and reduce the problem to the single-resource case.

Arrival order Lee and Hersh (1993) describe demand for each fare class as a Poisson process with time-varying intensity. This way, arbitrary arrival patterns of demand are possible. This formulation transforms the RM problem into a Markov Decision Problem, which can be solved via Dynamic Programming. For an overview on Dynamic Programming, we refer the reader to the book of Bertsekas (1995). Dynamic Programming was invented by Bellman, who noted that "[a]n optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" (Bellman, 1954). This is called Bellman's Principle of Optimality and reduces the problem to solving smaller and simpler subproblems. However, because of the connection of Bellman's Principle of Optimality with the Markov property of the demand process, using Dynamic Programming imposes restrictions on the modeling of demand, since the demand process needs to possess the Markov property as well. In fact, researchers have concentrated on an even smaller subgroup of discrete Markov processes, namely the Poisson and Bernoulli processes. The memorylessness of these two processes allows for stochastically independent increments of the demand process, which is a feature that fits the intuition of a demand process. This has led to the almost exclusive use of Poisson and compound

Poisson processes for continuous time and Bernoulli processes in discrete time in the RM literature (McGill & van Ryzin, 1999).

Batch orders Lee and Hersh (1993) incorporate batch orders into their dynamic programming approach. Batch orders are requests for multiple seats on the same flight, that can only be accepted or rejected as a whole. They show that in this case, the optimal policy may not have the same structural properties as in the single-order case. Other researchers have treated this problem as well (S. Brumelle & Walczak, 2003; Örmeci & Burnetas, 2004). They all show that the optimal policy's structure may deviate from the case without batch orders.

Cancellations Coping with customers' cancellation behavior inspired overbooking optimization, one of the first RM methods. In seat inventory management however, many researchers ignore cancellations in order to get feasible solutions of the optimization problem. Subramanian, Stidham, and Lautenbacher (1999) discuss methods to include cancellations in a dynamic programming approach. They observe that the state space grows prohibitively large, as soon as cancellation behavior is modeled as class-dependent. Therefore, they propose two heuristics that allow for a treatment of cancellations with class-specific cancellation rates in a dynamic optimization approach. Bertsimas and Popescu (2003) incorporate cancellations into various network RM techniques.

Dependent demand In recent years, advances in forecasting led to the development of dependent demand models. These new developments call for appropriate optimization techniques, since using independent demand optimization with dependent demand forecasts leads to suboptimal results (Gallego et al., 2009). Weatherford and Ratliff (2010) give an overview over the emerging optimization models, that can cope with dependent demand. Besides the multitude of techniques that require completely new structures in order to calculate an optimal policy in this setting, Fiig et al. (2009) describe a way to transform dependent demand so that traditional optimization techniques for independent demand may be used.

Strategic customers Dealing with strategic customers is a problem that researchers have started to approach in the last few years. In their survey article on customer modeling, Shen and Su (2007) include an overview on strategic customers in revenue management. They review both modeling and optimization approaches to the problem.

Dynamic pricing

To the best of our knowledge, the first work on Dynamic Pricing was done by Kincaid and Darling (1963). In their setting, a firm tries to find optimal prices for a single product. Demand arrives as a Poisson process with an arbitrary willingness-to-pay distribution. They find sufficient conditions for the existence of an optimal solution but do not present a method to find the optimal control.

Gallego and van Ryzin (1994) later studied a similar problem, where demand arrives as a Poisson or Compound Poisson process. They find an exact solution for the simple, yet insightful case of exponentially distributed willingness-to-pay. They also demonstrate that the optimal pricing policy can be approximated by a time-independent pricing policy. Additionally, they describe asymptotically optimal policies for several extensions of the problem, including seasonality and cancellations.

2.1.3 Inventory

For Dynamic Pricing, the structure of the optimal control is simply the price calculated by the optimization for each product. For quantity-based RM however, the form of the control may depend on the optimization technique. The simplest, albeit inefficient, form of control consists of a fixed allocation of capacity to each resource and is called partition control. Nested partition control is an optimal control that extends the partition control (Talluri & van Ryzin, 2004a). Virtual nesting represents a possibility to translate the nested partition control to the network case (Smith & Penn, 1988). Bid prices represent an elegant control, since they are easily extended to the network case. Talluri and van Ryzin (1998) have shown that although bid prices do not describe an optimal policy, they are asymptotically optimal.

2.1.4 Summary and Implications

In the past decades since its inception, RM has been a growing field of research. In this section, we have concentrated on research on RM in a monopoly, which still encompasses the overwhelming majority of research on RM. Starting with Littlewood's (1972) simple model, research on RM in a monopoly has led to the development of forecasting and optimization techniques that can deal with far more realistic customer behavior. This state of the art constitutes a fruitful basis for researchers who want to analyze competitive effects in RM (see Section 2.3.1). Similarly, the RM model in this thesis will profit from the research presented in this section, as we will rely on forecasting and optimization techniques described above.

2.2 Game Theory

Classical RM deals with the problem of a single decision maker trying to maximize revenue on a given market. As soon as more than one decision makers act on the same market, their decisions may influence each other. Thus, solutions of classical RM need not be applicable in a competitive setting. For such a problem, game theory provides a large variety of tools that help to understand the problem as well as to find solutions.

Game theory deals with the process of decision making of rational interacting individuals. Given the broad range of this topic, it comes as no surprise that researchers have tackled game theoretic problems for a long time.

Inspired by observations of competition of mineral water producers, Cournot (1838) was the fist researcher to analyze optimal strategies in an oligopoly. In his model, firms simultaneously set quantities of homogeneous goods, while prices are determined by the market as a linear function of the quantities. He showed that the optimal strategy results in a price higher than the marginal costs. In a review of Cournot's book, Bertrand (1883) argued that instead of setting quantities, firms would much rather set prices. Thus, in this setting, firms would set prices and sell as many quantities as the demand in the market requested. Contrary to the results of Cournot, he showed that the optimal strategy even in a duopoly consists of setting the marginal costs. His results imply that firms in an oligopoly do not have any market power, which deviates from the observed behavior in real life. Edgeworth (1897) addressed this discrepancy by introducing a capacity constraint. He showed that when sellers only have a limited amount of products to sell, the results of Bertrand do no longer hold. In fact he showed that there may exist no deterministic optimal strategy at all.

While researchers such as Cournot, Bertrand and Edgeworth intuitively used concepts of modern game theory, its foundations were laid by the work of von Neumann (1928), who started to formalize concepts of game theory.

2.2.1 Basics

An overview on the foundations of game theory is given by Osborne and Rubinstein (1994). Nevertheless, due to the great popularity of the field and the abundance of subtopics, we have not found an all-encompassing reference to the field. Additionally, many concepts of game theory will be left untouched by us. Therefore, in the following we will briefly present the necessary tools and ideas for our purposes.

Game theory has been applied to a great variety of decision making problems. We make no claim of completeness for our presentation. On the contrary, we will try to focus only on the facets necessary for understanding the problem of oligopolistic competition between service providers using revenue management.

Structure and Properties of a Game

To present a decision making problem in the game-theoretic framework, it is necessary to specify a payoff function for each player as well as a set of strategies available to him. Deterministic strategies are called pure strategies and have been treated already by the first researchers of game theory. Choosing probabilistically between pure strategies yields mixed strategies, which represent a strict generalization of pure strategies. A strategy profile is a set of strategies for each player in the game.

Dynamic games introduce a time dimension. In a dynamic game, at least one player can choose a strategy that conditions its actions on previous actions of himself or other players. If the game is set in continuous time, this is called a differential game.

There are two well-studied different information structures that influence the structure of strategies for dynamic games. If players can observe other players' previous actions and adapt their strategy accordingly, we speak of a closed-loop game. If players cannot observe other players actions and thus have to commit themselves to a strategy beforehand, this is called an open-loop game. Open-loop games are often easier to analyze compared to closed-loop games, which has led researchers to study open-loop games, even when a closed-loop representation would have been more appropriate (Fudenberg & Levine, 1988).

A dynamic game can be played either with a finite or an infinite horizon, depending on whether players believe that the game will be played a finite or infinite amount of periods. Setting a finite horizon is not necessarily the same as playing the game only a finite amount of periods. As long as the players do not know that the game will end, they will still play as in a game with infinite time periods.

The best-known form of dynamic games are the repeated games, also called supergames or iterated games. In a repeated game, players face the same base game, called a stage game, in each time period. A game consisting of only a single execution of the stage game is called a one-shot game or single-stage game.

In a repeated game, setting a finite horizon offers little insight compared to the one-shot game, since it can be demonstrated via backwards induction that any rational player will always choose the solution to the single-stage game at every stage (Rubinstein, 1979). But in infinite-horizon repeated games, the optimal solution for each player is not necessarily a repetition of the optimal solution of the stage game (Garcia & Smith, 1996). This has been noted for the first time by Chamberlin (1929), who argued that the repeated interaction can lead to tacit collusion, since the threat of retaliation in later stages may be enough to enforce a cooperative strategy. A generalization of this idea is known as the folk theorem, which has been discussed in many versions (Friedman, 1971; Fudenberg & Maskin, 1986; Rubinstein, 1979). Regardless of the variations, the folk theorem states that in a repeated game, any combination of strategies satisfying the minimax-property can be a subgame perfect equilibrium.

Solution Concepts

In game theory, the most famous solution concept is the Nash equilibrium. Introduced by Nash (1950), it implements a very intuitive meaning of optimality. A strategy profile constitutes a Nash equilibrium if no player has an incentive of deviating from his strategy as long as the other players follow the Nash equilibrium.

For dynamic games, the notion of the Nash equilibrium needs to be refined, since it ignores the sequential nature of the game. A subgame perfect equilibrium is a strategy that is a Nash equilibrium for the remaining part of the game at every intermediate point in time of a game. An appealing feature of a subgame perfect equilibrium is the fact that it cannot depend on non-credible threats as opposed to a Nash equilibrium. In RM, a non-credible threat can occur when one firm threatens to react to a competitor lowering his prices by engaging in a ruinous price war in the next stages. As long as this price war would hurt the threatening firm's revenue more than some other reaction, this threat is not rational and can be dismissed. Thus, removing non-credible threats serves to strengthen the intuition of a Nash equilibrium as a rationally optimal set of strategies. If a subgame perfect equilibrium consists of only strategies possessing the Markov property and the state contains only payoff-relevant information, the equilibrium is called a Markov Perfect Equilibrium (Maskin & Tirole, 2001). Depending on the information structure of the differential game, subgame perfect equilibria are called closed-loop equilbria or open-loop equilibria.

Another refinement of the concept of a Nash equilibrium is the notion of an Evolutionary Stable Strategy (ESS). Maynard Smith and Price (1973) gave the first definition of ESSs, but we will present here a variation by Thomas (1985) that highlights better the connection of the ESS with the Nash equilibrium: A Nash equilibrium is an ESS if not only there is no incentive to deviate from the Nash equilibrium to another strategy, but also there is an incentive to deviate from any other strategy to the Nash equilibrium. ESSs are especially insightful in the study of repeated games. If a strategy of the single-stage game is shown to be an ESS, then it will be successful in the repeated game as well (Maynard Smith, 1982).

2.2.2 The Prisoner's Dilemma

The prisoner's dilemma is one of the most famous examples of problems studied by game theory. It has first been presented by Flood (1958) and Dresher (1981) and has advanced to one of the most extensively studied topics in game theory.

The prisoner's dilemma is particularly attractive to researchers, since it models a very common problem in non-zero sum games, that has been observed in many practical applications (Axelrod, 1984): While mutual cooperation is Pareto-efficient and leads to the highest combined payoff for all players, mutual non-cooperative behavior is the Nash equilibrium.

Studying a simple model such as the Iterated Prisoner's Dilemma (IPD) can help to establish a better understanding of the situation in more complex problems displaying a similar behavior (Grüne-Yanoff, 2009).

Formulation of the Game

The most popular formulation of this game is about two arrested men that are interrogated separately by the police (e.g. Tucker, 1980). During interrogation, each suspect is offered a deal by the police. If he testifies against his partner and his partner does not cooperate with the police, he will be given a reward of one unit and his partner will be fined two units. If both confess, they will each have to pay a fine of one unit. However, if they both cooperate with each other and do not testify against each other, they will both go clear. If both men act rational, i.e. only care about their personal consequences, the only Nash equilibrium of this game results in both men testifying against each other.

More generally, the prisoner's dilemma is a symmetric two-player game, where each player can either cooperate or defect. In the prisoner's dilemma, mutual cooperation maximizes the joint payoff, whereas mutual defection is the unique Nash equilibrium. While the analysis of this game is trivial as a one-shot game, allowing players to learn from past moves and adapt their strategies appropriately greatly increases difficulty, but also insight. The repetition of the prisoner's dilemma with an infinite horizon is commonly known as the IPD and is used as a standard model for direct reciprocity.

Successful Strategies

Interest in the IPD was raised by Axelrod's (1980) famous tournaments, later covered in more detail by Axelrod (1984). In these tournaments, Axelrod invited game theorists to contribute computer programs that would play IPD games against each other. Even though many participants had entered quite sophisticated strategies, the first and second tournaments were won by the simplest and most cooperative participating strategy, submitted by mathematical psychologist Anatol Rapoport. This strategy, called Tit for Tat (TFT), starts cooperatively and then always reproduces what the other player has played the round before. By design, TFT can never obtain a higher payoff than its opponent. Its cooperative stance however led to consistently strong second places when paired with a cooperating player. One interpretation of these results put forward by Axelrod and Hamilton is that cooperation is likely to evolve naturally due to being profitable to egoistic individuals. Experiments with humans have since confirmed the results of the computer tournament (Wedekind & Milinski, 1996), adding to the significance of this experiment.

The importance of mathematical analysis of repeated games can be highlighted at the example of the IPD. The success of TFT in Axelrod's simulation does not prove its general efficiency and robustness. Indeed it has been shown that no pure strategy such as TFT can be evolutionary stable in the IPD (R. Boyd & Lorberbaum, 1987). In fact, an even stronger result holds: No finite mixture of pure strategies and no mixture of Tit for n Tats can be evolutionarily stable (Farrell & Ware, 1989). This may seem to contradict the idea of TFT as a successful strategy in the IPD, but the authors rather "interpret [their] negative results to suggest that evolutionary stability is too demanding a criterion". If the possibility of mistakes is added to the setting of the iterated prisoner's dilemma, this is not true anymore. In this case, any strategy that is always the uniquely best response to itself is an ESS, as long as error probabilities are positive, but sufficiently low, and the amount of players playing a different strategy is kept sufficiently low (R. Boyd, 1989).

Following his famous tournaments of the IPD, Axelrod (1984) implemented a so-called ecological simulation. In such a simulation, players interact with each other repeatedly and may change their strategy randomly after each round, with a bias towards more successful strategies. The pool of strategies in this simulation corresponded to the strategies entered in his tournaments. When searching for a more general strategy space, researchers have described strategies by their transition probabilities given the game's elapsed history. The resulting game can be described as an infinite-order chain (see e.g. Iosifescu & Grigorescu, 1990), i.e. a process depending on all previous stages s. Compared to Markov chains of finite order, infinite-order chains are not as well researched and not as easily analyzed. However, a slight modification greatly helps simplifying the situation without limiting its generality. Press and Dyson (2012) showed that at least in the IPD, a longer memory does not lead to superior strategies compared to strategies relying solely on the most recent observation. These single-stage memory strategies enable the description of the game as a Markov chain, which helps in the investigation of the game's long-term behavior. For a general class of Markov chains, a unique limiting measure—the so-called stationary measure—exists, against which the process converges in the long run (see e.g. Pinsky & Karlin, 2010, pp.199 – 266). Molander (1985) has studied the stationary measure for a particular set of strategies for the IPD, while Hauert and Schuster (1997) has computed the stationary measure for randomly chosen strategies. Note that due to the simplicity of the IPD, researchers have computed the transition matrices of the game without giving a thorough derivation (Boerlijst, Nowak, & Sigmund, 1997b; Molander, 1985; Nowak & Sigmund, 1990). The modeling of the IPD as a Markov chain was used by a group of researchers, who refined Axelrod's idea of an ecological simulation to so-called evolutionary simulations (Boerlijst et al., 1997b; Hauert & Schuster, 1997; Nowak &

Sigmund, 1992, 1993). In contrast to Axelrod's ecological simulation, these authors saved time and computational effort by relying on the stationary measure of each matchup, instead of simulating the interactions move by move. In these simulations, players were able to explore the whole strategy space, which was restricted to Markov strategies. In order to save time and computational effort, these authors relied on the stationary measure of each matchup, instead of simulating the interactions move by move. However, by allowing for additional players or a longer memory, the size of the strategy space can inhibit the exploration of the strategy space (Hauert & Schuster, 1997).

Since Axelrod's tournaments, the evolutionary approach has helped to discover other strategies that have been shown to perform well in the IPD. These have been mainly modifications of TFT, that addressed the weaknesses of this strategy. More precisely, TFT is vulnerable to errors (Wu & Axelrod, 1995). For example in the case of two TFT players competing against each other, a single accidental non-cooperative action of one of the players can lead to both players never cooperating again. Molander (1985) has shown that with noisy observations, two players using TFT will on average receive a similar payoff as two players using a random strategy. The first suggestion to solve this problem was made by Axelrod (1984) with the Tit for 2 Tats (TF2T) strategy, which retaliates only after two rounds of the other player's non-cooperation. This modification leads to more robustness against noise, while still performing quite well even without any noise, as was the case in the computer tournament. When analyzing the results of his first tournament, he found that TF2T would have even outperformed TFT. However, when entered in the second competition, TF2T was not as successful as TFT, due to the change of competitor strategies that were able to exploit TF2T's generosity (Axelrod, 1984).

Another variation of TFT that is more robust to errors is called Generous Tit for Tat (GTFT) (Nowak & Sigmund, 1992). Following this mixed strategy, a player cooperates with some fixed probability even though the other player has defected the round before. This kind of forgiveness can help avoid long vendettas in the presence of errors or imperfectly observable actions. The best probability of forgiveness has been computed by Molander (1985). Nowak and Sigmund (1992) have shown that GTFT dominates strategies that take the competitor's last action into consideration, although it needs TFT as a catalyst against aggressors.

TFT as well as GTFT base their decision solely on the opponent's last action. This can be changed by introducing a reputation for each player (Sugden, 1986). Each player starts with a good reputation. If a player defects without having been given a reason by his opponent, i.e. against an opponent with a good reputation, the player's reputation will drop to bad. If however a player defects against a competitor with a bad reputation, this does not change the player's reputation. For any player, cooperation will always restore a good reputation in the next stage.

A simple strategy using the reputation of each player is called Contrite Tit for Tat (CTFT) (R. Boyd, 1989). A CTFT player with a bad reputation will always cooperate to reestablish his good reputation, which explains the name of the strategy. Also, a CTFT player will always cooperate with an opponent with a good reputation. However, against a player with a bad reputation, a CTFT player with a good reputation will defect until the opponent cooperates, which will restore the opponent's good reputation. This procedure constitutes a possibility to deal with Tit for Tat's susceptibility to errors and in fact can constitute an ESS in the presence of errors (R. Boyd, 1989). In contrast to Generous Tit for Tat (see Nowak & Sigmund, 1992), Contrite Tit for Tat is good at invading population of defectors (Boerlijst et al., 1997b). Unfortunately, while the introduction of a reputation renders CTFT immune to errors in implementation of actions, it also introduces the possibility of errors during perception of an opponent's actions (Boerlijst et al., 1997b).

If players are allowed to not only react to their opponent's actions, but also to their own, another simple strategy emerges that has been shown to be even more successful and robust in an evolutionary sense for some settings (Nowak & Sigmund, 1993). The PAVLOV strategy in the prisoner's dilemma consists of cooperating if and only if both players played the same action in the run before. This is an example of a simple kind of strategy called win-stay, lose-shift, because it keeps the previous strategy if it has proved successful and switches strategies if the previous payoff was low. If errors are introduced, the strategy can exploit blind cooperation, while being itself robust against errors due to its error-correction in the symmetric setup (Kraines & Kraines, 2000). However, due to its interaction with aggressors, PAVLOV is only successful as long as the benefit of cooperation is high enough. PAVLOV profits from the presence of TFT-like strategies, since Tit for Tat is more effective against aggressors. But ultimately the ability to prevent blind cooperators from spreading and thus attracting aggressors leads to better performances for the PAVLOV strategy than TFT (Imhof et al., 2007). However, until TFT has paved the way, PAVLOV may perform poorly. Wu and Axelrod (1995) report that PAVLOV finished one of their tournaments as one of the worst strategies.

A combination of PAVLOV and CTFT has been dubbed Prudent Pavlov by Boerlijst et al. (1997b). This strategy is robust both against errors in the implementation of players' actions, but also to errors in the perception of these moves. Pelc (2010) follows another approach in creating a robust strategy against observation errors. The author demonstrates the existence of perfectly fault-tolerant strategies. However, these strategies require an unbounded memory and thus do not possess the Markov property.

Boerlijst, Nowak, and Sigmund (1997a) took a different point of view on the IPD. They found that there exist Equalizing Strategies, which allow a player to unilaterally set the score of his opponents, independently of the opponents' strategy. Later, Press and Dyson (2012) generalized this concept. They proved that the IPD allows the existence of a broader set of strategies, which allow a player to unilaterally set an affine relationship between both players' payoffs. The feasibility of any particular affine relationship depends on the choice of parameters of the affine relationship as well as on the payoffs of the game. These strategies are called Zero-Determinant Strategies. A particularly interesting special case are the Extortionate Strategies. Using such a strategy, a player can fix a linear relationship between both players' profits, where the profit of a player is his payoff minus the payoff for mutual defection. Press and Dyson showed that every possible Extortionate Strategies is feasible in the IPD, so that a player can choose to receive any multiple of the opponent's profit. For the special case, in which the player fairly sets his share of profits equal to the opponent's, they found the resulting strategy to be TFT. The power of Extortionate Strategies lies in the relationship with evolutionary players. Since a player, who is following an evolutionary approach, will aim to maximize his profit, he will inevitably maximize the extorting player's profit as well, albeit at a higher level. The only way out of this extortion is playing an ultimatum game. However, this requires a theory of mind and is not possible via evolution.

Given the amount of insight extracted from the IPD, it comes as no surprise that there have been many attempts at broaden its scope in order to gain a better understanding of similar situations. The IPD has been extended to the case of more than two players (Berkemer, 2006; Yao, 1996). Other modifications include the Alternating Prisoner's Dilemma (APD), in which players choose their moves in a non-simultaneous way, which has led to new dominating strategies (Frean, 1994). The extension of the amount of moves, that players may base their decision on, yields more successful strategies in the APD (Neill, 2001), but not in the IPD (Hauert & Schuster, 1997).

2.2.3 Summary and Implications

Game theory is applicable to any situation in which multiple decision makers interact. Thus, the concepts of game theory presented in this section lend themselves well for the problem of RM under competition, as we will show in Section 2.3.2. However, most insight has been gained from studying simple games such as the IPD, where the effects of repeated interactions could be studied in a thorough way. In this thesis, we will transfer ideas from the game theoretic treatment of the IPD to the RM scenario with repeated interactions between service providers.

2.3 Revenue Management under Competition

In the last decades, the ongoing work of researchers and practitioners has seen RM forecasting and optimization techniques come a long way. However, up to this day, most RM models do not take into account the effects of competition (Martínez-de Albéniz & Talluri, 2011). The effectiveness of these methods has been proven in monopoly settings only, but it is often assumed that they perform well in a competitive setting (Talluri & van Ryzin, 2004b, p.186). This is usually credited to the fact that estimating and forecasting in a competitive environment leads to data that implicitly accounts for competition. A case in point would be the above mentioned success story of the first use of RM by American Airlines (1987), where a very simple model performed well enough in a competitive setting to drive the competition out of business. However, Cooper, de Mello, and Kleywegt (2009) showed that while this implicit incorporation of competition in RM might in some cases be sufficient, in other cases it can lead to considerably suboptimal results.

Therefore, recent years have seen increased efforts to establish models that explicitly consider competitive effects in revenue management.

2.3.1 Amendments to the Monopoly Model

With the rise of the internet, prices of each competitor are much more transparent. Now customers have the possibility to compare prices for similar products across different competitors. This changes the nature of the competition, as S. P. Anderson and Schneider (2007) have shown that a costly search for prices can lead to even higher prices in a duopoly than in a monopoly. Additionally, each service provider now has the possibility to monitor prices from all their competitors quite easily. This information can be used as additional information about the competitor strategy compared to traditional RM forecasts. There are specialized firms who offer competitor price information on a varied range of level of detail (e.g. Infare Solutions, 2013), which can then be used in a RM system.

Lua (2007) analyzed the strategy of matching the lowest available fare of an airline on top of a standard RM system. In simulations with Passenger Origin-Destination Simulator (PODS), he found that price matching has a negative effect on the matching airline and a positive effect on the matched airline. Thus, this overly simplistic solution seems far from optimal.

A more systematic approach to price matching—albeit in a much simpler scenario—was used by Marcus and Anderson (2008). Under the assumption that the demand process and the competitor price process are known and follow simple ordinary differential equations, Marcus and Anderson analyzed the optimal strategy of a carrier with a service disadvantage. In their work, that carrier has to price lower than the competitor in order to attract the same level of demand due to customers perceiving its service as worse than the competitor's. The resulting optimal control takes the form of a simple bang-bang control.

Currie, Cheng, and Smith (2008) consider a generalization of this problem. Instead of the overly simplistic assumption that competitor prices are governed by ordinary differential equations, they allow for any function depending only on time. They formulate a capacity constraint that is meant to prevent overbooking, although in the formulation they use it, the constraint only states that there is no overbooking on average. This capacity constraint is then used to reduce the optimization problem to solving a set of differential equations with the help of calculus of variations and Lagrangian multipliers. Given an unresponsive competitor with a known price function, they find the best strategy for the rational carrier. However, they need to know competitor prices beforehand, which is a strong restriction of the model. Since competitor prices are assumed to be dependent on time only, this could be replaced by forecast of competitor prices.

Zhang and Kallesen (2008) present such a modification to a standard RM optimization technique, using of the availability of competitor prices. They observe competitor prices in order to fit parameters of a Markov chain, which helps them produce a forecast of the future competitor price. The resulting stochastic process depends only on time and observed competitor prices and thus disregards competitive interactions between suppliers. In their optimization problem, the authors use a dynamic program to find the best solution given the transition probabilities of the competitor's prices. Simulation results show that their method performs superior to standard approaches not considering competition, as long as the competing carrier remains unresponsive.

While Marcus and Anderson and Currie et al. model competitor prices as time-dependent exogenous functions, Ledvina (2011) uses a different approach, in which competitor prices are a result of some form of optimization of the competitor. If the competitor prices result from a control that would be optimal in a monopoly setting, Ledvina (2011) is able to calculate the best response. Ledvina requires the true demand to be known for both firms. However, the true demand for each provider in a duopoly depends on the strategy of the competitor as well. Thus, there is no true demand for a single firm, independent of the competitor's actions. But this is the input needed for the optimization of the irrational firm and its value. The choices made regarding the monopolistic demand may very well have a significant input on the carrier's strategy. In a realistic environment, these choices will depend on the forecast employed by the irrational provider, which calls for an analysis of different monopolistic strategies.

Both the works of Ledvina and Marcus and Anderson assume that demand in the market is known. Already in a monopoly setting, forecasting is one of the most difficult parts of any RM system, since it is impossible to observe potential customers that refrain from booking. Since in a competitive setting, the portion of observable customers further declines for each carrier, the task of forecasting is even more difficult.

A simple possibility to include competition in standard monopoly RM by using competitor information to enhance the forecast has been proposed by d'Huart and Belobaba (2012). The core idea is to form a belief about the competitor's booking situation by comparing its current lowest fare to historic observations of his lowest fare. This belief is used to modify the forecast in order to accommodate variations from the historic observations due to a change in competitive behavior. Simulation results show low but statistically significant gains over monopoly forecasts, as long as only one firm in the market uses this approach. However this positive effect vanishes, as soon as multiple airlines adopt the method.

2.3.2 Game Theory in Revenue Management

Game Theory has proven to be an effective tool for problems closely related to RM such as the newsvendor problem, which is essentially the single-resource RM problem with two fare classes over a single time period (Huang, Zhou, & Zhao, 2010; Jiang, Netessine, & Savin, 2011; Lippman & McCardle, 1997; Zhao & Atkins, 2008). In recent years, researchers have started to apply game theoretic ideas directly to RM problems. Most researchers try to find the Nash equilibrium of the single-stage game describing a single booking horizon. Sometimes the exact equilibrium is replaced by a heuristic that may be easier to compute.

In one of the first papers to explicitly incorporate competition on Dynamic Pricing, Dockner and Jørgensen (1988) study optimal dynamic pricing policies in an oligopoly without capacity constraints. They formulate a differential game and solve for an open-loop Nash equilibrium.

Netessine and Shumsky (2005) study the effects of competition in airline RM on the single-leg time-static model with two fare classes. In their model, demand may switch carriers, if one carrier does not offer the desired class, but will not switch classes. They find the existence of a pure strategy Nash equilibrium, in which each carrier protects more seats for the higher class than in the optimal solution for the monopoly.

Adding the effect of pricing, Zhao (2003) studies a similar problem, albeit with a different demand spill-over model. Here carriers can change prices for the high fare class as well as choose a booking limit. Zhao finds that the best strategy depends on the type of competition in the market. In his work, the effect of customers choosing the competitors' lowest price is called "price competition", and the effect of spill from one carrier to another due to stock-out of one carrier is denoted "seat inventory competition". The author points out that in a market dominated by price competition, carriers should price lower and protect less seats for the high fare class than in a monopoly, which is in line with classical economic literature. Similarly to Netessine and Shumsky's findings, airlines should price higher and protect more seats when seat inventory competition predominates. For a mix of both types, carriers should price the highest class lower than in a monopoly, but protect more seats for it.

Martínez-de Albéniz and Talluri (2011) analyze a discrete-time finite-horizon model of price competition under capacity asymmetries, where each firm offers a single perfectly substitutable product. In this model, customers always purchase the cheapest product as long as it is cheaper than their willingness-to-pay, which is common knowledge for each firm. Under the assumption that each firm can observe the inventory level of its competitor, the authors prove the existence of a unique subgame-perfect equilibrium in the duopoly setting.

The above papers assume that service providers know the booking limits of their competitors. However in reality, it is not possible to observe the inventory of a competing carrier, so that competitor's booking limits as well as remaining capacities cannot be assumed to be known (Talluri, 2003). Since competitor prices are easily observable since the rise of the internet, many researchers have focused on using prices as the available information in the modeling of their games.

Lin and Sibdari (2009) use another approach at the problem of unobservable inventory levels. Their model is a variation on the single-resource case, where each firms sells substitutable products using a single resource. They use a dynamic model in discrete time, in which customers arrivals are modeled as a Bernoulli process. Each customer's choice is modeled by a discrete choice model. In game-theoretic terms they analyze a discrete-time, finite-horizon game. The authors first calculate Nash equilibria of the game under the assumption of observable demand and inventory levels. In the next step, they present an interpolation heuristic for the case that inventory levels are only known in the beginning of the booking period.

The following articles in this section use competitor prices as the observable quantity.

Mookherjee and Friesz (2008) analyze a discrete-time, finite-horizon game with a known demand model in combination with other typical RM problems, such as overbooking and networks. The authors prove the existence and uniqueness of a pure strategy Nash equilibrium. They find that in the equilibrium, service providers price lower than in the monopolistic setting. In numerical experiments, they compute the revenue gap between cooperative and non-cooperative optimal control.

Perakis and Sood (2006) consider a discrete-time, finite-horizon stochastic game with continuous demand. In their variation of the problem, demand information that is extracted of observed prices still contains some level of uncertainty. They use ideas from robust optimization to find a unique open-loop solution and present a simple iterating learning scheme converging towards the equilibrium.

For a similar setting, another attempt at relaxing the assumption of knowing the demand model exactly in advance is made by Bertsimas and Perakis (2006). They consider a finite-horizon, discrete-time game, where firms compete on price for a single perfectly substitutable product. In their work, only the functional form of the demand model is known in advance, whereas its parameters have to be estimated by each firm during the booking period.

An extension to this approach is made by Simon (2007), who analyzes inequilibrium states of the game. Furthermore, Simon (2007) builds on the work of Gallego and van Ryzin (1994) on Dynamic Pricing, where the author formulates a game in continuous time and computes a closed-loop solution for a small example. Since this turns out to be infeasible for larger problems and general demand models, Simon introduces several heuristics improving on the open-loop solution consisting of either a simple feedback control or approximations of the closed-loop revenue function in the Dynamic Program. In the same setting, Ledvina (2011) calculates closed-loop equilibria for linear and exponential demand models.

Kwon et al. (2009) use an evolutionary game theory approach to model a continuous-time, finite-horizon game, for which parameters of the demand model have to be estimated by each carrier. Their demand model is based on deviations from an averaged reference price. The dynamics of this non-cooperative market model are formulated as a differential variational inequality, which can be transformed into an equivalent optimal control problem given the assumption that competitive prices are known beforehand. With functional form of the demand model known to each player, they use a Kalman filtering approach to estimate demand and solve the optimal control problem to obtain a open-loop solution.

Isler and Imhof (2008) suggest a way to introduce non-cooperative Nash equilibria into a RM simulation model by idealizing the behavior of a typical RM forecasting system. The authors construct a "psychic" forecast considering competitor availability information: After every booking request, each carrier learns which classes it could have sold given the customer's willingness-to-pay and the competitor's actual offers. This information is used to forecast future bookings. The simplicity of this approach guarantees usability in much more complex settings, where the actual Nash equilibria might be hard to compute. Isler and Imhof test their method in a discrete-time game with an exponential demand model, where the calculation of the exact Nash equilibria is tractable. They compare their technique to the actual Nash equilibrium and find it to perform sufficiently similar. In their setting, the Nash equilibria lead to a Competitive Spiral Down: Revenue decreases with increasing capacity as soon as a threshold capacity value has been reached, eventually converging to the Bertrand equilibrium in the case of no capacity constraints. Considering the discrepancy between their result and real-life observations, the authors reason that RM should be considered as a repeated game.

As a follow-up, Isler and Imhof (2010) present a solution to the problem of the Competitive Spiral Down, which they identify as a pricing problem. If both players agree not to choose prices below some threshold level, the price war can be limited. If this threshold value is chosen as the lowest price that a carrier should offer in the monopoly situation without capacity restrictions, the strategies end up close to the cooperative Nash equilibrium. In a repeated game, this tacit collusion corresponds to cooperative behavior.

Despite the remark of Isler and Imhof, there is not much research on the dynamics of RM under competition modeled as a repeated game. In a paper extending the seminal work on dynamic pricing by Gallego and van Ryzin (1994), Gallego and Hu (2007) introduce competition into the RM setting, thus creating a finite-horizon continuous-time stochastic game. Given that the demand model is known, demand intensities can be calculated from observed competitor prices. The authors find an open-loop solution for the single-stage game, in which the optimal prices have a similarly simple structure as in the monopoly. Furthermore, they compute the best strategies against irrational competitors that do not follow the optimal strategy. They also analyze the problem in an iterated game setting, where they prove a version of the folk theorem. However, they admit that their "result seems to be vague in that which equilibrium eventually is settled upon is determined by each firm's strategic behavior that is hard to model."

2.3.3 Summary and Implications

In the literature, there have been two different approaches to the problem of RM under competition: While some researchers have put forward amendments to the monopoly case of RM described in Section 2.1, others have presented applications of the game theoretic concepts presented in Section 2.2. However, for both approaches the literature concentrates on the single-stage case, although a multi-stage view is more appropriate for the problem. In fact, employing a single-stage view can prove very costly, since it can lead to the Competitive Spiral Down. On the other hand, the only approach to model RM competition with repeated interactions stresses the difficulties in modeling the players' strategic behavior and determining the game's stationary state. In this thesis, we will present a possibility to introduce repeated-game strategies into RM, so that we can investigate the limiting state of the game.

2.4 Simulation

Before any new strategy is put into practice, it first has to be evaluated to ensure its performance and robustness in a broad range of scenarios. If the interaction of this strategy with its environment is well understood and easily described analytically, an analytical evaluation offers the most complete insight. However, if a mechanism is not well understood, its impact has to be determined by experiments. The experiment can be executed in real life (e.g. Talluri, Castejon, Codina, & Magaz, 2010) or in a simulation environment (e.g. Belobaba & Hopperstad, 2004; Fiig et al., 2009).

However, experiments in real life have several disadvantages compared to simulations. In real life, an experiment may soon turn out very costly if the scale of the experiment is sufficiently large. If the experiment is kept small, the range of scenarios, against which the method is tested, is kept smaller as well. Also despite the idea of working with a control group in real life by Talluri et al., an experiment ceteris paribus is not possible in real life. In a simulation however, it is possible to test strategies ceteris paribus against a wide variation of scenarios without risking great costs.

Law and Kelton (2000) give a thorough guide to simulation, which helps the reader gain an understanding of challenges and solution concepts of simulations.

2.4.1 Simulations in Revenue Management

In RM there are many opportunities that fulfill the above points making a simulation necessary in order to gain a better understanding of the system. A typical use case of simulation in RM is the evaluation of forecast performance, since in the real world, the real demand can never be observed (Cleophas, Frank, & Kliewer, 2009b). This necessity has sparked interest in RM simulators for a long time.

About 20 years ago, researchers at Boeing developed probabilistic simulators as a means of simulating air travel passengers (*Decision Window Path Preference Model (DWM)*, 1994). Since this early simulator was aimed at providing decision support for fleet planning and scheduling, it featured a relatively sophisticated customer choice model, but no RM methods. Realizing that this model could not explain observed booking realizations, Boeing used this system as a basis to develop the PODS in collaboration with the Massachusetts Institute of Technology (MIT) (Hopperstad, 1995). Since then, PODS has served as decision support for research questions both from researchers and practitioners (e.g. Fiig et al., 2009; Gorin & Belobaba, 2004).

While PODS features a rich set of RM methods and is capable of simulating very large networks, criticism of its shortcomings has grown in recent years. In particular, the inaccessibility of the code to other researchers, an oversimplified customer model, inflexible booking class restrictions and simplistic schedules have inspired the creation of similar simulators. For example, the RM simulator REMIGIUS allows for dynamic capacity control in the optimization (Frank, Friedemann, & Schröder, 2007a). Following another approach, the Travel Market Simulator (*Travel Market Simulator*, 2013a) is built using a modular approach and released as open source. The creators of this simulator hope to attract researchers willing to contribute modules, so that the software can grow “to become the new generation PODS” (*Travel Market Simulator*, 2013b). Building on experience with PODS and REMIGIUS, Frank, Friedemann, and Schröder (2007b) describe a set of guidelines for simulations in revenue management. REvenue Management Training for Experts (REMATE) is a RM simulator developed by Lufthansa in cooperation with several universities and a software provider, that is built following the principles presented by Frank et al. (2007b). Earlier versions of REMATE have been described by Zimmermann, Cleophas, and Frank (2011) and Gerlach, Cleophas, and Kliewer (2013), while the state of implementation used in this dissertation will be described in Section 4.3. Compared to PODS, REMATE offers a more flexible customer model, schedule and booking class system. On the other hand, simulations in REMATE have to be run on considerably smaller networks.

2.4.2 Simulations in Game Theory

The amount of insight gained from Axelrod's simple computer tournament (1984) proves the worth of simulations to analyze repeated games. A simulation can often offer insights into the behavior of models that are too complex to examine in a mathematical model. In fact, research in simulations may stimulate further theoretical research in this area which aims at validating these results in a more rigorous fashion. However, due to the limiting assumptions of these mathematical models, agent-based simulations can in some cases even offer a more rigorous approach than mathematical modeling (Chattoe, 1996).

Gotts, Polhill, and Law (2003) give a survey over the advances in agent-based simulations for repeated games. Nevertheless, the authors stress the importance of theoretical approaches. The authors argue that the analysis of a simple simulation model should always be accompanied by examination of related theoretical research, while the use of more complex simulation models should include references to simpler models. This way theoretical analysis and agent-based simulations can profit from each other: Insights gained from simplified models can help explain phenomena in a more complex setting, while analysis of complex models helps validate ideas gained in a simplified setting as well as study interactions that have been ignored for the sake of simplicity.

2.4.3 Summary and Implications

Simulations have proven successful in the analysis of RM, where PODS has been used extensively in research and practice. In game theory, simulations such as Axelrod's (1984) computer tournament have helped raise interest and gain insights in the IPD. Nevertheless, an investigation led by simulations should always be complemented by theoretical analysis, if possible. In this thesis, we will use a combination of theoretical analysis and simulations to examine the competitive interactions of service providers. We will first use simulations with REMATE in Chapter 4. Then, we will turn to a mathematical analysis in Chapter 5, accompanied by simulations based on the analytical results.

3 Research Gap: Repeated RM Competition

Revenue Management (RM) has been credited with helping firms succeed in a competitive environment since the early days of balancing full fares against discounted fares. As one of the first adapters of revenue management, American Airlines managed to drive its low-cost competitor out of business with the help of its RM techniques (*The Art of Managing Yield*, 1987). Since then Belobaba and Wilson (1997) have shown that using a RM system is beneficial in a duopoly as well as in a monopoly. Thus, it appears clear that the benefits of using RM are not restricted to the monopoly case. However, the question for the best RM strategy in a competitive environment is still unanswered.

As outlined in the literature review in Section 2.3, the RM problem in a competitive environment has been approached from two disciplines: RM, reviewed in Section 2.1, and game theory, discussed in Section 2.2. From a RM perspective, researchers have examined the problem in a realistic setting and applied or adapted RM techniques that were originally developed for the monopoly case (see Section 2.3.1). Because of the high complexity of the setting, the evaluation of these approaches is usually restricted to simulations (see Section 2.4), since a mathematical analysis is often unfeasible. Several researchers have employed the Passenger Origin-Destination Simulator (PODS) to analyze competition between service providers using standard monopoly RM systems. Subsequently there have been different attempts to extend monopoly systems to account for competitive effects as well. Similarly to the simulation results of the monopoly RM systems, the performance of these expansions has been evaluated by comparison with other standard RM systems only. Therefore, the effects of interaction of competitive RM systems with each other have not been analyzed.

On the other hand, in Section 2.3.2, we have discussed a tremendous amount of research from a game-theoretic perspective, which explicitly accounts for interaction between the players. These authors have modeled games describing the RM problem of a single booking horizon, each describing a slightly different simplification of the original problem. In these papers, the evaluation process consisted of showing that the resulting strategies formed a Nash equilibrium. The resulting strategy profile is never put in comparison with standard RM techniques. Thus, the question of similarities and discrepancies of standard RM systems and game-theoretic solutions remains unanswered.

In Section 2.3, we found that the vast majority of research on RM under competition—relying on RM or on game theory—aim at finding an optimal solution for the problem during a single sales period. However, Isler and Imhof (2008) showed that the single-stage Nash equilibrium leads to the Competitive Spiral Down effect that can lead to the ruin of both players. Thus, each player faces a dilemma: He can either follow the non-cooperative solution, which will lead to the Competitive Spiral Down. Or he can implement cooperative behavior that approximates the joint optimum, but is vulnerable to non-cooperative aggressors. However, as Isler and Imhof (2008) pointed out, it is much more suitable to think of competition in RM as a repetition of similar events. This lends itself perfectly to modeling as a repeated game. Unfortunately, there is hardly any published research in that direction.

The dilemma faced by players engaged in the RM game is not unique. In fact, the most famous and best example of such a dilemma is formulated in the simple Prisoner's Dilemma, which may serve as an inspiration for the analysis of the RM game. In its repeated form as the Iterated Prisoner's Dilemma, this game has seen a tremendous amount of research that has produced several successful strategies (see the literature review on the Iterated Prisoner's Dilemma (IPD) in Section 2.2.2). Although Isler and Imhof (2008) and Gallego and Hu (2007) have acknowledged the similarity of the RM game to the IPD, this has not resulted in a new approach to the problem of RM under competition anywhere in the literature.

Research Question 1. *How can strategies of the IPD be adapted to the repeated RM game?*

Gallego and Hu (2007) provide us with an adapted version of the famous folk theorem for the repeated RM game, stating that, by choosing an appropriate discount factor, any payoff greater than the payoff for the single-stage Nash equilibrium can be obtained by a Nash equilibrium of the repeated game. Therefore, at least in the slightly simplified setting of Gallego and Hu, we are guaranteed that there exist strategies in the repeated RM game that solve the dilemma of the RM game by forming a jointly optimal Nash equilibrium. Unfortunately this does not help in the search for the best strategy.

Research Question 2. *Which strategy leads to a jointly optimal Nash equilibrium in the repeated RM game?*

To the best of our knowledge, the connection between monopoly-based RM systems, as they are currently wide-spread in practice and research, and game-theoretic solution approaches has not been made in the literature (see Section 2.3). The work of Ledvina (2011) represents a step in this direction as discussed in Section 2.3.1. Ledvina has computed the optimal strategy for a carrier under the assumption that its competitor uses a standard monopoly RM system. However, as pointed out in the literature review, this procedure depends on the monopoly-based RM system in question. Isler and Imhof (2008) showed in simulations that an idealized standard RM system leads to Nash equilibrium. Building on this method, Isler and Imhof (2010) presented a heuristics to approximate the jointly optimal solution as well. On the other hand, Cooper et al. (2009) demonstrated that a RM system trying to estimate demand as independent of the competitor does not necessarily converge to the Nash equilibrium and instead lead to different steady states such as the cooperative solution. Thus, the question, whether and how well a realistic standard RM system approximates these game-theoretic solutions, has not been answered yet.

Research Question 3. *How closely do standard RM methods approximate the game-theoretic solutions?*

The ongoing struggle to find optimal feasible strategies in competitive RM has led to service providers using simple irrational, yet competitive techniques such as price-matching and underpricing. In spite of the significant share of carriers using these methods, the literature offers no guidance on how to react to these strategies.

Research Question 4. *Which strategies are best suited to react to simple irrational strategies like price-matching or underpricing?*

In a realistic RM environment, it is impossible to completely avoid observation errors. Thus, while evaluating the success of strategies in the repeated RM game, it is essential to account for the possibility of flawed observations of the competitor's actions. The importance of errors is stressed by the fact that researchers have shown that the introduction of observation errors can severely change the dynamics of the game in the IPD (see literature review in Section 2.2.2).

Research Question 5. *What is the effect of observation errors on strategies in the repeated RM game?*

As discussed in Section 2.2.2, there is no singular best strategy in the IPD, since the results depend on the possibility of observation errors as well as on the frequency of competing strategies in the environment. Therefore, a rational player will choose the optimal strategy for his environment by learning from his repeated interactions with other competitors. In the literature review in Section 2.2.2, we have pointed out that for the IPD, such a learning scheme can be formulated as an evolutionary game, where each player reassesses his strategy after each round with the possibility to change towards more successful strategies. Due to the similarity of the IPD to the repeated RM game, a similar procedure may lead to insights about the success of RM strategies in different competitive environments.

Research Question 6. *Which repeated-game strategy is competitively robust in the sense that it fares best against a diverse set of competing strategies?*

However, we stressed Section 2.2.2 that such a learning strategy can be dangerous in the IPD. As Press and Dyson (2012) have shown, the IPD allows for Extortionate Strategies, which allow a player to unilaterally enforce to obtain an unfair share of the total payoffs. Against such a strategy, the best response is to cooperate, even though the extorting player does not always cooperate. Consequently, implementing such an Extortionate Strategy will lead a rational competitor to the cooperative solution, where the extorting player's payoff is maximized as well.

Research Question 7. *Do Extortionate Strategies exist in the repeated RM game?*

As discussed in the literature review, the RM problem under competition can be tackled either via simulations or analytically. The computational approach via simulations enables the researcher to embed the problem in a realistic and complex setting. This way, the researcher does not have to make possibly oversimplifying assumptions in his analysis. On the other hand, a mathematical analysis might not always be feasible, but it can provide deeper insights into the structure and the mechanics of the game. In this thesis, we will endorse both a computational as well as an analytical approach in order to examine the RM game as thoroughly as possible from as many angles as possible.

First, in Chapter 4, we will describe competition in a quite general RM context. We will use simulations to evaluate different strategies in a realistic setting, concentrating on questions 1 – 5.

In Chapter 5, we will examine the problem from an analytical point of view. In order to enable a mathematical analysis of the RM problem under competition, we will simplify the RM game from Chapter 4, mainly by focusing on the situation, where competitive effects are not dampened by capacity constraints. The resulting game will be both a simplification of the RM game and a generalization of the IPD. In this chapter, we will focus on questions 2 and 5 – 7.

4 RM Competition as a Repeated Game

In this chapter, we will present a model of RM-using service providers under competition. We will develop a heuristic to adapt strategies from the Iterated Prisoner's Dilemma (IPD) to the repeated RM game and evaluate different strategies with the help of simulations.

As indicated in the literature review in Chapter 2, researchers have approached the problem of competition between service providers from two different directions. One stream of research has discussed the problem in a realistic setting, using simulations to evaluate the performance of standard Revenue Management (RM) systems under competition (see Section 2.3.1). Another has examined simplified versions of the problem from a game theory perspective, usually solving the problem for the Nash equilibrium (see Section 2.3.2).

With the research divided in such a way, researchers as well as interested practitioners have had to balance advantages and disadvantages of the two research directions. While the treatment of realistic scenarios is intriguing, the dependence on an unmanageable amount of parameters can make a thorough analysis infeasible. In particular, the restricted set of tested strategies implies that these studies cannot answer the question for the best possible strategy. On the other hand, the game theoretic treatment usually provides an exact optimal solution through mathematical analysis. But in order to be able to calculate the Nash equilibrium, researchers have to make simplifying assumptions that can hinder the transfer of these results to realistic scenarios. This is a stark contrast to analyses using a simulation environment such as the Passenger Origin-Destination Simulator (PODS) or REvenue Management Training for Experts (REMATE), where scenario details can be tuned to create an almost arbitrarily complex scenario.

To bridge the gap between simulation-based and game theoretic analyses, we can adopt ideas from Isler and Imhof (2008), who showed that an idealized version of the forecast in a RM system can approximate the Nash equilibrium. Later, Isler and Imhof (2010) built on this approach and presented an approximation of the jointly optimal solution. As outlined in the literature review in Section 2.3.1, they find that the Nash equilibrium leads to the so-called Competitive Spiral Down: As long as the demand to capacity ratio is sufficiently low, service providers find themselves in a race to the bottom converging to the Bertrand Nash equilibrium. We will use their ideas to compute approximations of the single-stage non-cooperative and cooperative solutions.

Isler and Imhof (2008) remark that the RM game should not be treated as a single-stage game consisting of a single sales period, but instead as a repeated game. They argue that a repetition of the cooperative solution would solve the Competitive Spiral Down faced by the non-cooperative Nash equilibrium strategy of the one-shot game (Isler & Imhof, 2010). However, the similarity of this game to the IPD hints at different solutions. We will present a heuristic that will allow us to adapt successful strategies from the IPD to the repeated RM game. Since mistakes can happen in RM, it is important to implement a strategy that is robust against errors. In this chapter, we will analyze various strategies for the multi-stage RM game, both with and without the possibility of observation errors.

We will also reexamine Isler and Imhof's approximation of the single-stage game theoretic solutions, relying on an idealized forecast, which is only feasible in a simulation. Additionally, we will test whether a realistic forecast in a standard RM system will lead to similar results. For this purpose, we will use several standard RM methods of varying degrees of complexity to represent realistic RM systems. These methods were designed for the monopoly setting and consequently do not incorporate competitive effects in an explicit way. Because of the interaction of the two service providers, there is however an implicit consideration of the competitor's actions. The widespread use of these techniques is based on the belief that over time the implicit learning of competitive effects will suffice to reach an optimal strategy. However, Cooper et al. (2009) have shown that while this approach may converge to the non-cooperative Nash equilibrium, it may also converge to the cooperative solution or to an altogether different strategy. In this chapter, we will compare strategies using a standard forecast to strategies based on an idealized forecast similar to Isler and Imhof's. We will analyze the severity of the Competitive Spiral Down effect as well as the effect of using repeated game strategies on top of standard RM systems. Furthermore, we will use the approximations to the non-cooperative and to the cooperative solution of the single-stage as benchmarks, against which the standard RM techniques are measured.

Continuing our analysis of realistic and wide-spread competitive behavior, we will also investigate the effects of firms using irrational strategies like price matching. There are essentially two different possibilities to perform price matching. A service provider can either copy all the prices filed by its competitor, which is a pricing mechanism. Alternatively he can try to copy its competitor's availability situation, which is a RM mechanism and therefore more suited for our purposes. Price matching strategies are a simple way to explicitly integrate competition into a RM system. Consequently, these strategies have been used extensively in research and practice, as discussed in Section 2.3.1. Price-matching firms usually only use the information about the competitor's lowest available class, which can be gathered easily since the rise of the internet. Assuming a nested fare structure, this is enough to reach a sufficient precision about the competitor's availabilities.

In the following Section 4.1, we will give a mathematical model of RM in a single sales period under competition. We will build on this model to formulate strategies for the repeated interaction of the competitors in Section 4.2. Then, in Section 4.3, we will present the simulation environment REMATE, which we will rely on to yield computational results. In Section 4.4, we will design and evaluate simulation experiments that help us investigate the nature of RM with repeated competitive interactions. Finally, in Section 4.5 we will provide the reader with a conclusion of this chapter's results.

4.1 Modeling Revenue Management under Competition

In this section, we develop a model to describe competitive interactions resulting from RM decisions between two service providers S_1 and S_2. Both firms have a limited capacity of a single resource. They offer the same set of products $J = \{1, \ldots, n\}$, each consuming one unit of the resource, ordered in descending fare $f(j), j \in J$. S_1 and S_2 face each other repeatedly for stages $s \in \mathbb{N}^+$. In each of these stages, they have to decide simultaneously on the subset of products to make available for every time step t of a discretization of the sales period $\{0, \ldots, T\}$. Remaining capacity turns worthless at the end of the sales period T.

RM aims at finding optimal control strategies without being able to observe demand in an uncensored way. Thus, most RM systems rely on a forecasting engine, that tries to reconstruct true customer behavior from limited observations, and an optimization module, that uses this information to compute a control strategy. We will first describe the underlying demand process. Then, we will present different forecast methods that will be used in this chapter, before we will treat solutions of the single-stage optimization problem.

4.1.1 Demand

We assume that the arrival of demand is governed by a Poisson process with an inhomogeneous, but piecewise constant intensity. The behavior of each customer follows a utility-maximizing discrete choice model, conditional on the subset of available products with a lower price than the customer's willingness-to-pay. When faced with an offer set of available products of both service providers, the customer associates disutility values for the price and the product's restrictions. These disutilities are randomly generated, as is the willingness-to-pay of each customer. Choosing an alternative of the offer set is a two-step process:

1. The customer reduces the offer set to products with a lower price than his willingness-to-pay.
2. Of all these offers, the customer chooses the product with the lowest combined disutility.

As a result, we find for all availability situations of S_1 and S_2 the true demand D_k of product j of service provider S_k at time step t of stage s.

4.1.2 Forecast

Talluri and van Ryzin (2004a) have shown for quite general choice models that optimal control strategies take a nested form, i.e. all N optimal (depending on capacity constraints) subsets $J_i \subseteq J$ of offered products may be ordered so that $J_1 \subset \cdots \subset J_N$. Even for demand models that do not possess this property, Fiig et al. (2009) argued that the computational advantages of nested models justify using an approximate nested model instead of working with the original. As a special case of nested models, we restrict ourselves to availabilities that are nested by fare order. That means that service provider S_k always offers a subset $\{1, \ldots, j^{min}\} \subseteq J$, $j^{min} \leq n$, of classes ordered descending by price. Thus, the offer set is entirely described by its cheapest available product.

The true demand $D_k(j, j_k^{min}, j_l^{min}, t, s)$ of product j of service provider S_k at time step t of stage s depends on the availability of substitute products from provider S_k as well as from the competitor S_l, $l \neq k$. Here j_k^{min} denotes the cheapest available product of service provider S_k. Similarly to the true demand, control strategies depend on s and t. However, for the sake of clarity we omit this dependency in our notation whenever possible.

Unfortunately for the service providers, the true demand $D_k(j, j_k^{min}, j_l^{min}, t, s)$ cannot be observed directly. Instead, they have to rely on the observation $\hat{j}_k^{min}(t, s)$ of S_k's lowest available class j_k^{min} at time step t during stage s and the resulting bookings

$$b_k(j, \hat{j}_k^{min}(t,s), \hat{j}_l^{min}(t,s), t, s) = \begin{cases} D_k(j, \hat{j}_k^{min}(t,s), \hat{j}_l^{min}(t,s), t, s) & \text{if } j \leq \hat{j}_k^{min}(t,s) \\ 0 & \text{else} \end{cases} \tag{4.1.1}$$

From these sales data, each firm tries to estimate demand $d(j_k, j_k^{min}, t, s)$. In order to formulate a feasible optimization problem, estimated demand is independent of the competitor situation.

The development of a reliable forecast is a long-term process. For every stage s, the service provider computes an estimate, which is then combined with previous estimates to construct a smooth estimate for future stages. The technique of exponential smoothing is a simple method that uses a weighted average of the current observation $\hat{d}$ and the past reference d with $\alpha \in [0, 1]$ to produce an updated reference.

$$d(\cdot, s) = \alpha \hat{d}(\cdot, s) + (1 - \alpha) d(\cdot, s - 1) \tag{4.1.2}$$

This simple method has been shown to outperform alternative approaches under a variety of settings (Makridakis & Hibon, 2000; Weatherford & Kimes, 2003).

Independent demand Until fairly recently, the standard forecasting engines assumed demand to materialize independent of available alternatives. In an environment in which this assumption holds true, simple pick-up forecasting as described in Equation 4.1.3 can produce robust estimates of high quality (Weatherford & Kimes, 2003).

$$\hat{d}^I(j,\cdot,t,s) = \begin{cases} b_k\left(j, \hat{\jmath}_k^{min}(t,s), \hat{\jmath}_l^{min}(t,s), t, s\right) & \text{if } j \le \hat{\jmath}_k^{min}(t,s) \\ d^I\left(j,\cdot,s-1\right) & \text{else} \end{cases} \tag{4.1.3}$$

Separately for each product j, time step t and stage s, the estimation process of S_k counts observed bookings, as long as j was available. Else the reference is used as an estimate. This yields an estimate $\hat{d}^I$ for the demand in class j at time step t and stage s, which is used to update the reference d^I via exponential smoothing. The independence of this forecast is demonstrated by the independence of 4.1.3 of the chosen availability situation of S_k defined by the lowest class j_k^{min}.

Hybrid demand In the last years, dependent demand models have started to become both more important due to changes in traditional RM industries and more popular due to research advances. These models account for substitution effects between each firm's offered products. A common approach to this problem consists of combining an estimate using the independent demand model d^I and a purely price-dependent model d^D into a hybrid forecast d^H (E. A. Boyd & Kallesen, 2004). Bookings are classified as either price-oriented or product-oriented, depending on whether they occurred in the lowest available class. The previously described method may be applied to the product-oriented subset of bookings, whereas for the price-oriented bookings, usually a functional relationship between price and demand is postulated. This way, it is possible to infer information about demand for any product from sales of possibly different products. A particularly simple and popular technique is the so-called Q-Forecasting (Belobaba & Hopperstad, 2004), where customers' willingness to pay is assumed to follow an exponential distribution.

$$d^D(j, j_k^{min}, t, s) = \begin{cases} d_{base}(t,s)\exp\left(-c_e(t,s)\frac{f(j)}{f(n)} - 1\right) & \text{if } j = j_k^{min} \\ 0 & \text{else} \end{cases} \tag{4.1.4}$$

The base demand d_{base} and the elasticity c_e are estimated as the best fit, which can be achieved using different estimation methods (Bartke, 2013; Cléaz-Savoyen, 2005). In the following, we will describe the estimation technique used in the remainder of this chapter for the hybrid forecast d^H.

Since this is a method designed to work in a monopoly, we will without loss of generality take the perspective of provider S_k, but drop the index denoting the provider to shorten the notation.

First, all observed bookings are classified as price-sensitive or product-sensitive bookings. For this purpose, all bookings observed in the lowest available class will be denoted price-sensitive, the rest product-sensitive. The product-sensitive part will be estimated using the procedure outlined in Equation 4.1.3 for the independent demand d^I. For the price-sensitive part, we assume that demand obeys Equation 4.1.4, so that we only need to estimate two parameters: the elasticity c_e and the base demand $d_{base} = d(n)$.

In order to obtain more stable results for stage s, booking results from the last s_0 stages are pooled and used in the estimation process.

The base demand d_{base} is estimated by reverting Equation 4.1.4

$$\hat{d}_{base}(t,s) = \frac{1}{s_0} \sum_{r=s-s_0-1}^{s-1} \sum_{j=1}^{n} \exp\left(c_e(t,r)\frac{f(j)}{f(n)} - 1\right) b_{price}(j,t,r) \qquad (4.1.5)$$

before applying exponential smoothing to obtain d_{base}.

The estimation of the elasticity c_e requires a little more work. First, we define the indicator function

$$\delta(j,t,s) = \begin{cases} 1 & \text{if } \hat{\jmath}_k^{min}(t,s) = j \\ 0 & \text{else} \end{cases} \qquad (4.1.6)$$

indicating whether a product j was the cheapest available for a given time step t and stage s. This can be used to calculate the weight $w(j,t,s)$ of the observation of price-sensitive bookings of product j for all time steps t and stages $r \in \{s - s_0 - 1, \ldots, s-1\}$ as $w(j,t,s) = \sum_{r=s-s_0-1}^{s-1} \sum_{j=1}^{n} \delta(j,t,r)$. The weight turns out useful when averaging price-sensitive bookings during time step t over stages $r \in \{s - s_0 - 1, \ldots, s-1\}$:

$$\bar{b}(j,t,s) = \frac{\sum_{j,r} \frac{b_k(j,\cdot,\cdot,t,r,)}{d_{base}(t,r)} \delta(j,t,r)}{w(j,t,s)} \qquad (4.1.7)$$

Ideally, we would want to solve the optimization problem

$$\sum_{j=1}^{n} w(j,t,s) \left(\bar{b}(j,t,s) - \kappa(t,s)\exp(-c_e(t,s)\frac{f(j)}{f(n)} - 1)\right)^2 \to \min \qquad (4.1.8)$$

with the scaling parameter $\kappa(t,s)$ under the positivity constraint $c_e(t) > 0$. However, since this non-linear optimization problem is hard to solve, we instead solve the following problem in the log-space

$$\sum_{j=1}^{n} w(j) \left(\log\left(\bar{b}(j,t,s)\right)\right) - \log(\kappa(t,s)) + \exp\left(-c_e(t,s)\frac{f(j)}{f(n)} - 1\right)\Bigg)^2 \to \min, \qquad (4.1.9)$$

under the same positivity constraint $c_e(t) > 0$, where log denotes the natural logarithm.

For the case $\bar{b}(j, t, s) = 0$ of no price-sensitive bookings, it is necessary to introduce an auxiliary parameter. However, by choosing an appropriately high value for the runs used in the estimation s_0, this problem is less likely to appear. The optimization problem 4.1.9 allows an analytical solution, so that we find an elasticity parameter c_e for every time step t. To weaken the influence of outliers, these parameters are then used in a linear regression over the time steps t.

As Bartke (2013) pointed out, the estimation method presented here is rather naive, although it is quite similar to procedures outlined in the literature (Cléaz-Savoyen, 2005; Reyes, 2006).

Dependent demand Another dependent demand forecast due to Bartke (2013) is based on the filtering technique invented by Kalman et al. (1960) among others. As noted in the literature review in Section 2.1.1, Bartke observed that the RM problem of estimating a stochastic process based on censored observations is closely related to estimation problems in signal processing and adapted the well-known Kalman filtering technique to the RM forecasting problem.

Similarly to the previous methods, this forecast is oblivious to competition. Therefore, the presentation of this approach will focus on the monopoly case. In order to save some notation describing this estimation technique, we will take the point of view of service provider S_k, but avoid indexing by k.

The Kalman filtering approach for forecasting can be applied to a host of different demand models. In this case, the demand model is based on a graph describing the relations of booking classes of a single service provider (see e.g. Winter, 2010). Figure 4.1 shows an example of a booking class graph for four classes. This graph is ordered by price and created using product properties such as booking class restrictions. Each node represents a booking class of the service provider, while the edges represent connections of related booking classes. If the availability situation allows it, a part of the demand of a node will flow down outgoing edges to available classes. For each node, the service provider estimates the mean demand $D_k(j, j_k^{min} = j, t, s)$ in this class at time step t during stage s, if all lower classes were not available independently of the competition. For each edge, the firm estimates the mean buydown b, i.e. the demand that is willing to buy the lower class if given the possibility. Additionally to the estimation of the mean demand and buydown, the Kalman filter provides the firm with a covariance matrix Σ^d of all demand and buydown values.

As before, n is the number of classes of a single service provider. Additionally, the graph yields the number of buydown edges m. We introduce the reference vector $\hat{d}^K$, comprising both demand as well as buydown of S_k's products with $m + n$ entries, as well as the

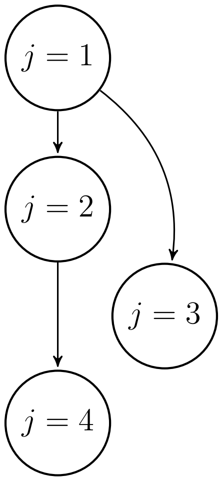

Figure 4.1: Booking class graph

booking vector $\tilde{b}$ with n entries. In order to save some notation, we will not explicitly denote the dependence on t, s and j^{min} wherever possible.

The basic assumptions of a Kalman filtering approach are that the demand follows a noisy Markov process that cannot be observed without an additional measurement error. In this model, the true demand vector $\widetilde{D}$ consisting of means of demand and buydown for S_k's classes is assumed to evolve via

$$\widetilde{D}(s+1) = \widetilde{D}(s) + w_s \tag{4.1.10}$$

with normally distributed random variables $w_s \sim \mathcal{N}(0, \Sigma^p)$. The $(m+n) \times (m+n)$ matrix Σ^p is called the process covariance matrix. According to this model, every booking contains a measurement error

$$\tilde{b}(s+1) = H(j^{min})\widetilde{D}(s+1) + v_s \tag{4.1.11}$$

$$\tag{4.1.12}$$

with $v_s \sim \mathcal{N}(0, \Sigma^b)$, where Σ^b is the $n \times n$ booking covariance matrix. Here, H is a $(m+n) \times n$ matrix, used to reconstrain demand based on the graph and the availability situation. Reconstraining is the process of finding the correct demand based on the availability situation. In Winter's model, the reconstrained demand of any class is the difference between the demand for this class and the sum of the buydown associated to all outgoing edges to available classes. In our case, the availability situation is completely described by the lowest available class j^{min} and therefore H depends simply on j^{min}. However, the reconstraining matrix H can be constructed in more general cases in a similar way.

Estimation of the process covariance matrix Σ^p and the booking covariance matrix Σ^b in a Kalman filtering model is usually hard and an ongoing topic of research (e.g.

Rajamani, 2007). In our implementation of the RM filtering problem, Σ^p is derived using Dijkstra (1959)'s algorithm to calculate the distance matrix Δ of nodes and edges in the graph. Note that Δ describes the shortest distance between any pair of nodes, edges or combination thereof using the booking class graph. We assume that proximity in the graph implies higher stochastic covariation, the absolute value of which can be regulated via two input parameters $\sigma_p > 0$ and $-1 < c_p < 1$.

$$\Sigma^p_{i,j} = \Sigma^p_{j,i} = \max(\hat{d}^K_i \hat{d}^K_j, 1)\sigma_p^2 c_p^{\Delta_{i,j}} \quad \text{for } i,j = 1,\ldots,n, j \geq i \tag{4.1.13}$$

Here, $\hat{d}^K_i$ denotes the i-th entry of $\hat{d}^K$, and likewise $\Sigma^p_{i,j}$ denotes the entry in the i-th row and j-th column of Σ^p.

The booking covariance matrix Σ^b is computed using the reconstraining matrix H, the current reference vector $\hat{d}^K$ and a positive user parameter $\sigma_b > 0$. The main assumption imposed on the structure of Σ^b is the stochastic independence of bookings in different booking classes.

$$\Sigma^b_{i,j} = 0 \quad \text{for } i,j = 1,\ldots,n, i \neq j \tag{4.1.14}$$

$$\Sigma^b_{i,i} = \max((H\hat{d}^K)^2_i, 1)\sigma_b^2 \quad \text{for } i = 1,\ldots,n \tag{4.1.15}$$

Then the update and prediction steps can be combined to yield a new reference for the next stage $s+1$. Note that in contrast to any other forecast described in this chapter, the Kalman filter does not use exponential smoothing to combine different point estimates to an updated reference. Instead, the Kalman filter is a Bayesian method and relies on the calculation of an optimal step size, the Kalman gain. This filter enables an analytical representation of the optimal gain, which helps the speed and robustness of the calculation of the new reference. The Kalman filter produces the minimum squared error estimates, as long as the model is correct and the errors are normally distributed. For errors following a different distribution, the Kalman filter at least generates the minimum squared error linear estimate.

As shown in Equations 4.1.16 – 4.1.19, the innovation covariance S enables us to find the optimal Kalman gain K. K can be used in combination with the innovation vector y to compute the new reference vector $\hat{d}^K(s+1)$ as well as the updated reference covariance $\Sigma^d(s+1)$.

$$S = H\Sigma^d(s)H^t + \Sigma^b \tag{4.1.16}$$

$$K = \Sigma^d(s)H^t S^{-1} \tag{4.1.17}$$

$$\hat{d}^K(s+1) = \hat{d}^K(s) + Ky \tag{4.1.18}$$

$$\Sigma^d(s+1) = (I_q - KH)\Sigma^d(s) + \Sigma^p \tag{4.1.19}$$

Here H^t denotes the transpose of the matrix H.

Note that the reference $\hat{d}^K$ contains values both for the demand of all booking classes as well as for the buydown between connected pairs of booking classes. The demand $d^K(j, j_k^{min}, s)$ for a single class j of S_k given an availability situation defined by the lowest class j_k^{min} can be found by applying the reconstraining matrix H to the reference vector $\hat{d}^K$

$$d^K(j, j_k^{min}, s) = (H(j_k^{min})\hat{d}^K(s))_j \tag{4.1.20}$$

Competitor-dependent demand To account not only for dependence on own substitute products, but as well on the competitor's available products is beyond the capabilities of current forecasting systems. However, any RM forecasting system automatically takes the competitor situation into consideration in an implicit way during the estimation process, because it is indistinguishable whether customers do not arrive due to their low willingness-to-pay or due to the competitor offering a preferable product. This indistinguishability leads to complications in the estimation process of any real-world forecasting system. It is however possible in a simulation to use knowledge of the true demand process in a so-called psychic forecast that represents an idealization of a real-world forecast. Similarly to Isler and Imhof's (2008) psychic forecast, the forecast d^P is constructed using the true demand as observed under the competitor's availability situation during time step t in stage s. However, whereas Isler and Imhof use maximum-likelihood estimation to estimate customer choice probabilities and propose an appropriate updating scheme, we follow a different approach similarly to the other forecasting techniques presented in this section. Recall that $\hat{\jmath}_l^{min}(t, s)$ denotes the observation of S_l's lowest available class j_l^{min} at time step t during stage s. In every time step t during each stage s we use the true demand given the observed availability of the competitor S_l as this stage's estimate for the demand

$$\hat{d}^P(j, j_k^{min}, t, s) = D_k(j, j_k^{min}, \hat{\jmath}_l^{min}(t, s), t, s) \tag{4.1.21}$$

at this time step and use exponential smoothing to update the estimates.

4.1.3 Optimization of Single Stage

Optimal control strategies for both independent and dependent demand forecasts can be found via dynamic programming. This means that for every combination of remaining capacity x and time t the value function $U_t(x)$ can be calculated in a recursive way. The boundary conditions for $U_t(x)$ are

$$U_T(x) = U_t(0) = 0, \tag{4.1.22}$$

since remaining capacity at the end of the sales period is worthless, as is remaining time with no capacity left.

We rescale the demand forecast d by refining the time discretization, so that at most one arrival is expected per time step. By abuse of notation, we keep denoting elements of this new discretization t. For notational clarity in our treatment of the single-carrier single-stage optimization, we drop the k indices, as well as the stage variable s. Furthermore, we omit the time variable t in d.

Let us first focus on the independent demand case. We calculate the value function $U_t(x)$ recursively via the following Bellman equation:

$$U_{t-1}(x) = \max_{j^{min}} \left\{ \sum_{j \leq j^{min}} d(j)(f(j) + U_t(x-1)) + (1 - \sum_{j \leq j^{min}} d(j))U_t(x) \right\} \tag{4.1.23}$$

$$= U_t(x) + \sum_j d(j)(f(j) - \Delta U_t(x))^+, \tag{4.1.24}$$

where we introduced the notation $\Delta U_t(x) = U_t(x) - U_t(x-1)$ and $x^+ = \max(x, 0)$. The optimal control policy takes the simple form

$$\text{accept request for product } j \iff f(j) \geq \Delta U_{t+1}(x) \tag{4.1.25}$$

Fiig et al. (2009) showed that the dependent demand case can be reduced to the independent demand case by substituting regular demand d and fares f with marginal demand $\tilde{d}$ and marginal fares $\tilde{f}$ respectively. We calculate $\tilde{d}$ and $\tilde{f}$ with the help of total demand Q and total revenue R:

$$Q(j) = \sum_{i \leq j} d(i, f(j)) \tag{4.1.26}$$

$$R(j) = \sum_{i \leq j} d(i, f(j))f(j) \tag{4.1.27}$$

$$\tilde{d}(j) = \Delta Q(j) \tag{4.1.28}$$

$$\tilde{f}(j) = \frac{\Delta R(j)}{\Delta Q(j)} \tag{4.1.29}$$

Note that for the independent demand case, the marginal fare is just the regular price and the marginal demand just the regular demand. The resulting Bellman equation takes the same form as in the independent demand case 4.1.23:

$$U_{t-1}(x) = U_t(x) + \sum_j \tilde{d}(j)(\tilde{f}(j) - \Delta U_t(x))^+ \tag{4.1.30}$$

Consequently, the optimal control strategy for dependent demand forecasts consists of a transformation of 4.1.25:

$$\text{accept request for product } j \iff \tilde{f}(j) \geq \Delta U_t(x) \tag{4.1.31}$$

In order to be able to find the non-cooperative Nash equilibrium, every player S_k needs complete information about the true demand D as well as the competitor's inventory level x_l, $l \neq k$. Using this complete information, the value function U^k of player S_k in a Nash equilibrium can be computed recursively via the Bellman equation

$$\begin{aligned} U_{t-1}^k(x_k, x_l) = \max_{j_k^{min}} \Bigg\{ & \sum_{j \le j_k^{min}} (D_k(j, j_k^{min}, j_l^{min}) f(j) + U_t^k(x_k - 1, x_l)) \\ & + \sum_{j \le j_l^{min}} D_l(j, j_k^{min}, j_l^{min}) U_t^k(x_k, x_l - 1) \\ & + \left(1 - \sum_{j \le j_k^{min}} D_k(j, j_k^{min}, j_l^{min})\right) U_t^k(x_k, x_l) \\ & + \left(1 - \sum_{j \le j_l^{min}} D_l(j, j_k^{min}, j_l^{min})\right) U_t^k(x_k, x_l) \Bigg\}. \end{aligned} \tag{4.1.32}$$

Here j_l^{min} is the result of S_l's optimization problem, since it is sufficient to react to the best response of other players in a Nash equilibrium. Consequently, the Bellman Equation 4.1.32 needs to be solved simultaneously for both players, which complicates the search for a Nash equilibrium. However, we know that the dynamics of the non-cooperative Nash equilibrium strategies can be replicated by using a psychic forecast implicitly accounting for dependency on competitor's products (Isler & Imhof, 2008). This psychic forecasts translates the information about demand and competitor's inventories into the forecast. After applying fare transformation to this forecast, we can put it in 4.1.30 and get control strategy 4.1.31.

The exact jointly optimal solution can be found by maximizing U^1+U^2. An approximation to the jointly optimal strategies is given if both service providers do not offer any product cheaper than an appropriate threshold product (Isler & Imhof, 2010). There exists $c \in \{1, \dots, n\}$ so that the cooperative policy can be approximated by

$$\text{accept request for product } j \iff \tilde{f}(j) \ge \Delta U_t(x) \wedge j \le c \tag{4.1.33}$$

The optimal threshold product $c \in J$ is found as the cheapest product a firm would offer in a monopoly without capacity constraints.

4.2 Multi-stage Strategies

The control strategies presented in Section 4.1 focus on the single-stage game only. However, as pointed out in Chapter 3, the view as a repeated game may be more appropriate to model RM under competition. As outlined in Chapter 3, we find close similarities between the RM game and the thoroughly studied IPD. We have argued in the literature review in Section 2.2.2 that a model as small and simple as the IPD can be used to establish a better understanding of the situation and help finding solutions to the underlying, more complex, problem (Grüne-Yanoff, 2009). Taking this perspective, we use this section to present strategies for the repeated game, mainly building on successful strategies from the IPD.

4.2.1 Heuristic to Adapt Strategies from the IPD to the Repeated RM Game

As shown in Figure 4.2, each player in the IPD chooses one of only two actions—cooperate and defect—for each stage. The payoffs satisfy the conditions $T > R > P > S$ and $2R > T + S$, so that mutual cooperation maximizes the combined payoffs, but mutual defection is the unique single-stage Nash equilibrium.

	Cooperate	Defect
Cooperate	R,R	S,T
Defect	T,S	P,P

Figure 4.2: Normal form of the prisoner's dilemma

In the RM game, each player uses a combination of forecasting and optimization to choose a policy that makes a subset of n products available during the booking horizon of every stage. As described in Section 4.1.3, this subset can depend on the time elapsed in the booking horizon as well as on the bookings observed by the player. In order to reduce complexity, we assume these subsets to be characterized by their cheapest product (see Section 4.1).

Although the repeated RM game offers far more possibilities in each stage than the IPD, the dynamics of the games are very similar. Choosing an aggressive price lower than the competitor leads to higher revenue than sharing revenues with the competitor. However, a more cooperative stance can lead to a higher payoff than mutual aggression. More specifically, we have shown that the use of a standard RM system leads to policy 4.1.31, which maximizes revenue in a monopoly. In a duopoly, the use of the psychic forecast d^P enables this policy to converge to the non-cooperative best response, thus the most efficient way of choosing an aggressive price (see Isler & Imhof, 2008). However, if both players follow this kind of strategy, revenues will drop far lower than in the case of both players following a more cooperative strategy such as 4.1.33 (see Isler & Imhof, 2010).

In order to adapt strategies from the IPD to the repeated RM game, we have to classify the actions of the RM game into cooperation and defection. This is necessary both for actions to play as well as for observing the competitor's action.

The effects of the policies 4.1.31 and 4.1.33 as described above provide us with natural candidates for providers to play 'cooperate' or 'defect'. Thus, we say a service provider S_k using demand forecast d plays $\mathrm{COOP}(d)$ in stage s at time t, if S_k follows the cooperative control strategy 4.1.33. Likewise, we say S_k plays $\mathrm{DEFECT}(d)$ in stage s and time t, if S_k follows the non-cooperative control strategy 4.1.31.

However, in contrast to the prisoner's dilemma, a service provider still needs to interpret the competitor's observed actions in order to determine whether the competitor was defecting or cooperating. As Isler and Imhof (2010) showed, in cases of very high demand, playing $\mathrm{DEFECT}(d^P)$ and $\mathrm{COOP}(d^P)$ may result in the same availability situation, making them indistinguishable to the competitor. Furthermore, keeping accurate and reliable information about competitor availabilities is a challenging task. To account for this, we introduce noise into observations in our model. Recall that $\hat{\jmath}_l^{min}$ denotes the observation of S_l's lowest available class j_l^{min}. Then, for a positive error parameter ε, there is a positive probability of an observation error

$$\mathbb{P}(\hat{\jmath}_l^{min} = j_l^{min} + 1) = \begin{cases} \varepsilon & \text{if } j_l^{min} < n \\ 0 & \text{if } j_l^{min} = n \end{cases} \tag{4.2.1}$$

$$\mathbb{P}(\hat{\jmath}_l^{min} = j_l^{min} - 1) = \begin{cases} \varepsilon & \text{if } j_l^{min} > 0 \\ 0 & \text{if } j_l^{min} = 0 \end{cases} \tag{4.2.2}$$

where we implicitly introduced the notation $j = 0$ for the situation in which no products are available at all. This observation can be used to interpret competitor behavior. We say that player S_k perceives competitor S_l to be cooperating at time t, if and only if the lowest available product does not underprice the cooperative threshold product c, i.e. if $\hat{\jmath}_l^{min}(t) \leq c$. This is decided independently for every time step $t \in \{0, \dots, T\}$ in the booking horizon.

Additionally players can assign each player a reputation as a tool to help interpret the competitor's actions. In the IPD, this concept was first introduced by Sugden (1986). The evolution of a player's reputation is displayed in Figure 4.3. Each player starts with a good reputation. Any player being observed as defecting against a competitor with a good reputation will lose his good reputation. The only way to regain a good reputation is by cooperating. Note that a player may retaliate against an opponent with a bad reputation without losing his good reputation.

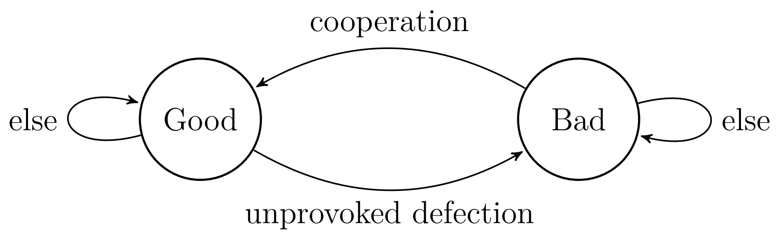

Figure 4.3: Evolution of reputation

4.2.2 Strategies

The interpretation of competitor behavior as cooperating or defecting allows us to adapt the strategies from the IPD presented in Section 2.2.2. In a fixed stage s, these strategies can exhibit different behavior at different time steps t, depending on the classification of competitor behavior at time step t of the previous stage $s-1$. Using COOP and DEFECT, we can construct repeated-game strategies as repetitions of single-stage strategies.

ALLD This is the repetition of the single-stage non-cooperative solution (Axelrod & Hamilton, 1981). We say S_k plays ALLD(d), if he always plays DEFECT(d). Note that ALLD does not make use of the categorization of competitor behavior. In a RM context, Isler and Imhof (2008) argued that a standard RM system should converge towards ALLD, but Cooper et al. (2009) showed that this depends on the forecast in use.

ALLC Similarly to ALLD, the repetition of the single-stage cooperative solution is called ALLC. We say S_k plays ALLC(d), if he always plays COOP(d). Similarly to ALLD, ALLC is independent of the categorization of the competitor's behavior. Isler and Imhof (2010) showed that both players using this strategy in a RM game leads to an approximation of the joint optimum and thus avoids the Competitive Spiral Down.

Tit for Tat Tit for Tat (TFT) is a strategy that has been shown to perform extraordinarily well in the IPD (Axelrod, 1984). A player plays TFT(d), if he starts with COOP(d) in all time steps of the first stage game. In the following stage games, he decides for every time step t whether to cooperate or defect. He plays COOP(d), if he perceived the competitor as cooperating in the previous stage, and plays DEFECT(d), if he perceived the competitor to be defecting. Despite its success in the IPD with perfect information, TFT is vulnerable to errors (Molander, 1985). Variations of TFT, that have been shown to be robust in the presence of errors, are Generous Tit for Tat (GTFT) and Contrite Tit for Tat (CTFT) (Wu & Axelrod, 1995).

Generous Tit for Tat In order to cope with noise, GTFT introduces a bias towards cooperation into the mechanics of TFT: GTFT cooperates with a fixed probability γ even though it should defect according to TFT. In the IPD, the optimal generosity probability for GTFT $\gamma = \min(\frac{2R-S-T}{R-S}, \frac{R-P}{T-P})$ depends on all four possible payoffs of the single-stage game Molander (1985). Thus, this strategy requires detailed knowledge of the expected revenue dependent on the availability situation of the competing service provider. This information is obviously beyond the scope of a realistic forecast, but it is not even included in the psychic forecast d^P, which covers only the case in which the competition plays DEFECT. We use psychic knowledge of the customers in the scenario to calculate the optimal generosity. In our simulations, we will use the optimal level of generosity of the IPD as well as different fixed generosity parameters and analyze their performance via means of simulation.

Contrite Tit for Tat CTFT is a robust TFT variant that does not need additional parameters, but relies on players' reputations instead (R. Boyd, 1989). If a CTFT player has a bad reputation, he will cooperate in order to regain a good reputation. Otherwise, if a CTFT player has a good reputation, he will cooperate against players with a good reputation, but defect against players with a bad reputation until they cooperate. Thus, CTFT is similar to TFT in that both rely on reciprocity to enforce cooperative behavior.

PAVLOV We will not restrict our analysis to TFT-variants. In the IPD, the PAVLOV strategy is an example of a win-stay, lose-shift strategy, where players try to avoid the both low payoff for mutual defection P and the sucker payoff S (Nowak & Sigmund, 1993). A player S_k following this strategy cooperates at time step t if the players have either both cooperated or both defected in the previous stage at this time step. In the RM game, this means that S_k plays COOP(d) at time step t during stage s, if $\hat{\jmath}_l^{min}(t, s-1) \leq c \wedge \hat{\jmath}_k^{min}(t, s-1) \leq c$ or if $\hat{\jmath}_l^{min}(t, s-1) > c \wedge \hat{\jmath}_k^{min}(t, s-1) > c$. Otherwise, a PAVLOV player will play DEFECT(d).

Matching A simple multi-stage strategy in the RM game not derived from the IPD is price matching, which has seen use in research and practice due to its simplicity (S. P. Anderson & Schneider, 2007; Hess & Gerstner, 1991). In the single-stage game, simultaneous price-setting prevents one player matching the competitor's actions. In the repeated game, a player can easily copy the competitor's last action. We say a player S_k plays MATCH, if he accepts a request for product j at time step t in stage s if and only if $j \leq \hat{\jmath}_l^{min}(t, s-1)$. This strategy does not require any demand forecast. Note that we call matching the RM mechanism of copying the competitor's last availability situations. In other publications, the term may refer to the pricing mechanism of copying all the prices filed by its competitor (e.g. Evans & Kessides, 1994; Nomani, 1990).

Underbidding Similarly to MATCH, we say S_k plays UNDER, if he accepts a request for product j at time step t in stage s if and only if $j \leq \hat{j}_l^{min}(t, s-1) + 1$. This is another strategy that does not rely on a demand forecast.

4.2.3 Properties of Successful Multi-stage Strategies

In this chapter, we analyze strategies for the repeated RM game. Due to its resemblance to the IPD, we want to build on the thorough treatment of the IPD in the literature described in Section 2.2.2. In the analysis of his IPD tournaments, Axelrod (1984, p.54) claimed that a successful strategy in the IPD should be

nice : cooperative unless provoked,

forgiving : able to cooperate after the opponent has defected,

retaliating : able to punish defectors,

clear : easy to understand for any opponent.

This characterization of successful strategies was motivated by the success of TFT in Axelrod's IPD tournaments. Therefore, it comes as no surprise that all Tit for Tat variations such as TFT, GTFT and CTFT possess these properties. The success of the PAVLOV strategy several years later however showed that Axelrod's properties can be relaxed by a successful strategy (Nowak & Sigmund, 1993). PAVLOV is neither as forgiving nor as retaliating as the variations of Tit for Tat. Instead, the success of this strategy is based on exploiting suckers.

Nevertheless, Axelrod's characterization is helpful to identify relevant strategies of the RM game. The ability to forgive and retaliate with a clear agenda is of special importance for human players. Forgiveness and retaliation are necessities for any competitive strategy, whereas clarity simply represents an attractive feature. Thus, it is not surprising that the majority of human players in the IPD choose either GTFT or PAVLOV as strategies (Wedekind & Milinski, 1996). Keeping in mind that the RM game is played between businesses, we note that competitive strategies are not the result of random mutation, but instead the consequence of human analysts' choices or of an automated system designed by humans. This is a difference to some applications of the IPD, which have aimed to model behavioral patterns in nature that may mutate completely at random. Therefore, we argue that the scope of opposition strategies in our analysis is wide enough, as it encompasses not only the most successful strategies (see e.g. Boerlijst et al., 1997b; Imhof et al., 2007; Nowak & Sigmund, 1993), but also those most often chosen by human opponents (Wedekind & Milinski, 1996). We can thus restrict ourselves to the study of the strategies presented in Section 4.2, although there is an infinite amount of possible strategies.

Among these strategies, we aim to find a strategy that is optimal in a sense that we will specify in the following. Obviously, a simple repetition of the single-stage strategies does not represent an ideal non-cooperative strategies of the repeated RM game. While ALLD cannot be exploited by any strategy, it displays the Competitive Spiral Down effect against aggressive competitors. On the other hand, ALLC completely avoids the Competitive Spiral Down effect, but is exploited by ALLD. Ideally, we would like to find a strategy of the repeated RM game combining the strengths of the single-stage equilibrium strategies, while avoiding their weaknesses.

In order to avoid the danger of exploitation, a prospective strategy should constitute a non-cooperative Nash equilibrium of the repeated game when paired against itself. In contrast to ALLD, the strategy should furthermore avoid the Competitive Spiral Down, ideally by reaching the joint optimum. Therefore, the ideal competitive strategy in RM should be part of a jointly optimal non-cooperative Nash equilibrium of the repeated game. Whether such a strategy exists for the repeated RM game is part of our investigation.

Thankfully, the folk theorem assures that a repeated game with an infinite horizon can have far more Nash equilibria than its constituent single-stage game. Fudenberg, Levine, and Maskin (1994) showed that imperfect public signals—such as flawed but public competitor price observations—allow that any feasible payoff vector Pareto-dominating mutual defection can be realized through a non-cooperative Nash equilibrium of the repeated game.

The dependence on past time steps via capacity constraints and on stages via exponential smoothing prevents an analytical discussion of repeated-game strategies. Furthermore, since there is no limit on the amount of possible strategies, we cannot compare all strategies against each other using stochastic simulations. Instead, we restrict ourselves to analyzing necessary conditions for a jointly optimal Nash equilibrium. To test whether a strategy is the best response against itself, we will run it against the most aggressive strategy possible, ALLD. If the strategy can be exploited by ALLD, it cannot be a part of a Nash equilibrium. Proximity to the joint optimum is tested by comparison of the symmetric setup of a strategy against itself with ALLC vs. ALLC.

4.3 Simulation Environment

In order to evaluate strategies in RM, we will use simulations throughout this chapter. As a tool for these simulations, the author has extended and used the simulation environment REvenue Management Training for Experts (REMATE). Although REMATE was created as an airline RM simulator, its use is not restricted to the airline case. In this section, we will give an overview of the relevant functionality of REMATE to this thesis, using the notation introduced in the previous sections.

This RM simulator is the product of a cooperation of Lufthansa with the universities of Paderborn, Heidelberg, Kaiserslautern and the Freie Universität Berlin. The main aim of REMATE is to provide insights via simulation, where testing in the real world might prove costly. The implementation was performed by a software development company. REMATE is written mostly in Java, using MySQL as database management system. The code is available to Lufthansa and researchers from cooperating universities, which allows researchers to build extensions on top of thoroughly tested versions that provide a vast functionality. Based on Law and Kelton (2000) and work with PODS, Frank et al. (2007b) formulated principles for the design of RM simulations, which served as guidelines in the design of REMATE. The simulator is subject to continual improvement and extensions, which renders previous descriptions by Zimmermann et al. (2011) and Gerlach et al. (2013) inaccurate. Thus, in this section we will give an overview of the state of the implementation of REMATE used in this dissertation.

From its inception, REMATE was devised as a multi-purpose simulation tool. As the name suggests, it should not only provide a platform for researchers, but also serve as a training and decision support tool for revenue managers. Therefore, the simulator includes modes specifically targeted at analysts. Guided by a simpler interface, they can perform analyses using scenarios that model a typical workday. Alternatively, they can test competitive strategies in games against each other. In these games, analysts occupy roles of different airlines in a single market, aiming at maximizing revenue of a set of flights for a single departure. However, since the game covers only a single departure, the long-term effects of competition cannot be analyzed. Thus, this method is not suited to investigate the aspects of competition focused on in this dissertation.

However, the functionality of REMATE is not reduced to these aspects targeting RM analysts. Additionally to answering specific requests from revenue managers, a researcher may also use REMATE to model and analyze a broad range of scenarios. While complexity needed to be reduced in favor of clarity for the use as a decision support tool for revenue managers, there is no need for such a restraint for the use as a research tool. Thus, it is possible to select arbitrary schedules, products, prices, capacity restrictions as well as several RM methods. Similarly to the decision support mode, scenario construction, simulation and analysis can be performed via a graphical user interface.

Structure

Figure 4.4 shows a simplified view of the RM process of two competitors in REMATE.

For each service provider, the basis of a scenario is provided by the supply consisting of schedules, products, prices and capacities. This, along with expected demand, is fed into an optimization routine, which generates optimal control values stored in an inventory. REMATE provides the possibility to modify these values manually via user influences,

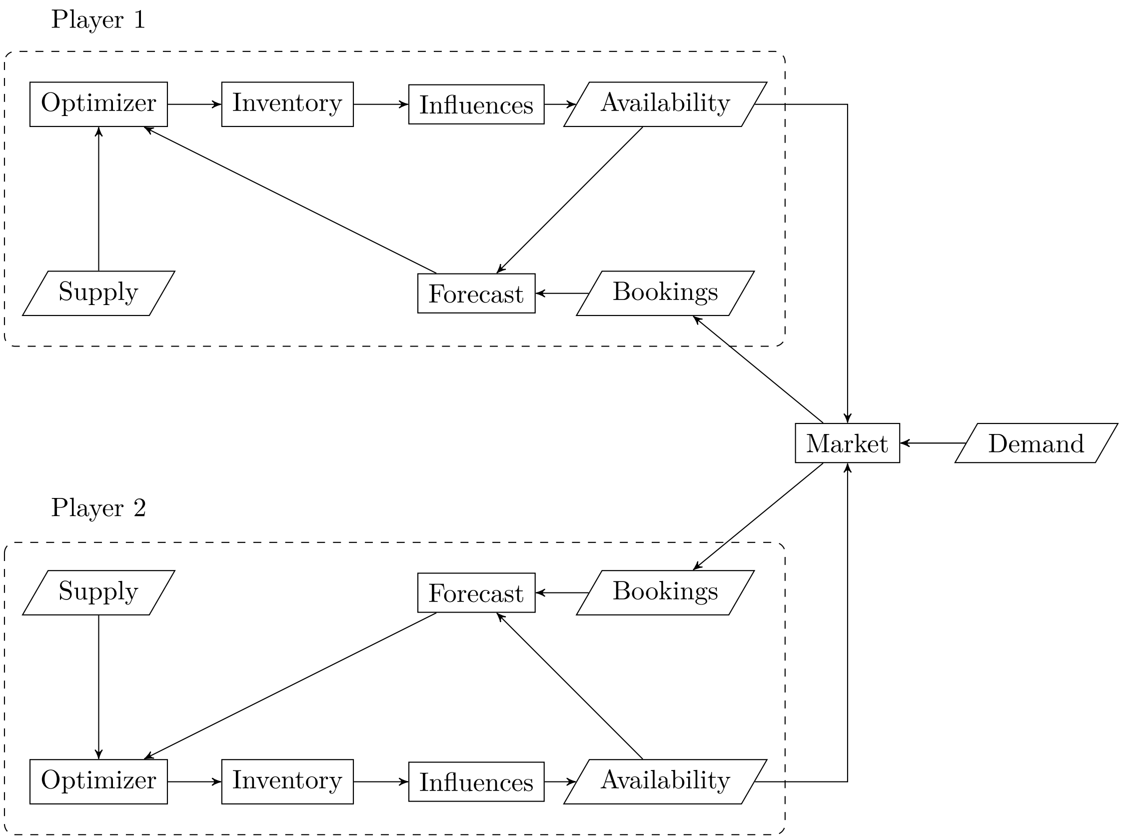

Figure 4.4: Flowchart of the RM process in REMATE

before they are met by the demand in the market module. Resulting bookings as well as the underlying availability situation are used to estimate and forecast demand for the following sales periods.

Supply

Like any model, REMATE is based on simplifications of the complexities of the real world, so that the RM problem becomes feasible. The core of any scenario consists of a set of service providers operating a schedule and offering products at various prices for a set of customers over a fixed booking horizon. For each simulation, REMATE runs repetitions of this sales period for a given number of runs.

For our purposes, any product consumes one unit of a resource and consists of a set of restrictions and a price. Being a RM simulator, REMATE has modeled the pricing process in a simpler way than real firms do. In REMATE, a product is linked globally with a set of restrictions, so that each service provider's products have the same restrictions independent of the market in which they are sold. For each simulation, each resource is tied to a fixed capacity configuration. In the airline case, this corresponds to a flight being served with an aircraft with a fixed amount of available seats.

Forecast

Although demand is spread continuously over time, the processes of both forecasting and optimizing require a discretization of time. This is done via sets of time steps. The forecaster yields expected demand between two time steps, while the optimizer performs a reoptimization given forecasted and actual bookings at each time step.

REMATE provides an implementation of the independent demand forecast d^I, the hybrid demand forecast d^H and the dependent demand forecast d^K described in Section 4.1.2. Additionally, the author implemented the psychic forecast d^P, presented in Section 4.1.2 as well.

All of these forecasting methods provide updating schemes for the estimated demand. In a simulation as well as in the real world, a starting point is required. Furthermore, the researcher may choose to use the initialization method throughout the simulation instead of estimating demand. This may serve as a benchmark in a monopoly, with the constant initialization providing an upper bound. The initialization method we employ in this chapter uses knowledge of the generated demand in the underlying scenario in order to provide an airline with the best possible fit of the employed demand model. For the initialization, demand per product and time step is computed as a linear combination of first and second choice products. Because of the significant differences in the underlying demand models of the forecasters, the demand per product and buydown between products is determined in a different way for each demand model. This is particularly problematic for independent demand models, which cannot represent the true dependence of the customer choice on the offer set. In this case, the initialization counts customers in the most expensive product that they can afford, which evidently is a very optimistic choice from the firm's point of view. Due to these differences between the customer model in REMATE and the demand models used in estimation, even this best case scenario is not perfect. However since it represents the best fit of the—possibly oversimplified—demand model, it represents a very good starting point. Therefore, this initialization method introduces a bias towards not changing anything or at least towards slow-changing methods and parameter choices.

Optimization

The optimization module in REMATE consists of several submodules taking care of finding the best capacity configuration for each flight, taking care of cancellations and no-shows by overbooking, and solving the seat inventory problem. In this sketch of the functionality of REMATE, we will focus on the seat inventory problem, since this mirrors the focus of this dissertation.

For this purpose, REMATE offers an implementation of the optimization techniques based on dynamic programming described in Section 4.1.3. These optimization methods cover dependent as well as independent demand models, so that they are applicable to all forecasting methods implemented in REMATE.

Additionally, it is possible to choose to use First Come, First Serve (FCFS) instead of an optimization technique. FCFS does not perform any optimization. Instead, it makes every product available, as long as the necessary resources are available. If a provider relies on the availabilities produced by FCFS, its performance may serve as lower bound on any strategy. Alternatively, the resulting availability situation can be adjusted manually. This complete manual control is useful to model low-cost-carriers.

Inventory

The inventory stores all necessary information to implement the control policies described in Section 4.1.3. It serves to calculate the availability situation based on control values calculated by the optimization module, prices and capacities from the supply as well as observed bookings. More precisely, the inventory used in our experiments relies on bid price vectors representing the value of a free seat given an arbitrary number of already observed bookings.

User Influences

In many simulation studies, the availabilities stored in the inventory after the optimization process are identical to the ones used in the market module (e.g. Belobaba & Wilson, 1997; Gorin & Belobaba, 2004). However, in real life, human analysts adjust availabilities on a daily basis in order to react to competition, unforeseen events or changes in the market structure (Mukhopadhyay, Samaddar, & Colville, 2007; Weatherford, 2009; Zeni, 2003). Since REMATE is intended to help in training and decision support of revenue managers, the simulator includes a possibility to alter the control values in the inventory after the optimization. We will use this module to implement the competitive strategies for the repeated game presented in Section 4.2.

Demand

Demand is generated following a Poisson process with an intensity that is piecewise constant between time steps in the booking horizon. The behavior of each customer follows a two-step process as described in Section 4.1.1: Given that the product's price does not exceed the customer's willingness-to-pay, the final product is chosen using a utility-maximizing discrete choice model.

Market

In the market module, booking and cancellation decisions are generated. These are the result of the generated customers following the demand model and the calculated availability situation of all providers.

4.4 Simulation Experiments

We analyze two symmetric service providers offering a single resource in a common market. The constant demand volume counts 30 customers, requests are uniformly distributed over 23 time steps. In order to approximate Isler and Imhof's (2008) results, the customers' willingness-to-pay is normally distributed with $\mu \approx 100$ and $\sigma \approx 40$. Furthermore, the products of each service provider S_k are associated with a restriction that serves to introduce uncertainty into the customers' decisions. This provider-specific restriction is normally distributed with $\mu_k \approx 2$ and $\sigma_k \approx 0.6$. As described in Section 4.2, demand and prices are set up similarly to the prisoner's dilemma. Therefore, the resource is sold at 16 different price points between 2 and 260 EUR as displayed in Table 4.1. These prices may seem unrealistic, but they guarantee that it is optimal for each service provider to undercut the competitor. However, as soon as the price offered falls below

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
f(j)	260	147	109	87	71	58	47	38	30	23	17	12	8	5	3	2

Table 4.1: Prices used in the simulation

the price of the cooperative threshold product c, both players' undercutting results in lower revenue for each.

Figure 4.5 shows the expected revenue for a service provider in a monopoly depending on the lowest available product. According to this graph, the cooperative threshold product is $c = 4$ with $f(c) = 87$ for our choice of parameters.

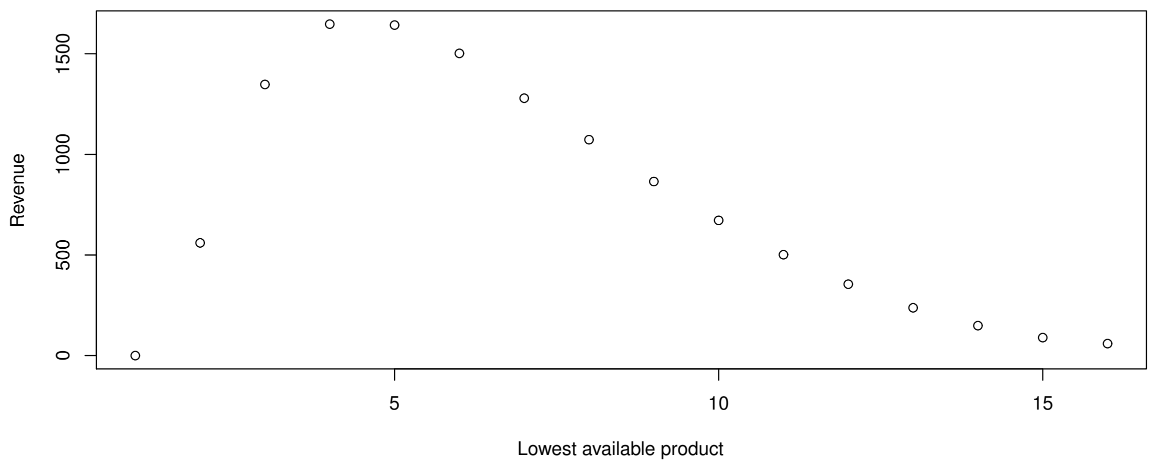

Figure 4.5: Revenue in a monopoly

Since we consider the dependence of competitive effects on the ratio of capacity and demand, we are interested in the revenue for capacity values from 1 to the amount of the total demand. For each possible capacity value, we execute 200 repetitions of the single-stage game. We will use the first 100 stages as a burn-in period and average the results over the last 100 stages, when the system has reached a relatively stable state. After a set of trial runs, we determined 200 and 100 respectively to be the perfect number for this purpose.

4.4.1 Psychic Forecast

Throughout this section, we will use the psychic forecast d^P presented in Section 4.1.2 with the exponential smoothing parameter $\alpha = 0.2$. To save some notation, we will omit the forecast when specifying strategies and write e.g. ALLD instead of ALLD(d^P).

Single-Stage Nash Equilibria

The heuristics using the psychic forecast d^P and the cooperative threshold product c described in Section 4.1 approximate the non-cooperative Nash equilibrium strategy and the cooperative, jointly optimal strategy of the single-stage game, leading to the strategies ALLD and ALLC for the repeated game.

In order to evaluate the quality of our approximation to the cooperative solution, we introduce the strategy JOI. In our symmetric setup, the exact jointly optimal solution is achieved when both players share the market fairly and each player acts optimally on his share. The strategy JOI describes the RM control of a service provider acting as a monopolist on half of the market equipped with a perfect knowledge of the demand. This perfect knowledge of the demand is achieved by using the psychic demand initialization of REMATE. Since there is no competition, this psychic forecast is sufficient to describe the demand, and we do not need our competitive psychic forecast d^P.

Similarly, we introduce the lower bound LOW, which represents the average revenue of a service provider, if both players always make the lowest price available. This behavior is equivalent to using FCFS as a RM control.

Figure 4.6 shows from left to right the average revenue over capacity of the simulation results of ALLC vs. ALLC, ALLC vs. ALLD and ALLD vs. ALLD. For comparison, we have added the revenue of a single service provider when both use LOW and when both follow the exact jointly optimal solution JOI.

As a repetition of the single-stage non-cooperative Nash equilibrium strategy, ALLD also represents a non-cooperative Nash equilibrium strategy of the repeated game. However, ALLD vs. ALLD leads to the Competitive Spiral Down. With growing capacity, each player has a greater incentive to undercut the competitor as a best response, leading to a very low revenue comparable to LOW. Note that this strategy, along with the incentive to undercut, is a direct consequence of the RM forecasting mechanism and may thus be interpreted as natural behavior for real RM systems.

ALLC vs. ALLC approximates JOI, thus avoiding the Competitive Spiral Down effect. However ALLC is exploited by ALLD, proving that ALLC is not the best response against itself in the repeated game.

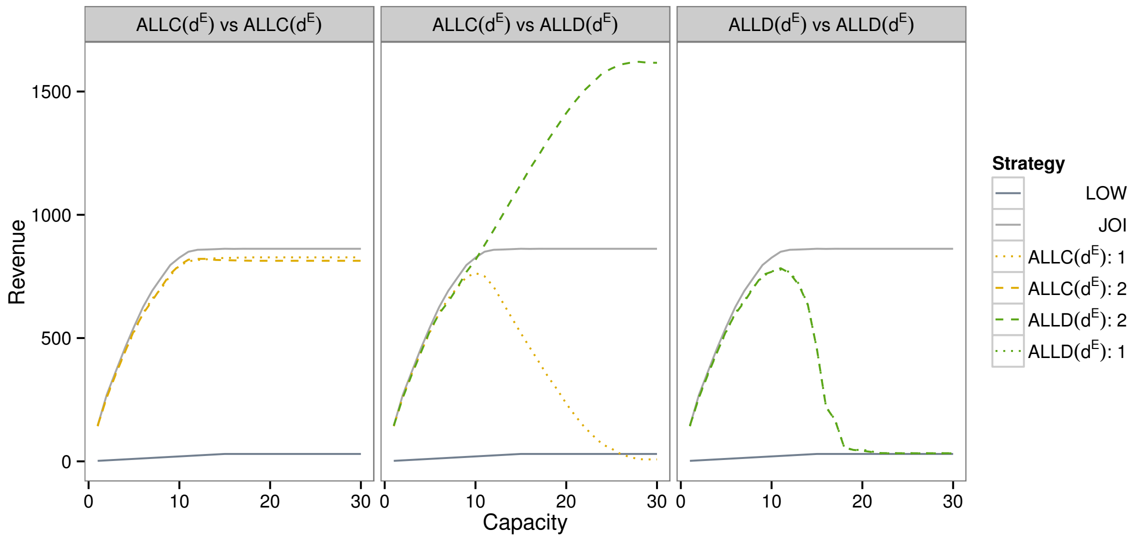

Figure 4.6: ALLD and ALLC

Although we have employed a different demand model as well as a slightly different implementation of the psychic forecast d^P, our computational results for ALLD vs. ALLD resemble the findings of Isler and Imhof (2008), and our results for ALLC vs. ALLC resemble those of Isler and Imhof (2010). We interpret this reproduction of results as a validation of the basic ideas behind the construction of the approximations of the single-stage solutions. This indicates that this approach can be used flexibly in different, possibly complex settings without amounting to a lot of implementational effort.

Tit-for-Tat

Plots 4.7, 4.8, 4.10, 4.11 and 4.12 display the average revenue of a service provider using various strategies against ALLC, ALLD and in the symmetric matchup. In these figures, we display the average payoff of ALLD vs. ALLD as an example of Competitive Spiral Down, as well as the revenue in ALLC vs. ALLC as an example of a jointly optimal strategy pair.

Without observation errors, i.e. $\varepsilon = 0$, Figure 4.7 shows that TFT fulfills the necessary conditions of an ideal strategy formulated in 4.2.3. The result of TFT vs. TFT is very close to ALLC vs. ALLC, so that there is no Competitive Spiral Down. Since it is not exploited by an aggressor, the RM version of TFT may well be a Nash equilibrium.

However, the introduction of a positive observation error probability $\varepsilon = 0.1$ leads to the Competitive Spiral Down seen in Figure 4.8. This shows that TFT is not robust against observation errors in the RM game. This effect is even more extreme than in

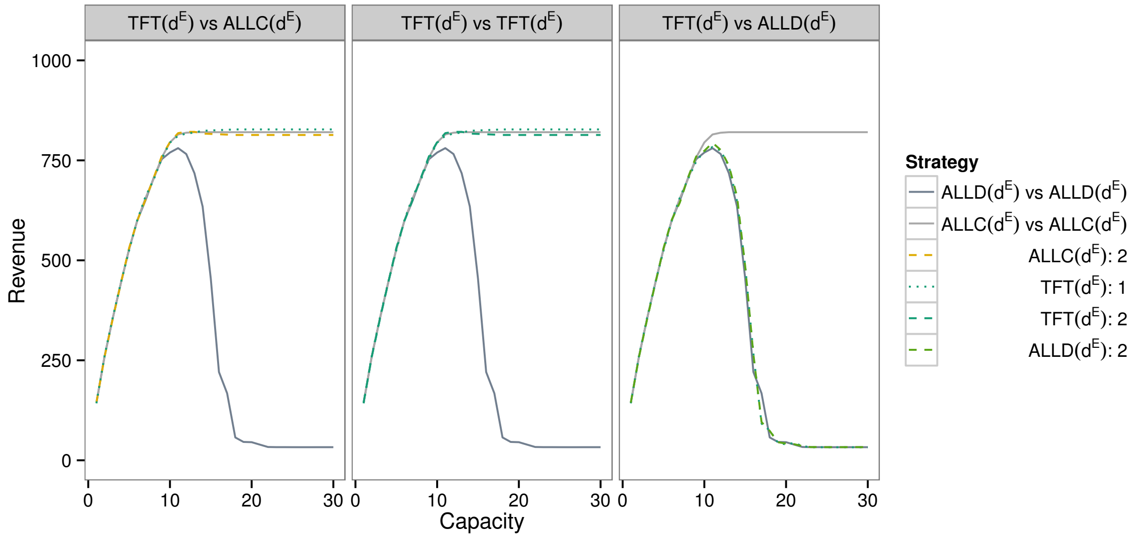

Figure 4.7: TFT for $\varepsilon = 0$

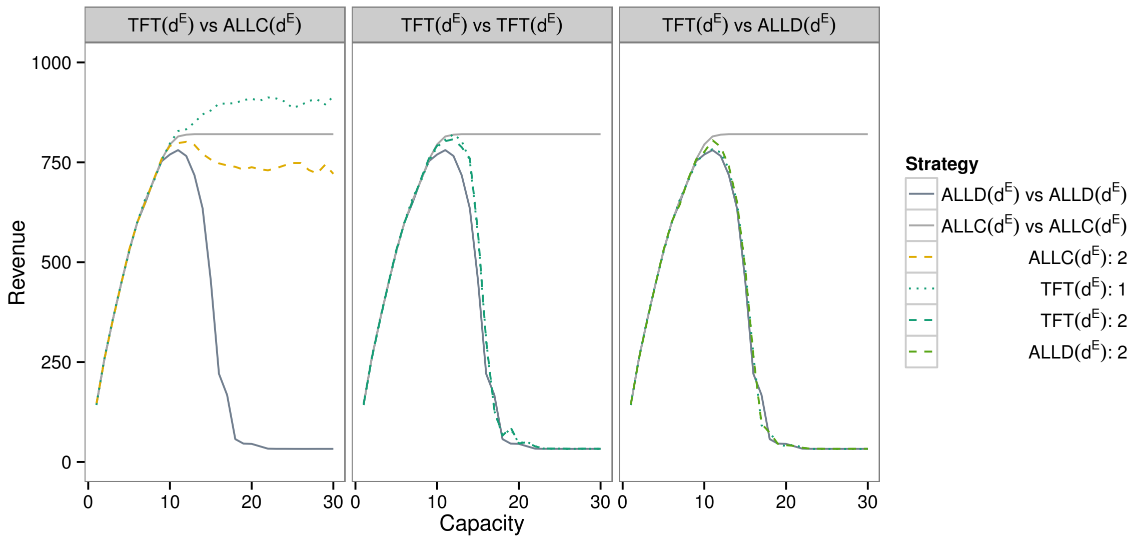

Figure 4.8: TFT for $\varepsilon = 0.1$

the IPD, where observation errors turn TFT vs. TFT into a random walk on the payoff space Molander (1985).

Generous Tit-for-Tat

To compensate for the observation error ε, we analyze the performance of more robust strategies such as GTFT. Note that Figure 4.8 shows that a lack of generosity leads to the Competitive Spiral Down, which is at its worst at the rightmost point of the graph, i.e. at the maximal capacity configuration. The introduction of generosity is supposed to weaken the Competitive Spiral Down effect, although this can lead to GTFT being exploited by aggressors.

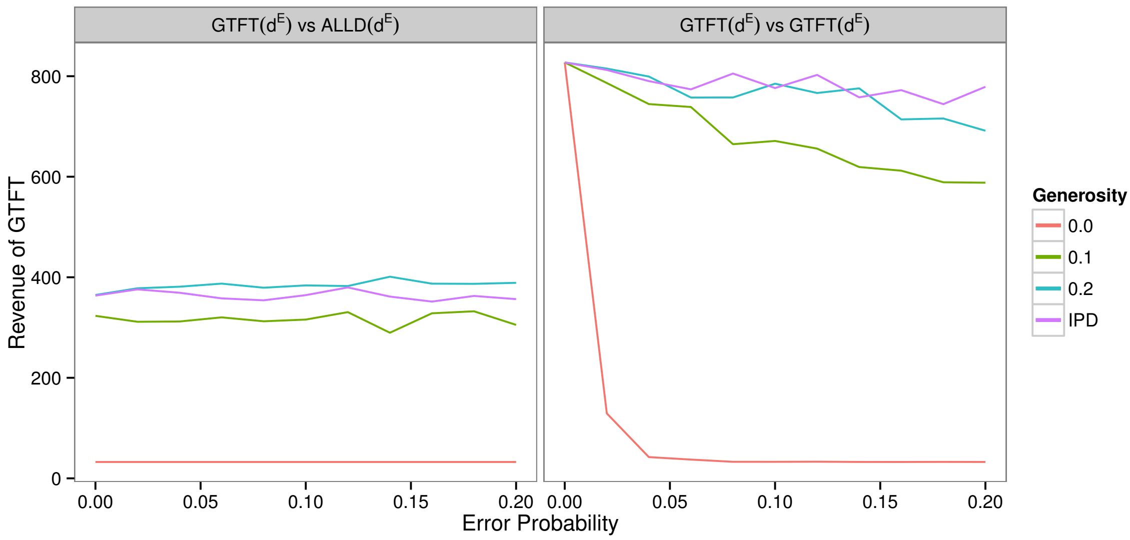

Figure 4.9: GTFT for range of generosity probabilities, $\varepsilon \in [0, 0.2]$ and CAP=30

Figure 4.9 shows the revenue of GTFT when paired against ALLD or GTFT at the maximal capacity configuration for a range of generosity parameters γ and observation error probabilities ε. Additionally to $\gamma \in \{0, 0.1, 0.2\}$, we have used the optimal level of generosity in the IPD as determined by Molander (1985) denoted by $\gamma = \gamma_{IPD}$. As is expected against an aggressor ALLD, the revenue of GTFT seems to be independent of the error probability for all parameter choices. For GTFT against GTFT, we find a decline in revenue for growing error rates for the fixed generosity probabilities, while the optimal generosity of the IPD seems to be relatively stable at a high level. For most error probabilities, the dynamic choice $\gamma = \gamma_{IPD}$ outperforms the static choices in the symmetric matchup, while we find the results of $\gamma = \gamma_{IPD}$ against an aggressor between the results of $\gamma = 0.1$ and $\gamma = 0.2$. In the rest of this chapter, we will always use $\gamma = \gamma_{IPD}$, whenever we show simulation results of GTFT.

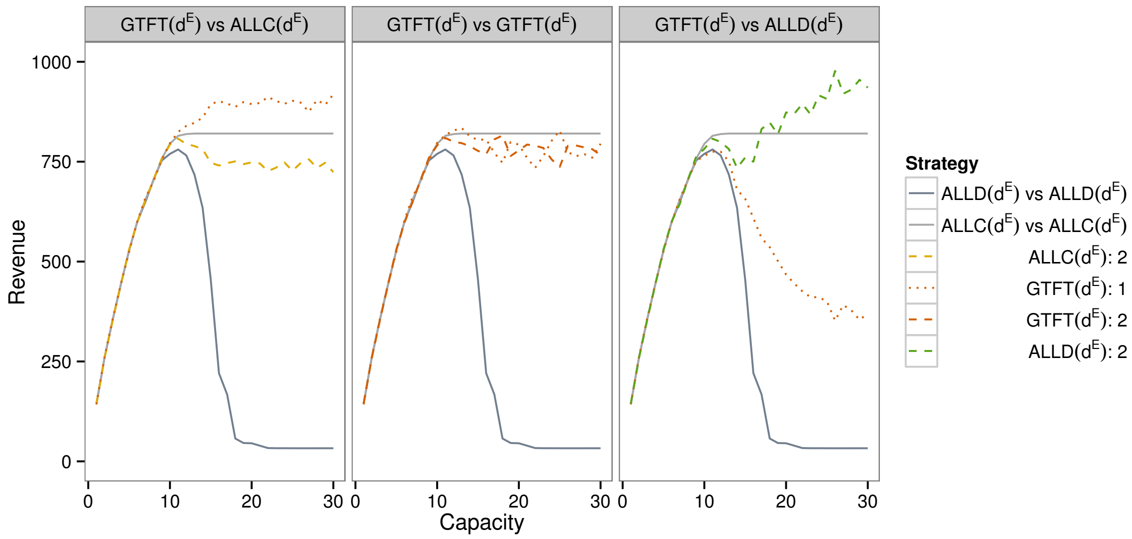

Figure 4.10: GTFT for $\varepsilon = 0.1$

Figure 4.10 shows that GTFT avoids the Competitive Spiral Down, but provides an incentive for the competitor to exploit its cooperative stance. The level of generosity can be used to control the trade-off between the two necessary conditions. Nevertheless, it is impossible to fulfill both conditions, since any positive level of generosity is exploited by an aggressive strategy such as ALLD, and Figure 4.8 shows that no generosity at all leads to a complete competitive spiral-down.

Contrite Tit-for-Tat

In contrast to GTFT, where each player can cooperate after a mistake of either player, each player corrects only his own mistakes in CTFT. For CTFT to be robust, it is necessary that both players base their decisions on the same possibly flawed observations, which corresponds to the importance of public signals for the folk theorem. Figure 4.11 shows that the introduction of reputation leads to a strategy that fulfills the necessary conditions for a jointly optimal Nash equilibrium of the repeated game formulated in Section 4.2.3.

Pavlov

Figure 4.12 displays the simulation results of the PAVLOV strategy for an observation error probability of $\varepsilon = 0.1$. Similarly to the Tit for Tat variations, PAVLOV approximates the jointly optimal solution in the symmetric matchup. However, PAVLOV vs. ALLD

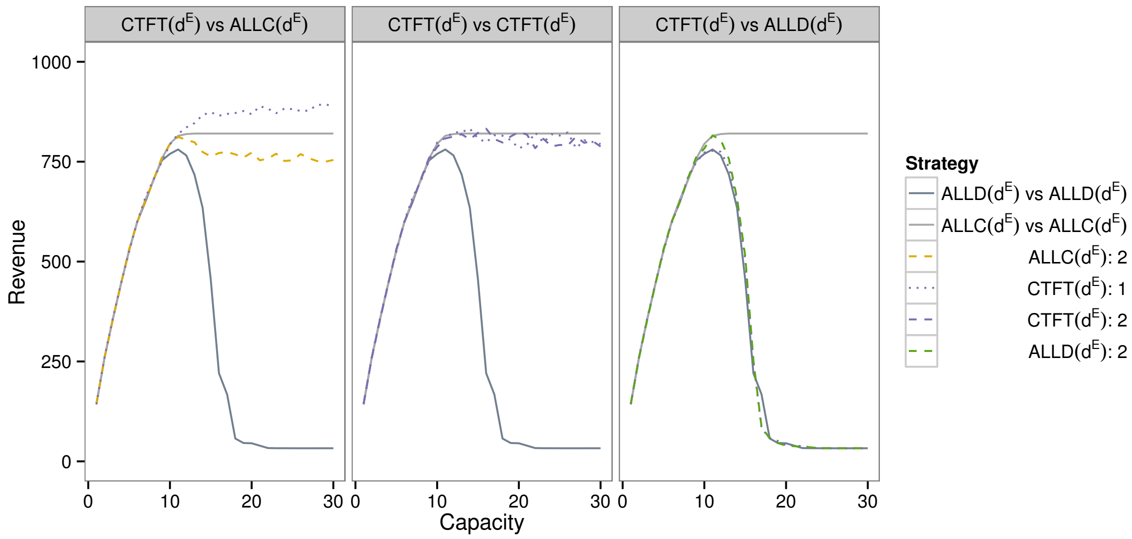

Figure 4.11: CTFT for $\varepsilon = 0.1$

shows that PAVLOV is no Nash equilibrium. On the other hand, PAVLOV exploits ALLC, resulting in a far higher revenue than any Tit for Tat variant against a cooperative player. Thus, in a predominantly cooperative environment, PAVLOV can outperform cooperative players as well as Tit for Tat players, despite not fulfilling our desired properties summarized in Section 4.2.3. Note that similarly to the IPD, observation errors are necessary for PAVLOV to exploit ALLC in the repeated RM game.

Robust Strategies vs. Non-robust Strategies

While Figure 4.8 demonstrated the fragility of TFT in the presence of observation errors, Figures 4.10, 4.11 and 4.12 showed the robustness of the GTFT, CTFT and PAVLOV strategies. However, these three robust strategies implement a different logic to achieve robustness. CTFT corrects only own mistakes, while GTFT can correct both players' mistakes. PAVLOV follows a completely different intuition, but its error-correcting features apply to both players' actions as well. In a symmetric matchup, there is no disadvantage in having each player correct his own mistakes. As a result, Figure 4.11 demonstrates a superior performance of CTFT compared to the results of GTFT displayed in Figure 4.10. However, players may encounter competitors employing multi-stage strategies without the necessary robustness. Therefore, a strategy performing well against less robust strategies may prove beneficial.

Figure 4.13 shows each of GTFT, CTFT and PAVLOV paired with a competitor using a simple TFT strategy. The results suggest that the generosity of GTFT is best suited to

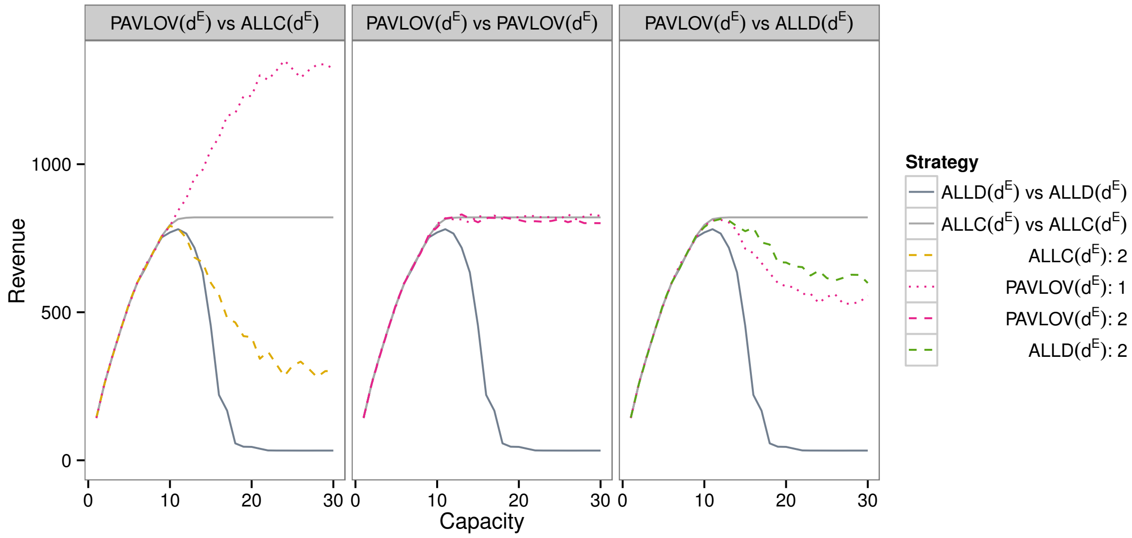

Figure 4.12: PAVLOV for $\varepsilon = 0.1$

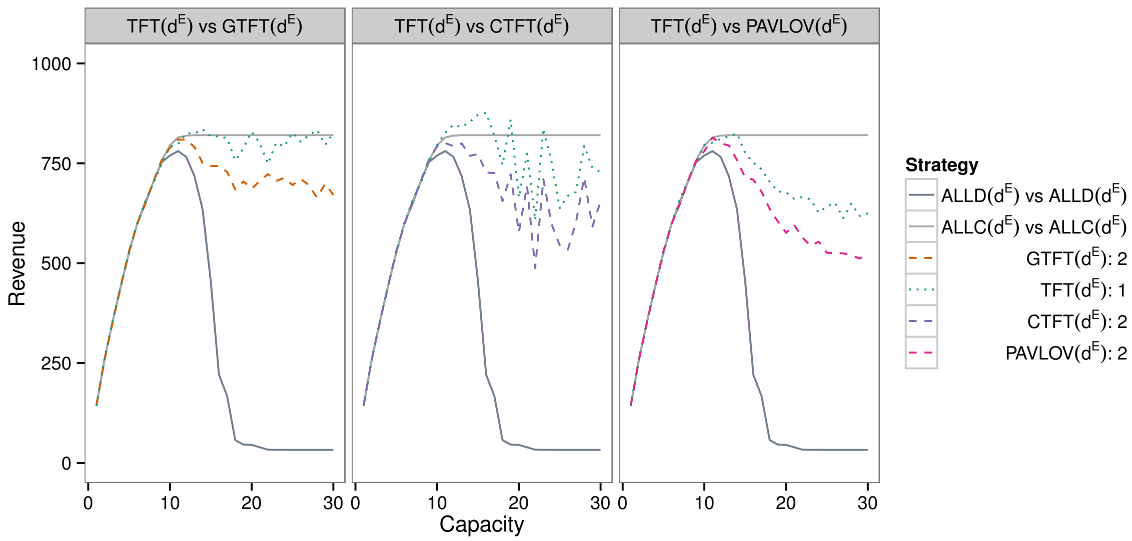

Figure 4.13: Robust strategies against TFT for $\varepsilon = 0.1$

this environment, since GTFT outperforms both CTFT and PAVLOV in this situation.

Irrational Strategies

Since they are not based on a forecast and instead only use competitor prices as input, both MATCH and UNDER depend completely on their competitor's strategy. As a consequence, MATCH necessarily leads to a similar outcome as the symmetric matchup. Thus, against strategies which are successful in a mirror matchup, MATCH can produce good results. In particular, Figure 4.14 demonstrates that MATCH approximates the non-cooperative solution against ALLC and the cooperative solution against ALLD.

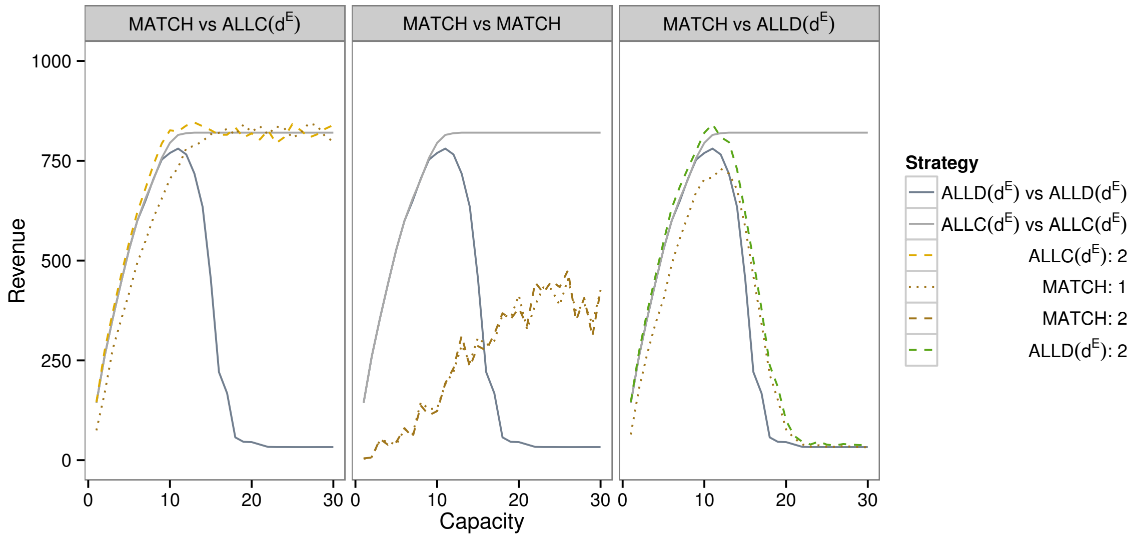

Figure 4.14: MATCH for $\varepsilon = 0.1$

However, in a mirror matchup against another price matching service provider, the results can be quite poor. As demonstrated by Figure 4.14, the effects of flawed competitor price monitoring cause MATCH vs. MATCH to become a random walk. Consequently, the revenue of the matching service providers is unlikely to come close to the optimal payoff achieved by the non-cooperative Nash equilibrium for high demand to capacity ratios. However, the payoff of two matching players can be higher than the payoff of a pair of ALLD or even TFT players for large capacities, since a random walk at least does not exhibit the Competitive Spiral Down behavior.

Figure 4.15 displays the results of an underpricing service provider in competition with cooperating and non-cooperating opponents and in the mirror matchup. Against ALLC, UNDER exploits its competitor's cooperative stance, which leads to a similar revenue as for ALLD against ALLC. When facing an always defecting ALLD opponent, UNDER is

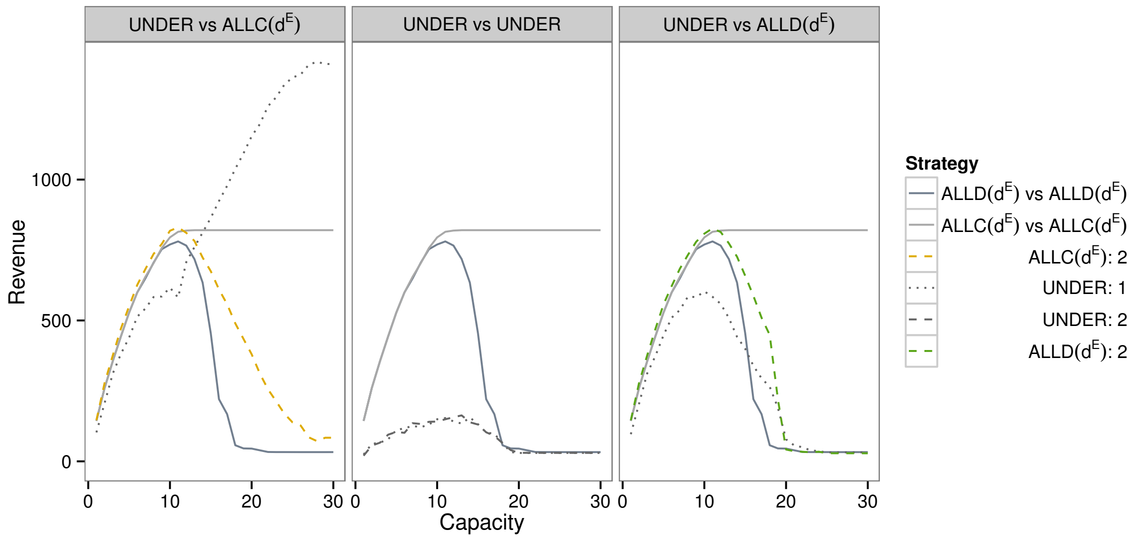

Figure 4.15: UNDER for $\varepsilon = 0.1$

exceedingly aggressive, leading to a slightly lower revenue than its competitor's. Against another underpricing service provider, both players end up with extremely low revenue close to the lower boundary set by LOW. In fact, only observation errors prevent UNDER vs. UNDER from falling on the level of LOW as displayed in Figure 4.6.

We examine the performance of previously successful strategies against irrational strategies. Since MATCH produces results close to the symmetrical mirror matchup, we do not provide simulation results for this case. For this case, keep in mind that in the mirror matchup with observation errors, TFT suffered due to its lack of robustness, whereas the other strategies GTFT, CTFT and PAVLOV approximated the jointly optimal outcome. In contrast to this, the results of a simulation pairing a rational firm using TFT, GTFT, CTFT or PAVLOV against an irrational firm relying on UNDER are not evident. Figure 4.16 demonstrates that against an underpricing competitor, the performance of all of the strategies TFT, GTFT, CTFT and PAVLOV is similar. For all these strategies, the rational providers earn far more than the irrational competitor, and slightly more than in the complete Competitive Spiral Down caused by ALLD vs. ALLD, which is displayed in the background. Thus, the strategies GTFT, CTFT and PAVLOV that have proved successful so far, represent an appropriate choice against an irrational competitor as well.

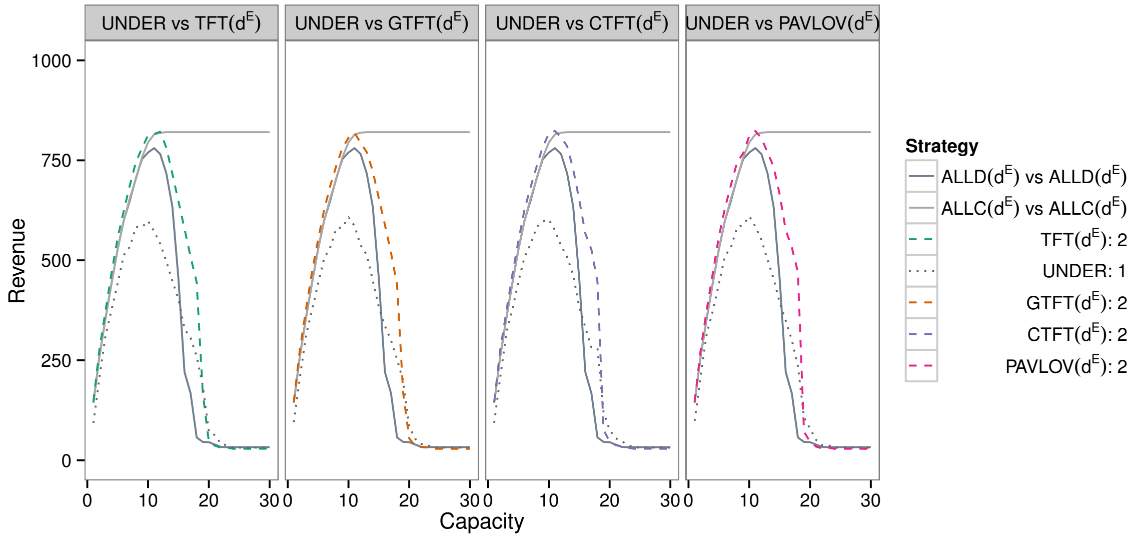

Figure 4.16: UNDER vs. various repeated game strategies for $\varepsilon = 0.1$

4.4.2 Standard Forecasts

In this section, we will analyze the duopoly using real world forecasts. In contrast to this, in the previous section, we relied on the psychic forecast d^P, yielding an approximation to the single-stage non-cooperative Nash equilibrium strategy. Combined with a tacit collusion scheme, this psychic forecast also produced an approximation to the single-stage jointly optimal solution. These single-stage strategies represent the basis of all repeated-game strategies presented in Section 4.2 except for the irrational strategies MATCH and UNDER. The psychic forecast d^P is therefore crucial to construct strategies such as TFT, GTFT, CTFT and PAVLOV. However, this forecast uses information about the real underlying demand, which cannot be observed by any service provider in a realistic setting. The psychic forecast d^P represents an optimum of knowledge, thus freeing us from the limitations of current techniques.

Real life RM forecasts strive to use past observations to produce an estimate of expected customer behavior that will lead to the highest possible revenue. As pointed out in Section 4.1.2, these methods were designed for the monopoly setting and consequently do not incorporate competitive effects in an explicit way, although there is an implicit consideration of the competitor's actions because of the interaction of the competing service providers. The widespread use of these techniques is based on the belief that over time the implicit learning of competitive effects will suffice to reach an optimal strategy. The goal of these techniques is to find the best response to given market conditions, possibly including one or more competitors. Success in this endeavor means that the service providers' forecasts converge to a Nash equilibrium, which highlights the importance of our previous analysis. Cooper et al. (2009) have shown that this approach may converge to the non-cooperative Nash equilibrium. However, they also showed that this can also converge to the cooperative solution or to an altogether different outcome.

In this section, we will examine whether RM systems using standard forecasts exhibit the same behavior in a competitive setting as systems using the psychic forecast d^P. In particular, we want to find out whether standard forecasts lead to the Competitive Spiral Down, which is necessary in order to approximate the single-stage Nash equilibrium strategy. Similarly, we are interested in the effect of tacit collusion used with standard forecasts. We investigate whether it is possible to use some of these forecasts to construct successful strategies for the repeated game. Furthermore, we want to analyze the performance of standard RM forecasts against the psychic forecast d^P. The performance against the psychic forecast can indicate how the standard forecast could fare against a competitor with a superior competitive RM system.

As representatives of standard RM forecasts, we will use the independent demand forecast d^I, the hybrid forecast d^H and the dependent demand forecast based on the Kalman filter d^K presented in Section 4.1.2. Thus, we have a range of forecast systems with a varying degree of complexity; from the simple independent demand forecast to more complex dependent demand forecasts.

As explained in Section 4.1.2, these estimation techniques require different parameters. For the independent demand forecast d^I, we will use the exponential smoothing parameter $\alpha = 0.2$, since this leads to stable results within a reasonable amount of stages. When using the hybrid demand forecast d^H, we will rely on the previous $s_0 = 5$ stages for the estimation in a given stage s and once more use the exponential smoothing parameter $\alpha = 0.2$. The dependent demand forecast d^K requires no exponential smoothing, since this is a Bayesian method based on Kalman filtering. Here, we will use the relative process deviation $\sigma_p = 0.05$ and correlation factor $c_p = 0.6$ to construct the process covariance matrix and the relative booking deviation $\sigma_b = 0.3$ to create the booking covariance matrix. These parameters have been shown to perform well in a variety of scenarios during testing with REMATE.

Similarly to the figures in the previous section, the x-axis represents the capacity of each service provider in the simulation. Every single combination of integer capacity value and revenue represents the average revenue of a service provider in a simulation consisting of 200 stages for this fixed capacity, where the first 100 stages serve as a burn-in period. Between integer values of capacity we use a linear interpolation to enhance the readability of the graph.

ALLD and ALLC

Similarly to the previous section, we start by examining ALLD and ALLC with standard RM forecasts. In order to facilitate comparisons between graphs, we continue displaying the revenue of mutual defection ALLD vs. ALLD as well as of mutual cooperation and ALLC vs. ALLD with the psychic forecast d^P in the background.

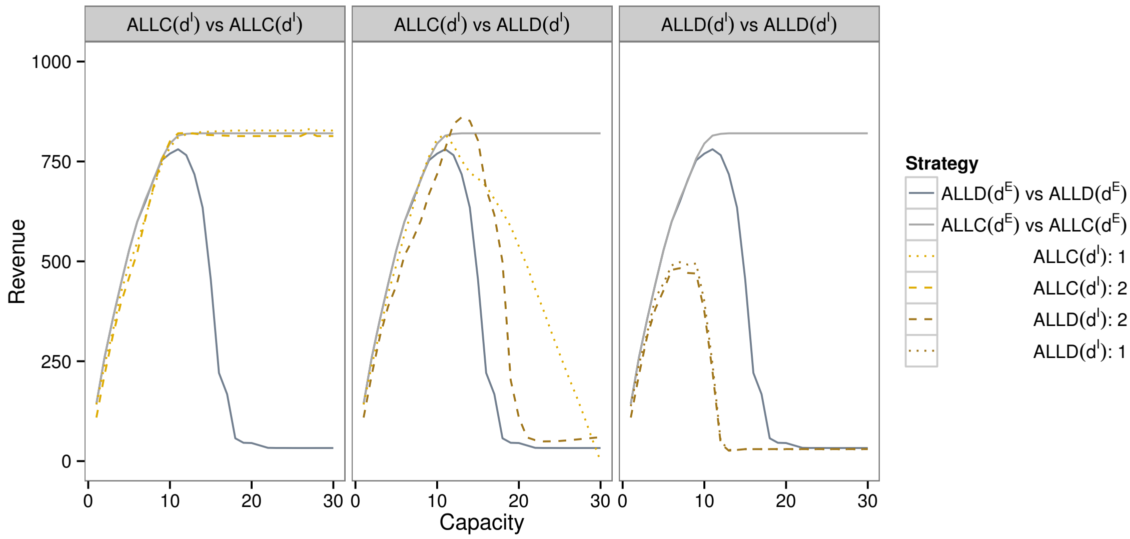

Figure 4.17: Independent demand forecast d^I

Figure 4.17 displays all combinations of ALLC and ALLD using the independent demand forecast d^I. In the mutually non-cooperative case ALLD(d^I) vs. ALLD(d^I), we find a stronger decline than explained by the Competitive Spiral Down effect. This is due to the inability of the independent demand forecast to model the customer choice process adequately, resulting in the classical Spiral Down effect as described by Cooper et al. (2006). Similarly, Isler and Imhof (2008) have also shown that the Spiral Down can be even more disastrous than the Competitive Spiral Down.

Nevertheless, mutual cooperation ALLC(d^I) vs. ALLC(d^I) works well, since the value of the cooperative threshold product c is known. In fact, for high capacity values, the revenue resulting of cooperation using the independent demand forecast is virtually identical to the revenue for mutual cooperation using the psychic demand forecast d^P. The graph in the center, displaying the exploitation matchup ALLC(d^I) vs. ALLD(d^I), shows a very different behavior to the psychic forecast case, where the defector ALLD(d^P) was able to exploit the cooperator ALLC(d^P). This difference in behavior is a consequence of the Spiral Down effect, which prevents ALLD(d^I) from keeping a price level just below the cooperating competitor using ALLC(d^I). Instead, the Spiral Down effect causes the price level to deteriorate, so that ALLD(d^I) cannot exploit a blindly cooperating competitor.

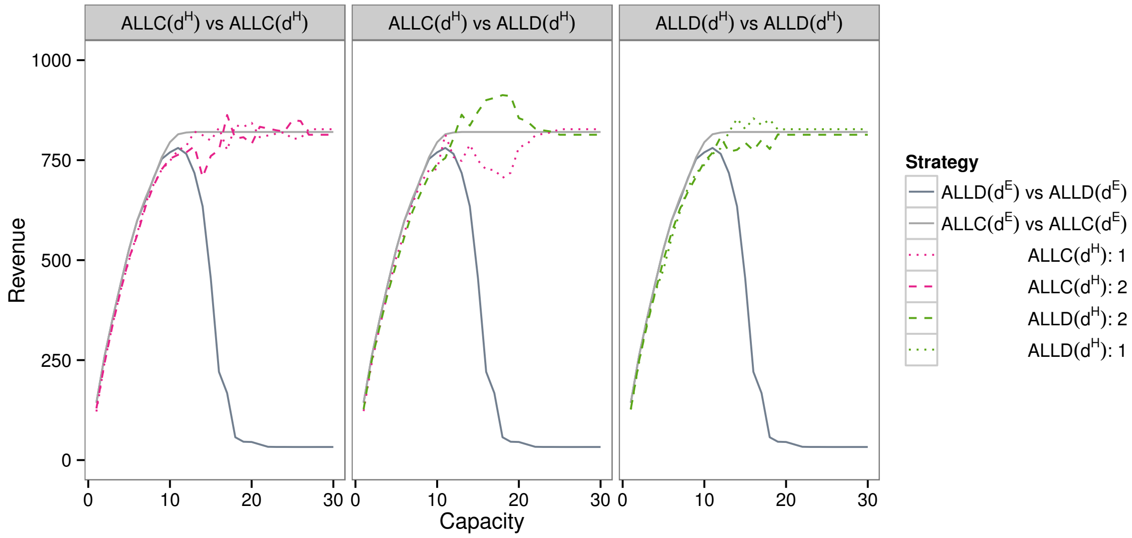

Figure 4.18: Hybrid demand forecast d^H

Figure 4.18 presents the simulation results for all ALLC(d^H) and ALLD(d^H) combinations. In the hybrid case, neither combination of ALLC(d^H) and ALLD(d^H) shows any signs of the Competitive Spiral Down effect. However, although this keeps revenues high, it is rather a sign of inaccurate estimation. In any case, the result is close to the jointly optimal solution. Thus, if all service providers were to provide this forecast, the Competitive Spiral Down effect would be rendered insignificant without implementing any elaborate multi-stage strategy such as those presented in 4.2.

Figure 4.19 shows the results of ALLC(d^K) and ALLD(d^K). In contrast to the results for d^I and d^H, the results for d^K are far closer to the results of the psychic forecast d^P. In particular, mutual defection leads to the Competitive Spiral Down, although the effect is weakened compared to the psychic case. Furthermore, we find that unilateral cooperation using ALLC(d^K) can be exploited by an aggressor using ALLD(d^K). Similarly to the other forecasts, the results for ALLC(d^K) vs. ALLC(d^K) show that mutual cooperation leads to similar results as in the psychic forecast case ALLC(d^P) vs. ALLC(d^P), indicating that the success of cooperation depends more on the knowledge of the correct value c for the tacit collusion than on the underlying forecast.

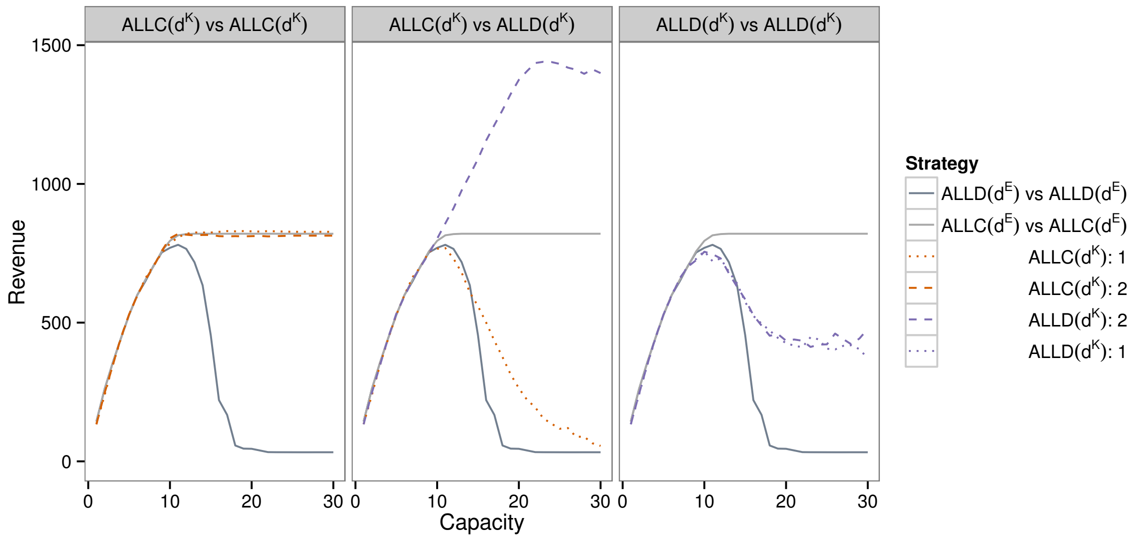

Figure 4.19: Dependent demand forecast d^K

Successful Strategies

Similarly to our analysis for the psychic forecast d^P, we will also investigate the performance of the strategies TFT, GTFT, CTFT and PAVLOV, that rely on our heuristic to transfer strategies from the IPD to the RM game. Since the previous results in this section showed that the other forecasts cannot adequately replicate the behavior of the psychic forecast d^P and are thus not appropriate for the use by these strategies, we will concentrate on the dependent demand forecast d^K,

Figures 4.20 – 4.23 show simulation results of the revenue of a pair of providers over varying capacity restrictions, where one provider uses TFT, GTFT, CTFT or PAVLOV against a competitor following the cooperative strategy ALLC(d^K), the non-cooperative strategy ALLD(d^K) and in the mirror matchup.

We find similar results to the experiments with the psychic forecast d^P in Section 4.4.1 displayed in Figures 4.8 and 4.10 – 4.12. Figure 4.20 shows that TFT(d^K) is not robust against observation errors, so that the mirror matchup leads to the Competitive Spiral Down. However, TFT(d^K) is cooperative against pure cooperation ALLD(d^K) and not exploited against the aggressor ALLD(d^K). In Figure 4.21, we observe that GTFT(d^K) prevents the Competitive Spiral Down in the mirror matchup, but is exploited by the aggressor ALLD(d^K). Figure 4.22 demonstrates that CTFT(d^K) is jointly optimal against cooperation and in the mirror matchup, while giving the competitor no incentive to defect as indicated by the poor result of ALLD(d^K). Thus, CTFT represents an ideal candidate for a repeated game strategy, whether it uses realistic forecasting technique

d^K or the idealized forecast d^P as shown in Figure 4.11. Finally, Figure 4.23 shows that PAVLOV(d^K) is able to exploit cooperation such as ALLC(d^K), but is vulnerable against aggression as demonstrated by ALLD(d^K).

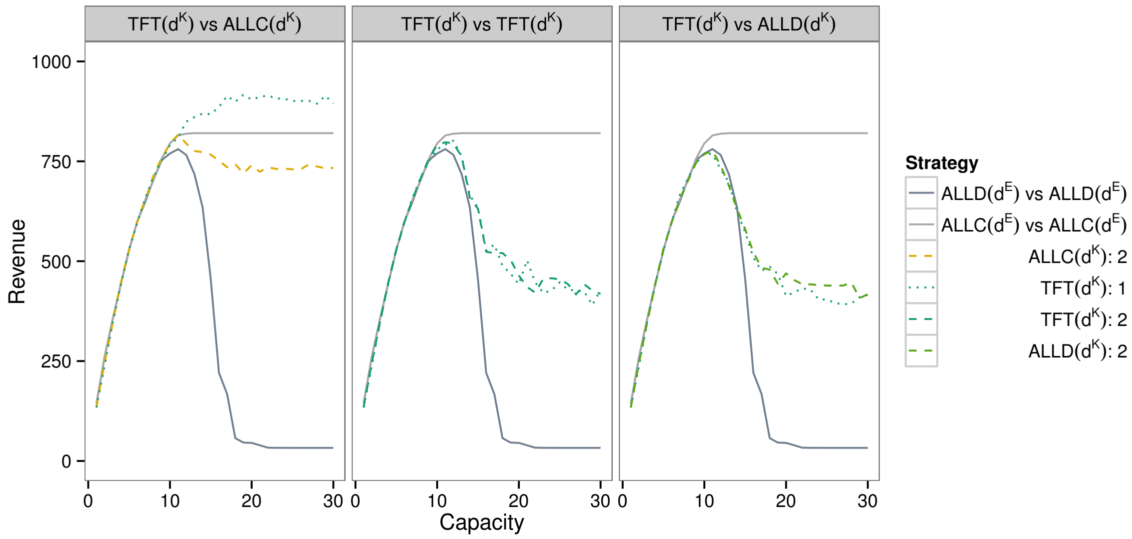

Figure 4.20: TFT using dependent demand forecast d^K with $\varepsilon = 0.1$

Standard Forecast vs. Psychic Forecast

While we have shown in this section that the behavior of standard forecasts can vary from that of the psychic forecast d^P, we want to examine their performance against a competitor with a superior forecast. Therefore, we will compare the standard forecasts to the psychic forecast d^P representing the ideal forecast. For this purpose, we will consider both providers using the non-cooperative strategy ALLD, which arises naturally if a provider does not implement any repeated-game strategy.

Figure 4.24 displays the results of this experiments for the standard forecasts d^I, d^H and d^K. In all three matchups, we find that the better forecast d^P leads to more revenue. Thus, service providers investing in newer and better forecasting methods automatically enhance their competitive performance, even if they do not implement explicit strategies for the repeated game. However, although the ideal forecast d^P always dominates the competitor, there are great differences between the outcomes depending on the standard forecast in use. We find that the independent demand forecast d^I leads to the Spiral Down and thus to a low payoff, although the provider using the standard forecast earns almost as much revenue as the provider using the psychic forecast. The provider using the hybrid demand forecast d^H gets exploited by the competitor, while the dependent

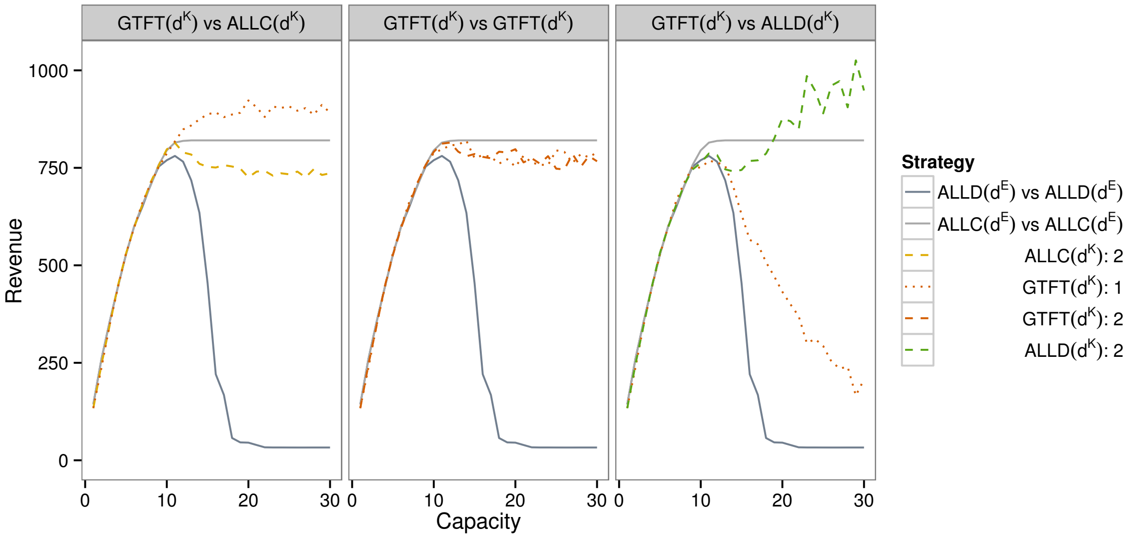

Figure 4.21: GTFT using dependent demand forecast d^K with $\varepsilon = 0.1$

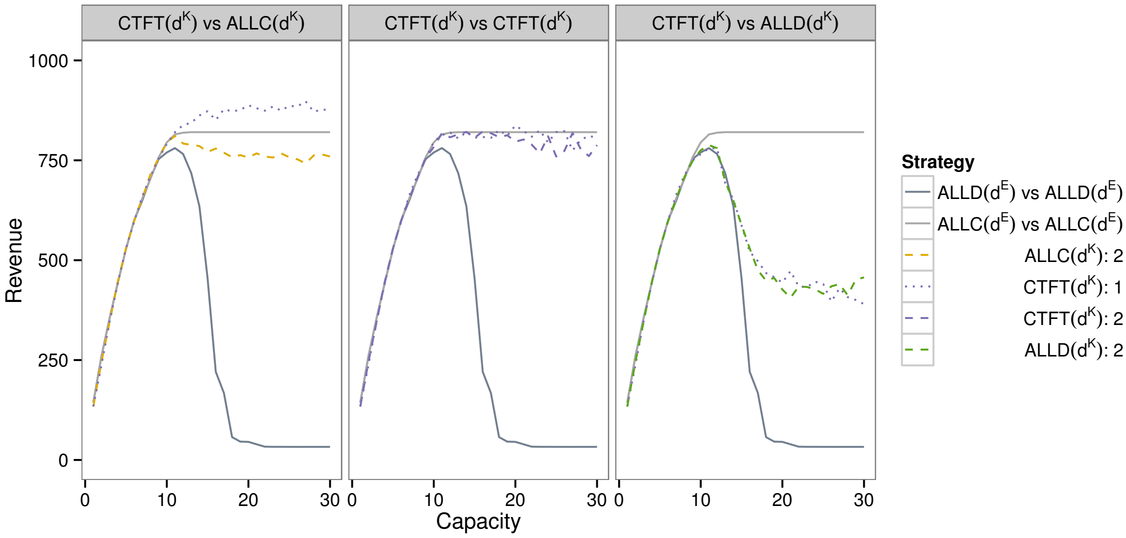

Figure 4.22: CTFT using dependent demand forecast d^K with $\varepsilon = 0.1$

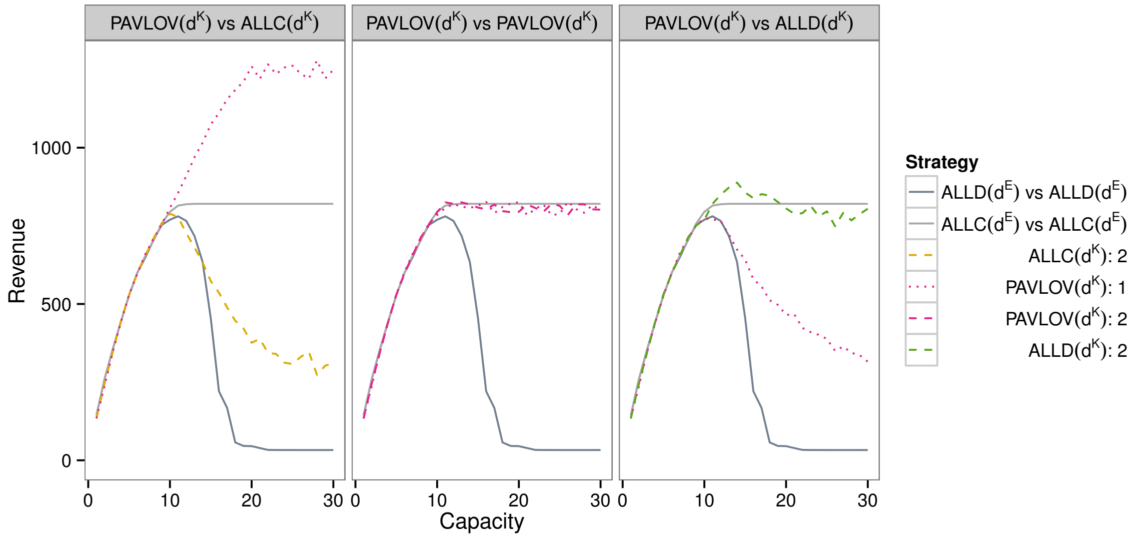

Figure 4.23: PAVLOV using dependent demand forecast d^K with $\varepsilon = 0.1$

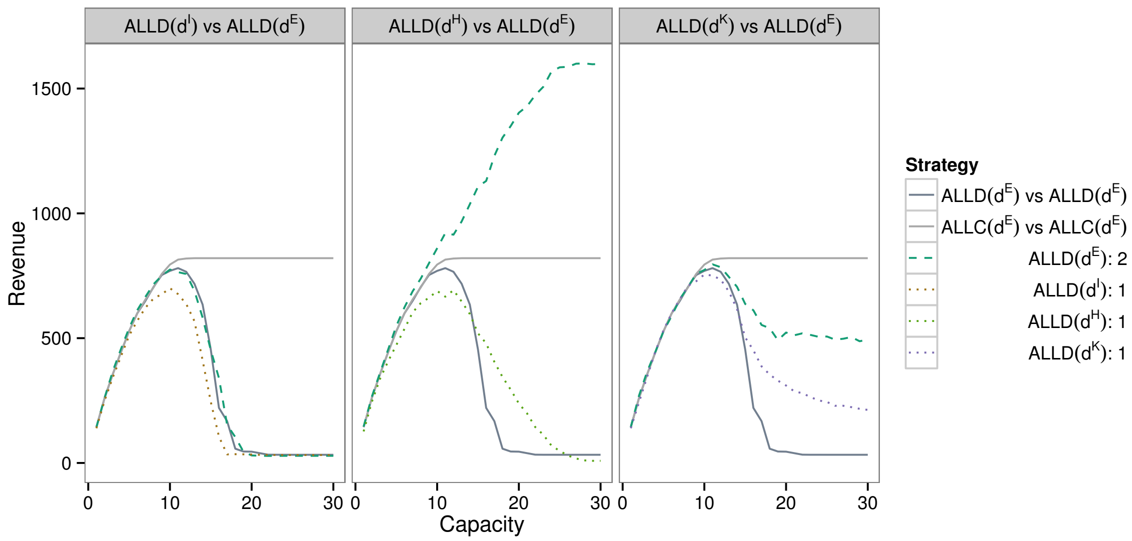

Figure 4.24: Standard forecast vs. psychic forecast

demand forecast d^K comes closer to the result of the psychic forecast. The provider using the dependent demand forecast still gets exploited, but performs reasonably well, considering its natural disadvantage against a psychic forecast.

Irrational Strategies

Similarly to the analysis for the psychic forecast d^P in Section 4.4.1, we will analyze the competitive interactions between a rational provider and an irrational competitor following either MATCH or UNDER. Since the dependent demand forecast d^K has outperformed the other standard forecasting methods in this section, we will concentrate on this forecast. Analogously to above, we will focus on ALLD as the most natural strategy of the repeated game.

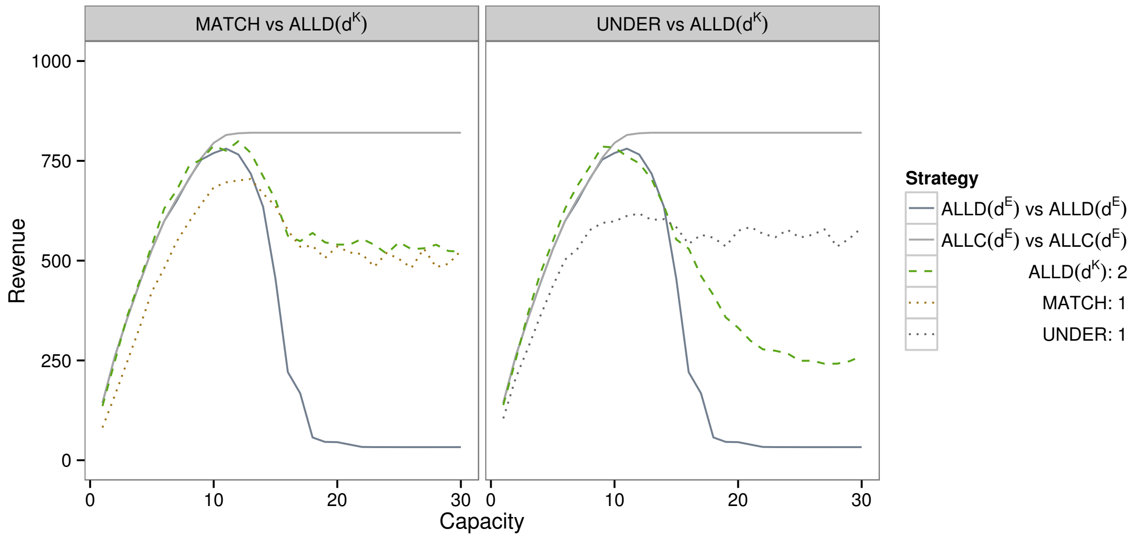

Figure 4.25: Standard forecast vs. irrational strategies with $\varepsilon = 0.1$

Figure 4.25 shows that MATCH leads to a similar result as the symmetrical case, as was to be expected. However, the matching provider receives a slightly lower payoff than his competitor for low capacities, which is caused by the great effect of observation errors for the MATCH strategy. As the result of ALLD(d^K) approximating the optimal solution for low capacities, the ideal response from the competitor would be to mirror this strategy. However, since MATCH has no own forecast to rely on, observation errors can cause an immediate deviation from the optimal strategy.

Against an underpricing provider, we can observe two different outcomes differentiated by the severity of the capacity restriction. As in the case of a matching provider, the rational provider receives a higher share of the payoff as long as capacities are low. In this case, the estimation procedure of the dependent demand forecast works well and leads to a good approximation of the optimal price level. Since UNDER deviates from this price level, the resulting revenue is distinctly lower. However, for higher capacities, the underpricing provider earns more revenue. This is due to the demand model used in estimation, which is independent of the competitor, becoming less and less correct as the capacity restriction vanishes: If the competitor has the capability to serve almost all or all the customers in the market, ignoring competition can become ruinous.

4.5 Summary

In this chapter, we provided a formal model for repeated interactions of two RM service providers and evaluated strategies in the repeated game via simulations.

In our analyses, we focused on research questions 1 – 5 posed in Chapter 3. Thus, we aimed to find a way to exploit the similarities of the IPD and the repeated RM game in order to find successful strategies for the repeated RM game. We wanted to reproduce the strategies' success using standard RM methods available to real-life service providers. Furthermore, we were interested in the strategies' behavior against irrational strategies and in the effect of observation errors.

In this section, we will summarize our approach in this chapter and give an overview of our findings.

First, we gave an introduction and motivation, where we outlined our approach of bridging the gap between game theoretic treatment of simplified models and realistic revenue management simulations. We pointed out the dilemma faced by RM employing service providers mentioned in Chapter 3: In a duopoly, both providers employing a RM strategy that focuses on maximizing only the provider's own revenue over the course of a single sales period leads to the Competitive Spiral Down, significantly reducing both players' revenues. On the other hand, a cooperative price selection avoids this problem, but is vulnerable against a more aggressive price selection of the competitor. In order to handle this dilemma, we stressed the importance of examining revenue management under competition as a repeated game, thus setting the agenda for the remainder of this chapter.

In Section 4.1, we gave an overview of our demand model as well as of the two main components of RM systems, namely a forecasting engine and an optimizer. We included realistic forecasts of varying degrees of complexity as well as an idealized, so-called "psychic", forecast relying on customer data. The idealized forecast is not possible in the real world, but in a simulation environment this forecast allows us to abstract from the imperfections of current forecasting technology. Finally, we described optimization techniques appropriate for each of the forecasts presented and the resulting control strategies.

In Section 4.2, we highlighted the similarities between the IPD and the repeated RM game. We developed a heuristic to transfer strategies from the IPD to the repeated RM game and used this to describe a host of repeated-game strategies for the repeated RM game, adapting the most successful strategies from the IPD. Furthermore, we described simple strategies inspired by the wide-spread RM practice of price matching.

In Section 4.3, we described the state of the art of the simulation environment REMATE, which was used to assess our proposed solutions to the dilemma in the repeated RM game. Since we had access to the source code of REMATE, we could build on this basis and add details to the implementation that enabled us to use it for the evaluation of competitive strategies in the remainder of this chapter.

Section 4.4 was devoted to the evaluation of the strategies presented in Section 4.2. We simulated a simple scenario of a duopoly of service providers with varying capacity restrictions using the simulation environment REMATE described in Section 4.3.

Before discussing simulation results, we outlined the properties a successful strategy in the repeated RM game should possess and how we would test for them in Section 4.2.3. In particular, a successful strategy should solve the dilemma outlined in Chapter 3 and thus be a part of a jointly optimal Nash equilibrium of the repeated RM game. However, since it is not possible to analyze all combinations of possible strategies via simulations, we presented a set of necessary conditions for a solution to the dilemma that we could test in our analyses.

In Section 4.4.1, we focused on simulation results for the psychic forecast d^P. Similarly to Isler and Imhof (2008), we found that the effect of competition depends heavily on the severity of the capacity restriction. For low capacities, competitive effects were almost irrelevant, with cooperative and non-cooperative behavior resulting in the same outcome. However, for high capacities the importance of competition grew significantly, so that pure non-cooperative behavior led to the ruinous Competitive Spiral Down effect. We were able to show that there exist strategies in the repeated RM game that seem to solve the dilemma, as they achieve the jointly optimal solution in a mirror matchup and do not let an aggressor exploit their cooperative stance. However, we demonstrated that the possibility of observation errors can profoundly change the behavior of strategies. Without errors, Tit for Tat represented a solution to the dilemma in the repeated RM game, but as soon as we introduced observation errors into the game, the mirror matchup of two TFT players led to the Competitive Spiral Down instead of the jointly optimal solution. Thus, observation errors have an even greater effect in the repeated RM game than in the IPD. Nevertheless, we showed that robust variations of TFT such as CTFT—and to a lesser degree GTFT—were able to transfer the desired properties of TFT to the case with observation errors. Additionally, we analyzed the PAVLOV strategy, which cannot be part of a Nash equilibrium, but has the possibility to exploit suckers. Finally, we analyzed the irrational strategies MATCH and UNDER based on price-matching. As expected, they were completely dependent on the competitor, on whose price they relied. Thus, these strategies could lead to good results, but had little control over their performance. We found that the strategies that had proven most successful in our analysis so far also performed best against the irrational strategies MATCH and UNDER: Each of the robust strategies GTFT, CTFT and PAVLOV fared well against irrational competition, whereas TFT suffered from its lack of robustness. When accounting for the possibility of the competitor following a simpler, non-robust strategy, we found GTFT to be the best strategy, which presented a change from the results of the symmetrical mirror matchups. Thus, choosing the right strategy may well depend on the environment.

Section 4.4.2 was devoted to Isler and Imhof (2008)'s hypothesis that standard forecasts should behave similarly to the psychic forecast d^P under competition. However, Cooper et al. (2009) showed that this is not necessarily true for arbitrary demand forecasts. In our analysis, we found great differences between the results of the standard forecasts presented in Section 4.1.2. As expected, using an oversimplified demand model like the independent demand forecast d^I is not sufficient to reproduce the results of the psychic forecast d^P. Instead, this led to the Spiral Down, caused by an inappropriate demand model, with even worse consequences than the Competitive Spiral Down. We showed that the widely-used hybrid demand model d^H prevents the Spiral Down, but is vulnerable against aggressors. For high capacities, the estimation procedure failed to react to the growing influence of the competitor's prices and proceeded to compute the same price elasticities as for low capacities. Thus, in a symmetric matchup the providers did not approximate the Nash equilibrium. Although this helped avoid the Competitive Spiral Down, the fact that the system did not reach the competitive best response can be dangerous against more aggressive competitors or simply competitors with a better estimation technique. Finally, we examined the forecast d^K, a prototype of a forecast based on a dependent demand model that has not seen extensive use in the industry yet. This forecast came closest to emulating results of idealized forecast d^P. Using the forecast d^K, we found that simulation results for our repeated game strategies turned out similar to the same strategies using the idealized forecast d^P. Thus, in order to render Isler and Imhof's hypothesis true and be able to use the results of Section 4.4.1, it seems necessary to use a high quality forecast with a sufficiently complex demand model. We also investigated the performance of standard forecasts against the psychic forecast d^P. As expected, the psychic forecast outperformed the standard forecasts, with a similar ranking of standard forecasts as before, emphasizing the importance of a high-quality forecast.

5 The Repeated RM Game as a Markov Process

In this chapter, we will build on the insights gained in the previous Chapter 4 to study a simplification of the repeated Revenue Management (RM) game to a Markov process. In contrast to the previous chapter, this variation allows a mathematical analysis, reducing our reliance on the use of simulations in this chapter.

The competition between two service providers leads to a high degree of complexity that hinders a thorough analysis. Even for the single resource case, an analytical treatment of the RM game is hard because of the interdependence of each player's states at different time steps and stages. In the previous Chapter 4, we used simulation as a tool to gain insight into the RM game despite its inherent complexity. Although the simulation approach is suited to analyze competitive interactions in complex scenarios, Bartke et al. (2013) showed that a high degree of complexity can complicate the analysis of the results and even introduce chaotic behavior. Consequently, we simplified the scenario to both firms using only a single resource in order to enhance the clarity of our results.

However, as stressed in the literature review in Section 2.4, computational results gained through simulation should always be accompanied by mathematical analysis. Ideally, simulation and mathematical analysis should complement each other. While simulation enables us to examine the problem in a more realistic setting, the results do not hold the same clarity as an analytic formula. On the other hand, mathematical analysis has the prospect of delivering provable relationships between scenario parameters. Unfortunately, the mathematical analysis of the repeated RM game in the form presented in Chapter 4 is not feasible.

Similarly to the computational results in Section 4.4, we will ignore network effects and focus on the single-resource case. In the single-resource RM game, each state is dependent on other states via two mechanisms: The learning mechanism during forecasting causes the state of the game during any stage s and time step t to depend on all previous stages at the same time step, whereas the optimization ties together all time steps t during one stage s via the capacity constraints. In this chapter, we will study a simplification of the original problem that will enable an analytical treatment of the RM game.

Throughout this chapter we will use the psychic forecast d^P. As we have argued in Chapter 4, this forecast represents an optimal case, towards which real life forecasts are striving. As simulation results in Section 4.4 showed, the Competitive Spiral Down is most extreme without capacity constraints. In this case—achieved when each service provider has sufficient capacity to accommodate the whole demand—the non-cooperative strategy ALLD(d^P) leads to the Bertrand Nash equilibrium. In this chapter, we will focus on this extreme point, where competitive effects are at their strongest. We will therefore drop the capacity constraint, enabling us to analyze a single time step t in isolation. The resulting game during each time step t can be described independently of all other time steps as an infinite-order chain, i.e. a process depending on all previous stages s. Unfortunately, infinite-order chains are not as well-researched as Markov chains. However, as discussed in Section 2.2.2, at least in the Iterated Prisoner's Dilemma (IPD), the restriction to single-stage memory strategies does not imply a loss of generality. Therefore, we use only the most recent observation to create d^P, which can be accomplished by setting the exponential smoothing parameter $\alpha = 1$.

With the information provided by the psychic forecast d^P, it is irrational to price higher than $f(c)$, rendering products with a higher price than the cooperative threshold product superfluous. To save some notation, we ignore products i with $i < c$ and assume that the players' joint payoff is maximized by the most expensive product, i.e. $c = 1$.

In summary, we will use the following assumptions to describe the RM game during each time step t by a Markov chain of order 1 over the stages s:

1. Each provider sells a single resource.
2. Both players use the psychic forecast d^P.
3. There is no capacity restriction.
4. Strategies are restricted to a single-stage memory.
5. The most expensive product is jointly optimal, i.e. $c = 1$.

The Markov version of the RM game is a generalization of the IPD. At each stage, instead of having only two options, each player has to decide which of his n products should be the lowest available. Since the transfer of ideas from the IPD to the RM game proved successful in Chapter 4, the even closer relationship between the RM game and the Markov version presented in this chapter leads us to believe that findings for the Markov game should be transferable even more easily. The great advantage of modeling the game as a Markov chain lies in the simplicity of the Markov concept. Markov processes in discrete time are thoroughly-studied stochastic processes with well-known properties (for an overview see Section 5.1.1).

Research on the IPD provides an array of thoroughly-studied simple strategies. However, a transfer of these strategies to the repeated RM game is not trivial, since even in the IPD, the construction of the game's Markov process for arbitrary strategies is cumbersome. In the study of the IPD, researchers have resorted to using transition matrices without providing a derivation, which led to a lack of insight into the structure of the transition probabilities. In the RM game, where the choice of an arbitrary amount of prices or even a continuous price range, observation errors and the concept of reputation further complicate the construction of the game as a Markov chain, an examination of the structure of the Markov transition kernels will be necessary.

As outlined in the literature review in Section 2.2.2, the formulation as a Markov chain helps examine the long-term behavior of the game, since for a general class of Markov chains, the stationary measure exists and is unique. This measure describes the process's long-term behavior, which has been used extensively in the analysis of the IPD. However, a similar analysis of stationary measures of the RM game has to build on a possibility to construct transition kernels of the RM game, which has not been done in the literature so far.

While the stationary measure describes the outcome of the game as long as nothing changes, researchers have tried to reproduce the players' pursuit of the optimal strategy, emulating behavior that can be found in environments of rational learning players. The basis of such an evolutionary simulation are the stationary measures of every combination of strategies, which are used to compute the game's long-term payoffs. Building on the formulation of the RM game as a Markov process, a similar approach may be useful to implement temporal dynamics into the players' strategies. However, even for the IPD, the strategy space can grow large, so that a thorough statistical analysis of the simulations can become difficult.

As mentioned in Section 2.2.2, Press and Dyson (2012) recently discovered a class of strategies in the IPD that can exploit evolutionary players. These so-called Extortionate Strategies allow a player to unilaterally enforce a linear relation between his and his opponent's payoffs. Following such a strategy against an evolutionary competitor enables the player to have the competitor drag him to the optimal payoff, all the while gaining a disproportionately high payoff. In Press and Dyson's paper and—to the best of our knowledge—the following research, it is always assumed that players can observe their competitor's moves perfectly. However, in the literature review in Section 2.2.2, we pointed out the influence of flawed observations in the IPD. As stated in Chapter 3, we will demonstrate the importance of observation errors in the RM game similarly to the IPD. Additionally, we will analyze the effect of observation errors on the existence of Extortionate Strategies.

In this chapter, we will first provide the mathematical background needed in the remainder of the chapter in Section 5.1. In Section 5.2, we will examine the structure of the Markov chain and derive the chain's transition matrices. In order to establish the link to the Iterated Prisoner's Dilemma, we will describe and analyze the payoff structure of the RM game in Section 5.3. In Section 5.4, we will put the considerations of the previous sections into practice by defining strategies and investigating their long-term behavior. In particular, we will first use the framework described in Section 5.2 to adapt the repeated-game strategies presented in Chapter 4 in section 5.4.1, then analyze the long-term behavior of the Markov chain for pairs of strategies in Section 5.4.2. In Section 5.4.3, we will examine the effect of errors in the determination of the jointly optimal product c used for cooperation. In order to help the reader get a better understanding of Section 5.4's results, we will give an extensive example in Section 5.4.4. Building on the insights from the previous sections, we will study an evolutionary game by means of simulation in Section 5.5, where we will introduce additional temporal dynamics into the game's strategies. Subsequently, we will use Section 5.6 to discuss the existence of Extortionate Strategies, which can exploit evolutionary behavior. As an excursus, we will apply our methodology to the case of continuous prices in Section 5.7. Finally, in Section 5.8 we will give a conclusion of this chapter, where we summarize and discuss our findings.

5.1 Mathematical Basics

In this section, we will give an overview of the mathematical basics needed in this chapter. We will focus on stochastic processes, which enable the analysis of the evolution of the RM game, and tensor products, which help describe competitive interactions in a linear way. A profound introduction into the field of probability theory including stochastic processes is given by Kallenberg (2002), while Pinsky and Karlin (2010) give a more application-oriented review of stochastic processes with a focus on Markov processes. Tensor products are introduced both mathematically rigorously and intuitively by Hungerford (1974), while a host of applications specifically of the Kronecker product is reviewed in by Loan (2000). Throughout this section, we will use examples to provide a connection between the mathematical concepts and the application in competitive revenue management.

5.1.1 Stochastic Processes

In the mathematical analysis of stochastic processes, we want to measure the probability of events. Given the set of all possible outcomes, we can characterize any event as a subset of a set J.

Example 1. *In a RM context, consider a single service provider with a set of products J who offers a subset of products characterized by the cheapest product. Then the subset $\{j\}$ with $j \in J$ corresponds to the event that the provider offers j as the cheapest product. More generally, the event A with $A \subset J$ means that the provider offers any $j \in A$ as the cheapest product.*

Unfortunately, Vitali (1905, for n=1) and Comfort and Gordon (1961, for arbitrary n) showed that it is impossible to apply the natural measure of length, area, volume and so on to all subsets of $\mathbb{R}^n$. Later, Banach and Tarski (1924) demonstrated the existence of so-called non-measurable sets in $J = \mathbb{R}^3$, for which no measure at all can be defined that is invariant to rotations and translations (for a more recent review see e.g. Stromberg, 1979). Therefore, we have to restrict ourselves to a set of subsets for which a probability measure can be assigned. For this purpose, we define a σ-Algebra $\mathcal{J}$ on a set J as a set of subsets $A \subset J$ fulfilling

$$\exists A \subseteq J : A \in \mathcal{J} \tag{5.1.1}$$

$$A \in \mathcal{J} \implies J \setminus A = \{j \in J \mid j \notin A\} \in \mathcal{J} \tag{5.1.2}$$

$$A_i \in \mathcal{J} \quad \forall i \in \mathbb{N}^+ \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{J} \tag{5.1.3}$$

Thus, $\mathcal{J}$ is a non-empty set of subsets of J that is closed under complementation and under countable union.

The definition of a σ-Algebra allows for a lot of different σ-Algebras for any but the most trivial sets J. For any J, the trivial σ-Algebra $\{\varnothing, J\}$ is the smallest possible σ-Algebra, while the power set $\mathcal{P}(J)$, denoting the set of all subsets of J, is always the biggest. If J is finite, it is possible to use the power set as σ-Algebra, since in this case any subset of J can be assigned a measure in a meaningful way. However, for uncountably infinite sets such as $\mathbb{R}^+$, we are limited in the choice of measures on the power set, due to the existence of non-measurable sets. Thus, for $J = \mathbb{R}^+$ we have to use a smaller σ-Algebra containing all the relevant subsets—but no non-measurable sets—instead of the power set. The natural choice is the Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^+)$, which is the smallest σ-Algebra that contains all the intervals of $\mathbb{R}^+$. This σ-Algebra will prove useful for the study of the RM game with continuous prices in Section 5.7.

The tuple $(J, \mathcal{J})$ of a set J equipped with σ-Algebra $\mathcal{J}$ is called a measurable space. As the name suggests, we can define a probability measure on a measurable space $(J, \mathcal{J})$ as a map $\mu : \mathcal{J} \to [0, 1]$, so that $\mu(\varnothing) = 0$, $\mu(J) = 1$ and for every countable collection of pairwise disjoint sets A_i, we have $\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i)$.

This enables us to define a probability space $(J, \mathcal{J}, \mu)$ as a set J equipped with a σ-Algebra $\mathcal{J}$ and a probability measure μ, or equivalently as a measurable space $(J, \mathcal{J})$ with a probability measure μ. When it is clear which probability measure is meant, the probability measure is often denoted by $\mathbb{P}$.

Of the many possible probability measures, a particularly simple example used in this dissertation is the Dirac measure. The Dirac measure δ_x is a probability measure that can be defined on any measurable space $(J, \mathcal{J})$ by

$$\forall A \in \mathcal{J} : \delta_x(A) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ else} \end{cases} . \tag{5.1.4}$$

On a discrete space, any probability measure can be defined as a convex combination of Dirac measures. This is no longer true for spaces based on $\mathbb{R}$.

Example 2. *For a single service provider offering a finite product set* $J = \{1, \ldots, n\}$, *we can use the* σ*-Algebra* $\mathcal{J} = \mathcal{P}(J)$. *Then, any probability measure takes the form* $\mu(A) = \sum_{j=1}^{n} c_j \delta_j(A)$ *for* $A \in \mathcal{J}$, *where the constants* c_j *are non-negative and sum to one,* $\sum_{j=1}^{n} c_j = 1$. *Here, the constants* c_j *represent the probability that the provider chooses product* j.

Example 3. *For a continuous state space* $J = \mathbb{R}^+$ *and* $\mathcal{J} = \mathcal{B}(J)$, *as we consider in Section 5.7, the Dirac measure is useful to choose a single specific element of the uncountably infinite product set, e.g. in order to set an exact price.*

With the help of measurable spaces and probability spaces, we can define random variables. Given two measurable spaces $(G, \mathcal{G})$ and $(J, \mathcal{J})$, we say a function $X : G \to J$ is measurable, if its preimage is included in the domain's σ-Algebra, i.e. $X^{-1}(A) \in \mathcal{G}$ for all $A \in \mathcal{J}$. If the domain of the function is even a probability space $(G, \mathcal{G}, \mu)$, a measurable function is called a random variable. The probability of any measurable outcome $A \in \mathcal{J}$ of the random variable X can be found as $\mathbb{P}(X^{-1}(A))$.

A stochastic process $(X_s)_{s \in S}$ is a sequence of random variables X_s, so that for every time index s, X_s is a measurable function $X_s : (G, \mathcal{G}, \mu) \to (J, \mathcal{J})$ from a probability space to a measurable space. Although this definition holds for general time sets S, we will restrict this outline to the discrete case $S = \mathbb{N}_0$, since we will not need the more general case in the following of this thesis.

In order to analyze a stochastic process, it is necessary to determine which parts of the process's history can influence its future behavior. A filtration $(\mathcal{G}_s)_{s \in S}$ is an increasing sequence of σ-Algebras $\mathcal{G}_s$ so that $s_1 \leq s_2 \implies \mathcal{G}_{s_1} \subseteq \mathcal{G}_{s_2}$. Each σ-Algebra $\mathcal{G}_s$ in a filtration can be thought of as the information available up to time s. Usually, the filtration is chosen so that each X_s is measurable with respect to $\mathcal{G}_s$, i.e. in order to calculate the probability of any event at a point in time s, it is sufficient to know the information available up to s without having to know the future. While there are many possible filtration candidates to choose when modeling a stochastic process, the natural choice is the natural filtration $\sigma(X_t, t \leq s) = \sigma\{X_t^{-1}(A) : A \in \mathcal{J}, t \leq s\}$ of process X_s generated by all values of X up to stage s.

A Markov process is a stochastic process that is memoryless in the sense that the future evolution of a process depends only on the present state of the process and not on its past. This so-called Markov property greatly facilitates the analysis of this class of stochastic processes. In this thesis, we will call a Markov process a Markov chain, if not only time S but also the state space J is discrete.

For a Markov chain, it is sufficient to formulate the Markov property for all elements of the state space as

$$\mathbb{P}(X_{s+1} = j_{s+1} \mid X_t = j_t, t \leq s) = \mathbb{P}(X_{s+1} = j_{s+1} \mid X_s = j_s), \tag{5.1.5}$$

where $\mathbb{P}(A \mid B)$ denotes the conditional probability of $A \in \mathcal{J}$ given $B \in \mathcal{J}$ and $j_s, j_{s+1}, j_t \in J$. Such a simple definition is not possible for continuous state spaces. Instead, we can use the natural filtration $\sigma(X_t, t \leq s)$ to generalize the Markov property 5.1.5 to

$$\mathbb{P}(X_{s+\tau} \in A \mid \sigma(X_t, t \leq s)) = \mathbb{P}(X_{s+\tau} \in A \mid \sigma(X_s)) \quad \forall A \in \mathcal{J}, \tau > 0 \tag{5.1.6}$$

for a Markov process on a general state space.

Example 4. *Consider a single service provider offering a finite set of products $J = \{1, \dots, n\}$ repeatedly over the stages $s \in \mathbb{N}^+$. At each stage s, the provider chooses a cheapest product depending on his choices in the previous stages. Therefore, at each stage s, the provider's cheapest product is a random variable with values in the measurable space $(J, \mathcal{P}(J))$. The collection of these random variables is a stochastic process that is also a Markov process if and only if the provider relies only on his action in the previous stage $s-1$ to choose a product in stage s.*

The evolution of the state of the Markov process is governed by its transition probabilities. We will focus on the time-homogeneous case, in which the transition probability between states of the Markov process is independent of the stage s. Therefore, we will omit the time-dependence in the notation of transition probabilities.

In the discrete case with finite $J = \{j_1, \dots, j_n\}$ and $\mathcal{J} = \mathcal{P}(J)$, it is sufficient and convenient to use a transition matrix M with entries $p_{i,l} = \mathbb{P}(X_s = j_l \mid X_{s-1} = j_i)$ to describe the Markov chain's transition probabilities. This matrix operates on the space of probability vectors $V = \{(v_1, \dots, v_n) \in [0,1]^n \mid \sum_{i=1}^{n} v_i = 1\}$ so that the evolution of the Markov process probability vector follows $\mu_{s+1} = \mu_s M$, where $\mu_s, \mu_{s+1} \in V$ and $\mu_s M$ denotes the matrix product. Note that in the discrete case, a probability vector is equivalent to a probability measure. If a measure μ satisfies $\mu M = \mu$, we call μ a stationary measure.

In the more general case of a possibly infinite state space J, it is impossible to represent the transition probabilities in matrix form. Instead, the more general concept of a transition kernel is used. A transition kernel K is a map on the measurable space $(J, \mathcal{J})$, for which

- $K(x, \cdot)$ is a probability measure for all $x \in J$ and
- $K(\cdot, A)$ is $\mathcal{J}$-measurable for all $A \in \mathcal{J}$.

We can find the probability measure μ_{s+1}, describing the state of the Markov process at time $s+1$, with the help of the transition kernel K and μ_s as $\mu_{s+1}(A) = (K\mu_s)(A) = \int_J K(x, A)\mu_s(dx)$ for any $A \in \mathcal{J}$. In this case of infinite state spaces, we call a measure μ a stationary measure, if $K\mu = \mu$. The case of infinite state spaces will be important in our analysis of continuous prices in Section 5.7.

Example 5. *For a single provider offering a finite product set with only two products $J = \{1, 2\}$ based on his previous price, the transition matrix takes the form*

$$M = \begin{pmatrix} p_{1,1} & 1 - p_{1,1} \\ p_{2,1} & 1 - p_{2,1} \end{pmatrix}. \tag{5.1.7}$$

In the pursuit of the stationary measure, it is helpful to analyze the elements of the state space. In the following, we will concentrate on the discrete case, since we will not need the general case in this thesis. This allows for a more simple notation. A state $j_1 \in J$ is accessible from $j_2 \in J$, if $\mathbb{P}(X_{s+m} = j_1 \mid X_s = j_2) > 0$ for $m \geq 0$. Using this concept, the state space J can be partitioned into communicating classes C. We say two states $j_1, j_2 \in J$ belong to the same communicating class, if they are accessible from each other. A communicating class C is called closed, if it cannot be escaped, i.e. for all $j_1 \in C, j_2 \in J$ we have

$$\mathbb{P}(X_{s+m_1} = j_2 \mid X_s = j_1) > 0 \implies j_2 \in C. \tag{5.1.8}$$

A Markov chain with a single closed communicating class is called irreducible. As the stationary measure of the Markov chain is concentrated on its closed communicating classes, we can only expect a unique stationary measure for irreducible Markov processes. A special case of communicating classes is created by so-called absorbing states. A state $j \in J$ is called absorbing, if $\mathbb{P}(X_{s+1} = j \mid X_s = j) = 1$. An absorbing state j is always the only member of its own closed communicating class $\{j\}$. Furthermore, we call a state $j_i \in J$ unreachable, if $\mathbb{P}(X_{s+1} = j_i \mid X_s = j_l) = 0$ for all $j_l \in J$. In other words, unreachable states are not accessible from any state in the state space.

Example 6. *Using the transition matrix from Example 5 with the values $p_{1,1} = 0.5$ and $p_{2,1} = 0$ we get*

$$M = \begin{pmatrix} 0.5 & 0.5 \\ 0 & 1 \end{pmatrix}. \tag{5.1.9}$$

Since the state 1 is not accessible from state 2, we find two separate communicating classes $\{1\}$ and $\{2\}$. However, only the class $\{2\}$ is closed and in fact consists of the absorbing state 2. Here, the stationary measure—represented as a vector—is $\pi = (0, 1)$.

We can combine probability spaces to help describe stochastic processes on a larger state space. We call the natural combination of probability spaces $(J_1, \mathcal{J}_1, \mu_1)$ and $(J_2, \mathcal{J}_2, \mu_2)$

the product space $(J_1 \times J_2, \mathcal{J}_1 \otimes \mathcal{J}_2, \mu_1 \otimes \mu_2)$. Here, the tensor product σ-Algebra $\mathcal{J}_1 \otimes \mathcal{J}_2$ denotes the smallest σ-Algebra that contains all subsets $\{A_1 \times A_2 \mid A_1 \in \mathcal{J}_1, A_2 \in \mathcal{J}_2\}$ of the cartesian product $J_1 \times J_2$, whereas the product measure $\mu_1 \otimes \mu_2$ is the unique measure that satisfies $\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ for all $A_k \in \mathcal{J}_k$. The original measures μ_k are called marginal measures and can be reconstructed with the help of the product measure via integration $\mu_1(A) = \int_{J_2} \mu_1 \otimes \mu_2(A \times dj_2)$. In the following Section 5.1.2, we will examine tensor products in more detail.

Example 7. *Consider competition between two service providers S_1 and S_2, each offering a set of products J_k, $k = 1, 2$. As long as each player S_k acts independently of the competitor, we can find a probability space $(J_k, \mathcal{J}_k, \mu_k)$ that completely describes and evaluates S_k's actions. The measurable space $(J_1 \times J_2, \mathcal{J}_1 \otimes \mathcal{J}_2)$ contains all possible events that can result from combinations of both players' actions, while the product measure $\mu_1 \otimes \mu_2$ yields the probability of both players' actions if both providers act independently from each other.*

5.1.2 Tensor Products

In many natural contexts, a researcher can encounter relationships between entities that depend linearly on more than a single entity. As an example, this can happen in game theory when the outcome of the game depends linearly on both players' strategies. If the entities can be embedded in vector spaces, these relationships may be expressed as multilinear maps from many vector spaces to a single vector space. However, the study of multilinear and even bilinear maps on vector spaces is by far not as simple and well developed as the study of linear maps. Tensor product offers a possibility to transform bilinear—and, by extension, multilinear—maps to linear maps. Although the concept exists in more general contexts, we only cover tensor products on real vector spaces in this section, since this is sufficient for this thesis.

We define a tensor product of two real vector spaces as a real vector space that enables the use of linear maps instead of bilinear maps. This is called the universal property of the tensor product. More precisely, given two real vector spaces V_1 and V_2 we define a tensor product of V_1 and V_2 to be a pair (T, t), where T is a real vector space and $t : V_1 \times V_2 \to T$ is a bilinear map with the following property: Given any bilinear map $Q : V_1 \times V_2 \to W$, there exists a unique linear map $L : T \to W$ such that $Q = L \circ t$, where $\circ$ is the composition of maps. Writing $T = V_1 \otimes V_2$ and $t(v_1, v_2) = v_1 \otimes v_2$ for $v_1 \in V_1, v_2 \in V_2$, this means that we have a unique linear map L with $Q(v_1, v_2) = L(v_1 \otimes v_2)$.

For finite-dimensional real vector spaces, the construction of the tensor product is simple. Given two real vector spaces V_1 with basis $\{e_1, \ldots, e_n\}$ and V_2 with basis $\{f_1, \ldots, f_m\}$, a basis of the tensor product $V_1 \otimes V_2$ can be found as

$$\{e_1 \otimes f_1, \ldots, e_1 \otimes f_m, e_2 \otimes f_1, \ldots, e_2 \otimes f_m, \ldots e_n \otimes f_1, \ldots, e_n \otimes f_m\}. \tag{5.1.10}$$

Thus, in this choice of basis every basis vector $v_1 \otimes v_2$ of $V_1 \otimes V_2$ is associated with a pair (v_1, v_2), where v_i is a basis vector of V_i.

The tensor product also operates on linear maps on vector spaces. Given linear maps $\phi_1 : V_1 \to W_1$ and $\phi_2 : V_2 \to W_2$, the tensor product $\phi_1 \otimes \phi_2$ is defined as

$$\phi_1 \otimes \phi_2 : V_1 \otimes V_2 \to W_1 \otimes W_2 \tag{5.1.11}$$
$$v_1 \otimes v_2 \mapsto \phi_1(v_1) \otimes \phi_2(v_2). \tag{5.1.12}$$

Example 8. *Consider repeated competitive interactions between two service providers. Each provider offers a finite product set $J_1 = J_2 = \{1, \dots, n\}$ based on his own action in the previous stage independently of the competitor. The transition probabilities $\phi_1 : V_1 \to V_1$ and $\phi_2 : V_2 \to V_2$ for each player's state operate linearly on the space of probability vectors $V_k = \{(v_1^k, \dots, v_n^k) \mid \sum_{j=1}^{n} v_j^k = 1\}$ associated with provider S_k. Then, the transition probabilities for the combination of both players' actions are given by the tensor product $\phi_1 \otimes \phi_2$.*

In order to employ matrix notation for the tensor product $\phi_1 \otimes \phi_2$ of the linear maps $\phi_1 : V_1 \to W_1$ and $\phi_2 : V_2 \to W_2$, we need to choose a basis of the tensor products $V_1 \otimes V_2$ and $W_1 \otimes W_2$. If we use the standard choice of basis 5.1.10 of the tensor product of vector spaces, the so-called Kronecker product uses matrix representations of the linear maps ϕ_1 and ϕ_2 to yield a matrix representation of the tensor product $\phi_1 \otimes \phi_2$. The Kronecker product represents an efficient and simple way of calculating the tensor product of linear maps. Consequently, in this thesis we will always choose the standard basis of the tensor product of vector spaces, so that we can employ the Kronecker product. Let V_1, V_2, W_1, W_2 be finite-dimensional vector spaces with bases

$$b(V_1) = \{e_1, \dots, e_n\} \tag{5.1.13}$$
$$b(V_2) = \{f_1, \dots, f_q\} \tag{5.1.14}$$
$$b(W_1) = \{g_1, \dots, g_m\} \tag{5.1.15}$$
$$b(W_2) = \{h_1, \dots, h_p\}, \tag{5.1.16}$$

$V_1 \otimes V_2$ and $W_1 \otimes W_2$ tensor product vector spaces with bases

$$b(V_1 \otimes V_2) = \{e_1 \otimes f_1, \dots, e_1 \otimes f_q, e_2 \otimes f_1, \dots, e_2 \otimes f_q, \dots, e_n \otimes f_1, \dots, e_n \otimes f_q\} \tag{5.1.17}$$
$$b(W_1 \otimes W_2) = \{g_1 \otimes h_1, \dots, g_1 \otimes h_p, g_2 \otimes h_1, \dots, g_1 \otimes h_p, \dots, g_m \otimes h_1, \dots, g_m \otimes h_p\}, \tag{5.1.18}$$

and $\phi_1 : V_1 \to W_1$ and $\phi_2 : V_2 \to W_2$ two linear maps. Given the chosen bases, $A = (a_{ij})$ denotes the $m \times n$ matrix representation of ϕ_1, while B is the $p \times q$ matrix representation of ϕ_2. Then, the matrix representation $A \otimes B$ of $\phi_1 \otimes \phi_2$ given the chosen bases can be computed as the Kronecker product

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}. \tag{5.1.19}$$

Example 9. *In a repeated competitive interaction between S_1 and S_2, each player offering the product set $J_1 = J_2 = \{1, 2\}$, providers choose their product based on their own previous action, independently of each other. The transition matrices of each player can be written as*

$$M_1 = \begin{pmatrix} p_1 & 1-p_1 \\ p_2 & 1-p_2 \end{pmatrix}, M_2 = \begin{pmatrix} q_1 & 1-q_1 \\ q_2 & 1-q_2 \end{pmatrix}. \tag{5.1.20}$$

The Kronecker product

$$M_1 \otimes M_2 = \begin{pmatrix} p_1\,q_1 & p_1\,(1-q_1) & (1-p_1)\,q_1 & (1-p_1)\,(1-q_1) \\ p_1\,q_2 & p_1\,(1-q_2) & (1-p_1)\,q_2 & (1-p_1)\,(1-q_2) \\ p_2\,q_1 & p_2\,(1-q_1) & (1-p_2)\,q_1 & (1-p_2)\,(1-q_1) \\ p_2\,q_2 & p_2\,(1-q_2) & (1-p_2)\,q_2 & (1-p_2)\,(1-q_2) \end{pmatrix} \tag{5.1.21}$$

represents the transition matrix on the product space $J_1 \times J_2$. This matrix gives the transition probabilities between all combinations of single-player states if S_1 follows M_1 and S_2 follows M_2 independently of each other.

As mentioned in Section 5.1.1, the notion of tensor products has also been applied to probability spaces, where $\mathcal{J}_1 \otimes \mathcal{J}_2$ denotes the tensor product σ-Algebra of σ-Algebras $\mathcal{J}_1$ and $\mathcal{J}_2$, and $\mu_1 \otimes \mu_2$ denotes the product measure of measures μ_1 and μ_2.

Among the many useful features of dealing with linear maps on vector spaces, the adjugate matrix is particularly interesting to us. The adjugate or classical adjoint of a $n \times n$ matrix A is

$$\mathrm{adj}(A)_{i,j} = (-1)^{i+j} \det(m(A, j, i)), \tag{5.1.22}$$

where $m(A, i, j)$ is the matrix produced by removing the i-th row and the j-th column of A. The adjugate matrix has the property that

$$A\,\mathrm{adj}(A) = \det(A) I_n, \tag{5.1.23}$$

with I_n denoting the $n \times n$ identity matrix. This property can be proved using Cramer's rule (see e.g. Hungerford, 1974) and will be exploited in Section 5.6.

5.2 Construction of the Markov Chain of the RM game

For research of the endearingly simple IPD, the analysis of the game's Markov chain for a given pair of strategies was feasible—although cumbersome—without thorough derivation of the transition matrices (see e.g. Boerlijst et al., 1997b; Molander, 1985; Nowak & Sigmund, 1990). However, due to the growing dimensions of the strategy space, such a naive procedure quickly becomes impractical for more complex games such as the RM game.

In this section, we will examine the structure of the RM game as a Markov chain. We will build on the description of the matrices describing the constituent dynamics of the game such as each player's strategy, observation errors and the evolution of players' reputation to derive the form of the game's transition matrices.

Since we have not yet specified the payoff structure of the game, the analysis in this section applies to a more general class of games than just the RM game. More specifically, the construction described in this section can be carried out for any 2-player symmetrical Markov game with an n-dimensional strategy space.

5.2.1 Prerequisites

As described in Section 4.1.2, any availability situation of a service provider S_k is entirely characterized by its lowest available product $j_k^{min} \in J_k$, where $J_k = \{1, \ldots, n\}$ denotes the set of products of S_k.

By abuse of notation, we write for the elements of $(J_1 \cup \{0\}) \times (J_2 \cup \{0\})$

$$(i, j) = (j_1^{min} = i \wedge j_2^{min} = j), \tag{5.2.1}$$

where $j_k^{min} = 0$ denotes the state that no product is available from S_k.

Since we have removed the capacity constraints, service providers cannot run out of resources and we can ignore the possibility $j_k^{min} = 0$. Consequently, we can use $J_1 \times J_2$ as the state space of the Markov chain $(X_s)_{s \in \mathbb{N}_0}$ describing the RM game. Since the state space is finite, we can use the σ-Algebra $\mathcal{P}(J_1 \times J_2)$, where $\mathcal{P}(J_1 \times J_2)$ is the power set of $J_1 \times J_2$. For the same reason, we can express the probability measure on the state space as a probability vector. The space of possible probability vectors of the Markov chain is the tensor product $V_1 \otimes V_2$, where

$$V_k = \left\{ (v_k^1, \ldots, v_k^n) \in [0,1]^n \,\middle|\, \sum_i v_k^i = 1 \right\}, \quad k = 1, 2, \tag{5.2.2}$$

denotes the space of probability vectors describing the probability measure of player S_k's lowest available product with $\mathbb{P}(j_k^{min} = j) = v_k^j$. Note that the lower index indicates the service provider and the upper index the product. The probability distribution of the Markov process X_s evolves via a transition matrix operating on $V_1 \otimes V_2$.

The tensor product guarantees that the probability $\mathbb{P}((i,j)) = v_1^i v_2^j$ of state (i,j) corresponds to the basis vector $e_i \otimes e_j$ in $V_1 \otimes V_2$, where e_i denotes the i-th standard basis vector of $\mathbb{R}^n \supset V_k$, $k = 1, 2$. However, the order of the basis elements $e_i \otimes e_j$ is important for the form of the matrix representation of linear maps on the state's probabilities. In the following, we will use

$$\{e_1 \otimes e_1, \dots, e_1 \otimes e_n, e_2 \otimes e_1, \dots, e_2 \otimes e_n, \dots, e_n \otimes e_1, \dots, e_n \otimes e_n\} \tag{5.2.3}$$

as the basis of the underlying vector space $\mathbb{R}^n \otimes \mathbb{R}^n \supset V_1 \otimes V_2$. This choice of basis allows us to use the Kronecker product to represent the tensor product of linear maps on V_1 and V_2 in matrix form.

In the RM game, each service provider S_k follows a strategy for the repeated game, which we will denote by σ_k. Without observation errors, the interaction of these strategies yields the transition matrix

$$M_{\sigma_1,\sigma_2} : V_1 \otimes V_2 \to V_1 \otimes V_2. \tag{5.2.4}$$

5.2.2 General Markov Strategies

Any Markov strategy σ_k of player S_k can be expressed as

$$\overline{M}_{\sigma_k} : V_1 \otimes V_2 \to V_k. \tag{5.2.5}$$

We call strategies represented in this form general Makov strategies.

The strategy matrix can differ for different players playing the same strategy due to our choice of basis. But the matrix for player S_2 can be deduced from the strategy matrix for player S_1. Let $\overline{M}_{\sigma_1}$ be the strategy matrix for player S_1. Then, the matrix for player S_2 following the same strategy $\sigma_1 = \sigma_2$ is $B \cdot \overline{M}_{\sigma_1}$, where B is the matrix of the change of basis

$$B : V_1 \times V_2 \to V_1 \times V_2 \tag{5.2.6}$$

$$e_i \otimes e_j \mapsto e_j \otimes e_i. \tag{5.2.7}$$

The matrix B accounts for the fact that the roles of own and competitor prices are reversed for player S_2 compared to player S_1.

The transition matrix M_{σ_1,σ_2} depends on both players' strategies, where both players choose their prices simultaneously based on the same input:

$$M_{\sigma_1,\sigma_2} : V_1 \otimes V_2 \to V_1 \otimes V_2 \tag{5.2.8}$$

$$(v_1 \otimes v_2)^t \mapsto \left((v_1 \otimes v_2)^t \cdot \overline{M}_{\sigma_1}\right) \otimes \left((v_1 \otimes v_2)^t \cdot B \cdot \overline{M}_{\sigma_2}\right) \tag{5.2.9}$$

With the forking matrix

$$Q : V_1 \otimes V_2 \to V_1 \otimes V_2 \otimes V_1 \otimes V_2 \quad (5.2.10)$$

$$e_i \otimes e_j \mapsto e_i \otimes e_j \otimes e_i \otimes e_j \quad (5.2.11)$$

that provides both players with the same input, we can write M_{σ_1,σ_2} as a combination of matrix products and tensor products:

$$M_{\sigma_1,\sigma_2} = Q \cdot \left(\overline{M}_{\sigma_1} \otimes (B \cdot \overline{M}_{\sigma_2})\right) \quad (5.2.12)$$

5.2.3 Reactive Strategies

If players choose their next action independently of their previous action, we can construct the transition matrix in an easier way. Following common notation, we call these strategies reactive strategies. Reactive strategies are an important subset of possible strategies, so that in some previous studies of the IPD, researchers have restricted themselves to the analysis of reactive strategies (e.g. Nowak & Sigmund, 1990, 1992).

Each player S_k reacts to its competitor S_l's previously offered lowest product j_l^i by choosing one lowest product j_k^m of his own for the next stage, which is found by the map M_{σ_k}:

$$M_{\sigma_k} : V_l \to V_k \quad (5.2.13)$$

With the help of the map

$$B_k : V_l \to V_k \quad (5.2.14)$$

$$e_i \mapsto e_i, \quad (5.2.15)$$

which accounts for the fact that each player needs its competitor's prices as input, we can write each player's strategy as

$$M_{\sigma_k} = \widetilde{M}_{\sigma_k} \circ B_k \quad (5.2.16)$$

with the linear map

$$\widetilde{M}_{\sigma_k} : V_k \to V_k. \quad (5.2.17)$$

Given our canonical choice of basis for V_1 and V_2, M_{σ_k} and $\widetilde{M}_{\sigma_k}$ have the same matrix representation.

The Markov process of this game takes the form of a Markov chain, in which its single-player strategies provide the components used in the construction of the game's transition matrix. In fact, the transition matrix can be computed as a tensor product of the single-player strategies

$$\begin{aligned} M_{\sigma_1,\sigma_2} &= M_{\sigma_1} \otimes M_{\sigma_2} && (5.2.18)\\ &= \left(\widetilde{M}_{\sigma_1} \circ B_1\right) \otimes \left(\widetilde{M}_{\sigma_2} \circ B_2\right) && (5.2.19)\\ &= \left(\widetilde{M}_{\sigma_1} \otimes \widetilde{M}_{\sigma_2}\right) \circ B, && (5.2.20) \end{aligned}$$

where $\otimes$ denotes the tensor product and B is the matrix of the change of basis

$$\begin{aligned} B : V_1 \times V_2 &\to V_1 \times V_2 && (5.2.21)\\ e_i \otimes e_j &\mapsto e_j \otimes e_i. && (5.2.22) \end{aligned}$$

Since each $\widetilde{M}_{\sigma_k}$ is a linear map from V_k to V_k and due to our choice of basis for $V_1 \otimes V_2$, we can use the Kronecker product to compute the tensor product in 5.2.20.

Ignoring the possibility of observation errors, the distribution vector thus evolves in the following way:

$$\begin{aligned} M_{\sigma_1,\sigma_2} : V_1 \times V_2 &\to V_1 \times V_2 && (5.2.23)\\ (v_1 \otimes v_2)^t &\mapsto (v_1 \otimes v_2)^t \cdot B \cdot \left(\widetilde{M}_{\sigma_1} \otimes \widetilde{M}_{\sigma_2}\right). && (5.2.24) \end{aligned}$$

In case we need to test a reactive strategy with another strategy that depends on its player's previous action, we have to describe both strategies in their general form. Due to our choice of basis, we can find the form of the general Markov strategy for the first player as

$$\overline{M}_{\sigma_1} = 1_n \otimes \widetilde{M}_{\sigma_1}, \tag{5.2.25}$$

where $1_n = (1, \ldots, 1)^t$ is the $n \times 1$ matrix with only ones as elements and $\otimes$ denotes the Kronecker product. As mentioned in Section 5.2.2, the form of the matrix for the second player can differ. It can be constructed with the help of the basis permutation B as $B \cdot \overline{M}_{\sigma_1}$.

5.2.4 Observation Errors

As explained in Section 4.2, we want to account for the possibility of observation errors. The lowest available price of the observed availability situation price is shifted lower or higher from the real availability situation, each with probability ε.

$$\begin{aligned} E : V_k &\to V_k && (5.2.26)\\ e_1 &\mapsto e_1(1-\varepsilon) + e_2\varepsilon && (5.2.27)\\ e_i &\mapsto e_i(1-2\varepsilon) + e_{i-1}\varepsilon + e_{i+1}\varepsilon \quad 1 < i < n && (5.2.28)\\ e_n &\mapsto e_n(1-\varepsilon) + e_{n-1}\varepsilon && (5.2.29) \end{aligned}$$

In summary, we can construct the transition matrix of the Markov process of two competing service providers in the presence of errors using each player's strategy and the error transition matrix E:

$$M^{\varepsilon}_{\sigma_1,\sigma_2} = (E \otimes E) \cdot M_{\sigma_1,\sigma_2} \tag{5.2.30}$$

where $\otimes$ denotes the Kronecker product. First, each service provider's observations are subject to errors. Then, these observations are combined and used as basis for each player's strategic decision via M_{σ_k}.

5.2.5 Reputation

Some strategies require the introduction of a reputation as described in Section 4.2.1. Each player S_k is assigned a reputation of $r_k \in R_k = \{g, b\}$, where $r_k = g$ represents a good and $r_k = b$ a bad reputation. For every stage $s \in \mathbb{N}_0$, the Markov process X_s of the RM game is now a $(J_1 \times J_2 \times R_1 \times R_2)$-valued random variable. We denote the vector of the probability distribution of r_k by $\rho_k \in P_k$, where $P_k = \{(\rho_k^1, \rho_k^2) \in [0,1]^2 \mid \rho_k^1 + \rho_k^2 = 1\}$. With the basis $\{e_g, e_b\}$ of $\mathbb{R}^2 \supset [0,1]^2$, where e_g corresponds to good and e_b to bad reputation, a probability vector in P_k can be written as $(\rho_k^1, \rho_k^2) = \rho_k^1 e_g + \rho_k^2 e_b$.

Combining this with the model without reputation, we find that the space of probability vectors of the Markov chain with reputation is $V_1 \otimes V_2 \otimes P_1 \otimes P_2$. The evolution of the distribution of reputation is defined by the map R for $p, r \in \{g, b\}$:

$$R : V_1 \otimes V_2 \otimes P_1 \otimes P_2 \to V_1 \otimes V_2 \otimes P_1 \otimes P_2 \tag{5.2.31}$$

$$e_i \otimes e_j \otimes e_p \otimes e_r \mapsto \begin{cases} e_i \otimes e_j \otimes e_g \otimes e_g & \text{if } i,j = 1 \\ & \text{ or } i = 1, j > 1, p = b, r = g \\ & \text{ or } i > 1, j = 1, p = g, r = b \\ e_i \otimes e_j \otimes e_g \otimes e_b & \text{if } i = 1, j > 1, p = g \\ & \text{ or } i = 1, j > 1, p = b, r = b \\ & \text{ or } i,j > 1, p = g, r = b \\ e_i \otimes e_j \otimes e_b \otimes e_g & \text{if } i > 1, j = 1, r = g \\ & \text{ or } i > 1, j = 1, p = b, r = b \\ & \text{ or } i,j > 1, p = b, r = g \\ e_i \otimes e_j \otimes e_b \otimes e_b & \text{if } i,j > 1, p = r \end{cases} \tag{5.2.32}$$

Recall that in this chapter, $c = 1$ denotes the product maximizing the players' combined payoffs, while i corresponds to the cheapest product of S_1 and j to the cheapest product of S_2. Note that since the players' products are sorted in descending order by their price f, $i < j$ is equivalent to $f(i) > f(j)$. The indices p and r refer to the reputation of S_1 and S_2 respectively and can take values of g for a good or b for a bad reputation.

Thus, a player will gain a bad reputation for defecting without provocation, i.e. against a competitor with a good reputation. Players can restore their good reputation by cooperating, and in all other cases, the players' reputations remain unchanged.

The introduction of reputation leads to general Markov strategies $\overline{M}_{\sigma_k}$ that may depend on the probability of all states in the previous stage:

$$\overline{M}_{\sigma_k} : V_1 \otimes V_2 \otimes P_1 \otimes P_2 \to V_k \tag{5.2.33}$$

With these general single-player strategies, we can construct the transition matrix as

$$\begin{aligned} M_{\sigma_1,\sigma_2} : V_1 \otimes V_2 \otimes P_1 \otimes P_2 &\to V_1 \otimes V_2 \otimes P_1 \otimes P_2 && (5.2.34)\\ x^t &\mapsto x^t \cdot Q \cdot \left(\overline{M}_{\sigma_1} \otimes \left((B \otimes A) \cdot \overline{M}_{\sigma_2}\right) \otimes \pi_P\right) && (5.2.35)\end{aligned}$$

with the forking matrix

$$\begin{aligned} Q : V_1 \otimes V_2 \otimes P_1 \otimes P_2 &\to (V_1 \otimes V_2 \otimes P_1 \otimes P_2)^{\otimes 3} && (5.2.36)\\ x &\mapsto x \otimes x \otimes x, && (5.2.37)\end{aligned}$$

the projection

$$\begin{aligned} \pi_P : V_1 \otimes V_2 \otimes P_1 \otimes P_2 &\to P_1 \otimes P_2 && (5.2.38)\\ e_i \otimes e_j \otimes e_p \otimes e_r &\mapsto e_p \otimes e_r && (5.2.39)\end{aligned}$$

and the permutation matrices

$$\begin{aligned} B : V_1 \otimes V_2 &\to V_1 \otimes V_2 && (5.2.40)\\ e_i \otimes e_j &\mapsto e_j \otimes e_i && (5.2.41)\end{aligned}$$

and

$$\begin{aligned} \widetilde{B} : P_1 \otimes P_2 &\to P_1 \otimes P_2 && (5.2.42)\\ e_p \otimes e_r &\mapsto e_r \otimes e_p. && (5.2.43)\end{aligned}$$

By abuse of notation, we keep writing $\overline{M}$ for general Markov strategies and Q for the matrix that provides the other maps with the input.

Of course, for reactive strategies we can construct the transition matrix simply as

$$M_{\sigma_1,\sigma_2} = \left(B \otimes \widetilde{B}\right) \cdot (M_1 \otimes M_2 \otimes I_4), \tag{5.2.44}$$

where I_n is the n-dimensional identity matrix. However, since in this case neither strategy depends on reputation, we will save some notation and ignore reputation.

At every stage s, the reputation will be updated based on the previous actions including observation errors. Then, each player sets prices according to his strategy and the updated reputation:

$$M^r_{\sigma_1,\sigma_2} = R \cdot M_{\sigma_1,\sigma_2} \tag{5.2.45}$$

$$M^{\varepsilon,r}_{\sigma_1,\sigma_2} = (E \otimes E \otimes I_4) \cdot R \cdot M_{\sigma_1,\sigma_2} \tag{5.2.46}$$

Here, the upper index r indicates the consideration of reputation in the transition matrix.

5.3 Payoff Structure in the Markov RM game

As mentioned in Section 4.2 of the previous chapter, demand and prices of the RM game are setup like in the IPD, which results in a similar dilemma for the players: Each service provider has an incentive to underprice his competitor, since there is always a price level that leads to higher revenues than sharing earnings with the competitor at the higher price level. However, if both players keep underpricing each other, they will end up with the unprofitable Competitive Spiral Down effect.

In this section, we will describe the payoff structure of the Markov RM game that allows us to maintain these properties. We will show that the dilemma occurs naturally for a host of demand models, as long as the prices are set up in a myopic way. In particular, the pricing structure needs to provide prices that make narrowly underpricing the competitor a profitable option, at least for a single stage.

In the prisoner's dilemma, the payoffs are characterized by two inequalities:

$$T > R > P > S \tag{5.3.1}$$

$$2R > T + S \tag{5.3.2}$$

As Boerlijst et al. (1997a) put it, "R stands for the *reward* for mutual cooperation, P is the *penalty* for mutual defection, T is the *temptation* payoff for unilaterally defecting and S the *sucker* payoff for being exploited."

For $k = 1, 2$ we introduce the maps $Y_k : J_1 \times J_2 \to \mathbb{R}^+$ describing the revenue made by service provider S_k. In our symmetric setting, we have $Y_1(i, j) = Y_2(j, i)$. In order to have the RM game represent a generalization of the IPD, we have to adapt Inequalities 5.3.1 and 5.3.2 to the state space $J_1 \times J_2$.

5.3.1 Assumptions

For a generalization of $T > R$, we have to keep in mind that the RM game offers each player many possibilities to undercut its opponent due to multiple price points. Assuming that the price levels are chosen appropriately for the demand model, narrowly undercutting is always profitable. However, this cannot be guaranteed for undercutting by multiple classes, since a player may end up dropping prices so far that the advantage of attracting the demand from his competitor is outweighed by the disadvantage of earning far less per customer. It is therefore only necessary that a player is always better off being the only player undercutting than having both players drop their prices.

$$Y_1(i+1, i) > Y_1(i, i) \tag{5.3.3}$$

$$Y_1(i+m, i) > Y_1(i+m, i+m) \tag{5.3.4}$$

The first Inequality 5.3.3 shows that it is optimal for each service provider to marginally undercut its competitor, thus we have $T > R$ for undercutting by a single class. Unilaterally undercutting by multiple classes as in Inequality 5.3.4 is still better than mutually choosing the low price, therefore we always have $T > P$. However, it is not guaranteed that the payoff is higher than the payoff for mutual cooperation.

In the RM game without capacity constraints, the cooperative threshold product c denotes the product that maximizes the joint revenue of both players. Since in this chapter, the most expensive product is jointly optimal, we have $Y_1(c, c) = Y_1(1, 1) > Y_1(i, i)$ for any $i > 1$. As a regularity assumption on the prices and the demand model, we assume that there exists no other local maximum of both players' joint revenue.

$$Y_1(i, i) > Y_1(i+m, i+m) \quad \forall i, m \geq 1 \tag{5.3.5}$$

Consequently, as soon as the price of the offered products falls below the price of $c = 1$, both players undercutting results in lower revenue for both players, $R > P$.

Intuitively, there is not much reason for any customer in a sensible demand model to buy a product at a higher price than the lowest offered, which leads to next to no expected purchases.

$$Y_1(i, i+m) < Y_1(n, n) \tag{5.3.6}$$

Inequality 5.3.6 refers to the sucker payoff, which in our symmetric setup is always lower than complete mutual defection, thus $P > S$.

Finally, the last inequality in the prisoner's dilemma's payoff needs to be fulfilled for every class with a lower price than the cooperative threshold product $c = 1$.

$$2Y_1(i, i) > Y_1(i, i+m) + Y_1(i+m, i) \quad \forall i \geq 1 \tag{5.3.7}$$

Inequality 5.3.7 creates the dilemma that the joint revenue of mutual cooperation is higher than that of unilateral defection, although unilateral defection is optimal for each player due to Inequality 5.3.3.

The Inequalities 5.3.3 – 5.3.7 guarantee the connection of the RM game to the IPD. For $n = 2, c = 1$, these inequalities can be simplified to the familiar Inequalities 5.3.1 and 5.3.2 of the prisoner's dilemma.

5.3.2 Payoff Calculation

The payoff s_k for each player S_k can be calculated as the scalar product of the stationary distribution and a reward vector y_k:

$$s_k = \pi^t \cdot y_k \tag{5.3.8}$$

$$y_1 = \sum_{i=1}^{n}\sum_{j=1}^{n} Y_1(i,j) e_i \otimes e_j \tag{5.3.9}$$

$$y_2 = \sum_{i=1}^{n}\sum_{j=1}^{n} Y_2(i,j) e_i \otimes e_j \tag{5.3.10}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} Y_1(j,i) e_i \otimes e_j \tag{5.3.11}$$

$$= B \cdot y_1 \tag{5.3.12}$$

where B is the change of basis

$$B : V_1 \times V_2 \to V_1 \times V_2 \tag{5.3.13}$$

$$e_i \otimes e_j \mapsto e_j \otimes e_i \tag{5.3.14}$$

introduced in Section 5.2.

5.3.3 Examples

In this section, we will examine a variety of demand models used frequently in the RM literature (Talluri & van Ryzin, 2004b, pp.301–332). We will find that for many demand models, the assumptions 5.3.3 – 5.3.7 applying to the revenue can be satisfied by appropriately chosen prices.

In the following demand models, we will assume that the demand D_k for S_k's products is concentrated on the lowest product of each player at any stage s, so that we have $D_k(j, j_k^{min}, j_l^{min}) = 0$ for $j \neq j_k^{min}$. Thus, in order to shorten notation, we can concentrate on the lowest products and write $D_1(i,j) = D_1(i, j_1^{min} = i, j_2^{min} = j)$ and $D_2(i,j) =$

$D_2(j, j_2^{min} = j, j_1^{min} = i)$. This notation allows the formulation as a map $D_k : J_1 \times J_2 \to \mathbb{R}^+$, which demonstrates the connection to the revenue function Y_k. In fact, the revenue for player S_k for a product constellation $(j_k^{min} = i, j_l^{min} = j)$ can be calculated as

$$Y_k(i,j) = D_k(i,j) f(i). \tag{5.3.15}$$

Without loss of generality, we will take the point of view of the first player, similarly to Section 5.3.1.

Fixed valuation A simple model to describe customer behavior is a fixed valuation model (see e.g. Martínez-de Albéniz & Talluri, 2011). Customers can buy any class j with a lower or equal price than an arbitrary class $\tilde{c}$ and no class with a higher price. Obviously, $\tilde{c}$ is the jointly optimal product in this model. Without loss of generality, we can ignore the more expensive products and assume that $\tilde{c} = c = 1$. In this model, customers always buy the cheapest product in the market, and if both players offer the same price, each gets half the demand. The demand function can be written as

$$D_1(i,j) = \begin{cases} D_0 & \text{if } i \geq c, i > j \\ \frac{1}{2} D_0 & \text{if } i \geq c, i = j \\ 0 & \text{else} \end{cases} \tag{5.3.16}$$

with a constant D_0. In this model, the Inequalities 5.3.3 – 5.3.7 are fulfilled if and only if $2f(i+1) > f(i) > f(i+1)$.

Uniformly distributed valuation A uniform distribution of customers' valuation leads to a linear demand model. Keeping the assumption that customers buy the cheapest class, and choose randomly between competitors if both players offer the same price, we find the demand function

$$D_1(i,j) = \begin{cases} D_0 \left(1 - \frac{f(i)}{f(n)}\right) & \text{if } i > j \\ \frac{1}{2} D_0 \left(1 - \frac{f(i)}{f(n)}\right) & \text{if } i = j \\ 0 & \text{else} \end{cases} \tag{5.3.17}$$

with a constant parameter D_0. More generally, a linear form of demand leads to

$$D_1(i,j) = \begin{cases} D_0 - \psi f(i) & \text{if } i > j \\ \frac{1}{2}(D_0 - \psi f(i)) & \text{if } i = j \\ 0 & \text{else} \end{cases} \tag{5.3.18}$$

with constant parameters D_0 and ψ. Note that choosing $\psi = \frac{D_0}{f(n)}$ leads to the model characterized by Equation 5.3.17. Similarly to the fixed valuation model, we find that a game using this model satisfies Inequalities 5.3.3 – 5.3.7 if and only if the prices are chosen in the range $2f(i+1) > f(i) > f(i+1)$.

Linear demand Alternatively, the demand can be assumed to be linear in both players' prices (see e.g. Ledvina & Sircar, 2011), so that the demand function takes the form

$$D_1(i, j) = D_0 - \psi(f(i) - f(j)). \tag{5.3.19}$$

Inequalities 5.3.3 – 5.3.7 are fulfilled for every choice of prices, as long as $\psi > \frac{D_0}{f(n)}$.

Utility maximization Isler and Imhof (2008) used a demand model with exponential valuation. Each customer tries to maximize his utility, where a customer's total utility consists of the base utility U_0 denoting the customer's willingness-to-pay, the utility U associated with player S_1 and the utility $\bar{U}$ associated with S_2. These utilities are assumed to be distributed independently and according to the exponential distributions

$$U_0 \sim \beta \exp(-\beta x), \tag{5.3.20}$$
$$U \sim \alpha \exp(-\alpha x), \tag{5.3.21}$$
$$\bar{U} \sim \bar{\alpha} \exp(-\bar{\alpha} x). \tag{5.3.22}$$

The demand for a product is the result of a choice process, in which the customer weighs up the alternatives. If the first player offers the lowest product i and the second player offers j, a customer purchases product i if it does not exceed his valuation

$$U_0 - U - f(i) > 0 \tag{5.3.23}$$

and it is better than the competitor's product

$$U_0 - U - f(i) < U_0 - \bar{U} - f(j). \tag{5.3.24}$$

For their simulation, Isler and Imhof used the parameters $\beta = 0.01$ and $\alpha = \bar{\alpha} = 0.5$ and the prices displayed in Table 5.1.

If we ignore the products with a higher price than the jointly optimal product $j = 5$ with $f(j) = 100$, we can verify that this demand model with Isler and Imhof's parameters and prices fulfills Inequalities 5.3.3 – 5.3.7.

j	1	2	3	4	5	6	7	8	9	10	11	12	13
f(j)	300	240	180	140	100	65	40	28	20	14	10	7	5

Table 5.1: Prices used by Isler and Imhof (2008)

Normally distributed valuation In the demand model used in Section 4.4 to simulate the RM game with capacity constraints, customers always buy the cheapest class in the market. Similarly to previous examples, demand is evenly split if both players offer the same products at the same price. Customers' willingness-to-pay is determined by a normal distribution with mean $\mu \approx 100$ and standard deviation $\sigma \approx 40$. We find that Inequalities 5.3.3 – 5.3.7 define a range of possible prices. In particular, the set of prices used in Section 4.4 and reproduced in Table 5.2 fulfills Inequalities 5.3.3 – 5.3.7, as long as we ignore the products above the jointly optimal product $j = 4$ with $f(j) = 87$.

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
f(j)	260	147	109	87	71	58	47	38	30	23	17	12	8	5	3	2

Table 5.2: Prices used in the simulation in Section 4.4

5.4 Limiting Behavior of Repeated-Game-Strategies

Throughout this section, each service provider will use the psychic forecast d^P, so that the pair of strategies of both players playing DEFECT(d^P) converge to the non-cooperative Nash equilibrium strategies of the single-stage game.

Without capacity constraints and dependence solely on the previous stage, DEFECT reduces to undercutting the competitor's previous availability situation j by making the cheaper product $j + 1$ available. COOP consists of playing DEFECT, without offering any product cheaper than the jointly optimal product c. Since we assume $c = 1$ in this chapter, COOP reduces to choosing exactly the jointly optimal product.

5.4.1 Strategies

In Section 4.2, we have introduced strategies for the multi-stage RM game that served as a basis for our further analysis of competitive interactions between service providers in Chapter 4. Many of these strategies were adapted from the IPD, which can be seen as a special case of the Markov version of the RM game formulated in Chapter 5. In this section, we will show how the strategies used in Chapter 4 can be formulated in the framework presented in Section 5.2.

ALLD The strategy ALLD is simply a repetition of the non-cooperative single-stage strategy DEFECT. Since we have removed the capacity constraints in this chapter, the symmetric case of two ALLD playing competitors leads to the Bertrand Nash equilibrium, in which both players always choose the lowest price. As ALLD depends only on the competitor's previous move, it can be formulated as a reactive strategy with a single-player strategy matrix:

$$M_{ALLD} : V_l \to V_k \tag{5.4.1}$$
$$e_i \mapsto e_{\min(i+1,n)} \tag{5.4.2}$$

ALLC Similarly, ALLC is a repetition of the cooperative single-stage strategy COOP, which will lead to tacit collusion in the symmetric case. This strategy is independent of any player's actions. However, we present it here in reactive strategy form, since this is the most compact form of describing strategies in our framework:

$$M_{ALLC} : V_l \to V_k \tag{5.4.3}$$
$$e_i \mapsto e_1 \tag{5.4.4}$$

TFT Tit for Tat (TFT) is a reactive strategy that has performed extraordinarily well both in the IPD as well as in the RM game of Chapter 4, as long as there were no observation errors. A TFT player will play DEFECT, if the competitor priced lower than the level of tacit collusion given by the cooperative threshold product c. Otherwise, he will play COOP.

$$M_{TFT} : V_l \to V_k \tag{5.4.5}$$
$$e_i \mapsto \begin{cases} e_1 & \text{if } i \leq c \\ e_{\min(i+1,n)} & \text{else} \end{cases} \tag{5.4.6}$$

GTFT A more robust version of TFT is Generous Tit for Tat (GTFT), which is also a reactive strategy with the single-player strategy matrix

$$M_{GTFT} : V_l \to V_k \tag{5.4.7}$$
$$e_i \mapsto \begin{cases} e_1 & \text{if } i \leq c \\ e_{\min(i+1,n)}(1-\gamma) + e_c\gamma & \text{else} \end{cases}. \tag{5.4.8}$$

Here, γ is the generosity probability, i.e. the probability with which a GTFT-playing service provider will play COOP, although its competitor has played DEFECT in the previous stage. This is the only deviation from the TFT strategy.

CTFT Another robust modification of TFT is Contrite Tit for Tat (CTFT). Since it uses reputation, CTFT is not a reactive strategy, and has to be modeled in the general form for Markov strategies on the lifted state space $J_1 \times J_2 \times R_1 \times R_2$. For player S_1, the strategy is given by

$$\overline{M}_{CTFT} : V_1 \otimes V_2 \otimes P_1 \otimes P_2 \to V_k \tag{5.4.9}$$

$$e_i \otimes e_j \otimes e_p \otimes e_r \mapsto \begin{cases} e_1 & \text{if } p = b \vee r = g \\ e_{\min(j+1,n)} & \text{if } p = g \wedge r = b \end{cases}. \tag{5.4.10}$$

Thus, CTFT plays DEFECT, if its player has a good reputation and its competitor a bad reputation. Otherwise it cooperates.

PAVLOV In the IPD, PAVLOV is an example of a win-stay, lose-shift strategy, that works fundamentally different from TFT and its modifications. Instead of trying to punish aggressors and cooperate with the rest, PAVLOV will try to avoid the two least profitable situations: exploitation at the hands of the competitor and mutual defection. After being exploited, PAVLOV will switch to playing DEFECT, and after a round of mutual defection, PAVLOV will switch to playing COOP. Otherwise, PAVLOV will repeat playing COOP or DEFECT respectively. This mechanism cannot be formulated as a reactive strategy, since it depends on both players' previous actions. Instead, we need the general Markov strategy form on $J_1 \times J_2$ with the single-player strategy matrix

$$\overline{M}_{PAVLOV} : V_1 \otimes V_2 \to V_k \tag{5.4.11}$$

$$e_i \otimes e_j \mapsto \begin{cases} e_1 & \text{if } i, j = 1 \\ & \text{ or } i, j > 1 \\ e_{\min(j+1,n)} & \text{else} \end{cases}. \tag{5.4.12}$$

MATCH Simple price matching can be formulated as the reactive strategy MATCH with the single-player strategy matrix

$$M_{MATCH} : V_l \to V_k \tag{5.4.13}$$

$$e_i \mapsto e_i. \tag{5.4.14}$$

UNDER Similarly to MATCH, underpricing is a reactive strategy called UNDER with the single-player strategy matrix

$$M_{UNDER} : V_l \to V_k \tag{5.4.15}$$

$$e_i \mapsto e_{\min(i+1,n)}. \tag{5.4.16}$$

Note that since we have removed the capacity constraints of the game and use the psychic demand forecast d^P in this chapter, UNDER is the same as the strategy ALLD. Recall that these strategies are different in the more general case of Chapter 4, where the single-stage non-cooperative best response does not need to be the same as underpricing the competitor.

5.4.2 Stationary Measures

In this section, we will examine the long-term behavior of the Markov chain constructed in Section 5.2 using strategies from Section 5.4.1. This is similar to Section 4.4.1, where we investigated the long-term behavior of the RM game with capacity constraints with the help of simulations. Instead of using simulations, we will analyze the stationary measures of the Markov chains for different combinations of strategies. Since there are too many combinations to present here in a compact form, we will focus on notable combinations, similarly to Section 4.4.1. In particular, we will examine the symmetric matchup of two players using the same strategy. This will help us to determine how close strategies are to the joint optimum as well as their robustness against erroneous observations. Additionally, the behavior of PAVLOV will require our special attention.

ALLC vs. ALLC In the matchup of two unconditional cooperators, the cooperating state $(c, c) = (1, 1)$ is absorbing and accessible from all other states $(i, j) \in J_1 \times J_2$. Since ALLC does not take the competitor's previous actions into account, this is true independently of the possibility of observation errors. Thus, the unique stationary measure is the Dirac measure $\pi = \delta_{\{(1,1)\}}$, representing almost sure tacit collusion.

ALLC vs. ALLD For a cooperative ALLC player against an aggressive ALLD player, the state $(c, c+1) = (1, 2)$ is absorbing and accessible from all other states in the state space, independently of observation errors. This results in the unique stationary measure $\pi = \delta_{\{(1,2)\}}$. In the state $(1, 2)$, the ALLD narrowly underprices the ALLC player, which leads to an exploitation of the cooperative player.

ALLD vs. ALLD For two defectors that repeat playing the non-cooperative Nash equilibrium of the single stage, the state of the lowest prices (n, n) is absorbing and accessible from all other states $(i, j) \in J_1 \times J_2$. Since this is true with and without observation errors, the unique stationary measure in both cases is $\pi = \delta_{\{(n,n)\}}$. In this case, the stationary measure demonstrates the effect of the Competitive Spiral Down, where both players end up with the unprofitable repetition of the single-stage Bertrand Nash equilibrium.

TFT vs. TFT Without errors, the Markov chain of the symmetric case TFT vs. TFT possesses three closed communicating classes in $\{(1,1)\}$, $\{(1,n),(n,1)\}$ and $\{(n,n)\}$. Therefore, the stationary distribution depends on the starting distribution. Since there are no errors, neither player will deviate from mutual cooperation or mutual defection. Therefore, the state of mutual cooperation $(1,1)$ is accessible only from itself, while (n,n) is accessible only from all states of mutual defection $\{(i,j) \mid i,j > 1\}$. The third closed communicating class $\{(1,n),(n,1)\}$ is accessible from the remaining states $\{(1,j) \mid j > 1\} \cup \{(i,1) \mid i > 1\}$. Given a starting distribution of mutual cooperation that typically characterizes TFT, the appropriate stationary measure is $\pi = \delta_{\{(1,1)\}}$.

The introduction of observation error drastically changes the shape of the stationary measure. We analyze the cases $n = 2$ and $n > 2$ separately.

$n = 2$ This case is equivalent to the IPD, where $n = 2, c = 1$. For the IPD, Molander (1985) showed that observation errors lead to a random walk on the state space with the unique stationary measure $\pi = \frac{1}{4}\left(\delta_{\{(1,1)\}} + \delta_{\{(1,2)\}} + \delta_{\{(2,1)\}} + \delta_{\{(2,2)\}}\right)$.

$n > 2$ This is the most important case for the RM game, since service providers can file arbitrary prices. Consequently, we can assume that there is more than a single price level below the price of the cooperative optimum. As before without observation errors, the state (n,n) is absorbing and thus generates the closed communicating class $\{(n,n)\}$. However, in contrast to the case without observation errors, (n,n) is accessible from all other states and thus generates the only closed communicating class of the Markov chain. Since stationary measures are concentrated on closed communicating classes, the unique stationary measure is the Dirac measure $\pi = \delta_{\{(n,n)\}}$. Thus, in this case, any possibility of observation errors will ultimately lead to the Competitive Spiral Down, in which both players play the single-stage Bertrand Nash equilibrium at every stage.

GTFT vs. GTFT Without observation errors, the state of mutual cooperation $(1,1)$ is absorbing. Due to GTFT's generosity, this state is also accessible from all other states. Consequently, the stationary distribution of the symmetric case GTFT vs. GTFT without errors is cooperation $\pi = \delta_{\{(1,1)\}}$.

With observation errors, we have to distinguish two cases similarly to the matchup TFT vs. TFT.

$n = 2$ This case is equivalent to the IPD, which has been extensively studied by Molander (1985), although the author does not explicitly give the stationary measure, which comes to

$$\pi = C\left(\delta_{\{(1,1)\}} + \frac{(1-\gamma)\varepsilon}{(1-\varepsilon)\gamma - \varepsilon}\left(\delta_{\{(1,2)\}} + \delta_{\{(2,1)\}}\right) + \frac{(1-2\gamma+\gamma^2)\varepsilon^2}{(1-2\varepsilon+\varepsilon^2)\gamma^2 + (1-\varepsilon)2\varepsilon\gamma + \varepsilon^2}\delta_{\{(2,2)\}}\right). \tag{5.4.17}$$

with a constant C.

$n > 2$ There is only a single closed communicating class, since generosity causes the state of mutual cooperation $(1,1)$ to be accessible from any state in the state space. Starting from $(1,1)$, observation errors can trigger defection, so that the closed communicating class in this case is $\{(i,j) \mid i,j \in \{1,3,4,5,\dots,n\}\}$. Note that no GTFT player will ever choose 2, since products $i > 1$ are only chosen as a means of retaliation. In this case, the competitor was observed to be playing $i > 1$ and thus the reaction consists of choosing $i + 1 > 2$.

CTFT vs. CTFT In the scenario of two CTFT players without observation errors, the state of mutual cooperation with good reputations $(1,1,g,g)$ is absorbing and accessible from all states in the state space with reputation $J_1 \times J_2 \times R_1 \times R_2$. Therefore, the unique stationary measure is $\pi = \delta_{\{(1,1,g,g)\}}$.

Even with observation errors, we do not need a case-by-case analysis. The CTFT players each correct their own mistakes. If one player is mistakenly observed to be defecting, he gets assigned a bad reputation and is punished for one stage, before both return to cooperating. If both players are observed to be defecting, they each get assigned a bad reputation and immediately cooperate. Thus, there is only a single closed communicating class $\{(1,1,g,g),(1,1,b,b),(1,\min(3,n),b,g),(\min(3,n),1,g,b)\}$, on which the stationary measure is concentrated. Since this closed communicating class only contains states with the two products $i = 1$ and $j = \min(3,n)$, this is equivalent to the IPD and we can compute the stationary measure explicitly. For the case $n = 2$ of the IPD, where $\min(3,n) = n$, we find

$$\pi = \frac{1}{1+2\varepsilon}\left(\left(1-\varepsilon^2\right)\delta_{\{(1,1,g,g)\}} + \varepsilon^2\delta_{\{(1,1,b,b)\}} + \varepsilon\delta_{\{(1,2,b,g)\}} + \varepsilon\delta_{\{(2,1,g,b)\}}\right), \tag{5.4.18}$$

and for $n > 2$, where $\min(3,n) = 3$, we have

$$\pi = \frac{1}{1+2\varepsilon}\left(\left(1-\varepsilon^2\right)\delta_{\{(1,1,g,g)\}} + \varepsilon^2\delta_{\{(1,1,b,b)\}} + \varepsilon\delta_{\{(1,3,b,g)\}} + \varepsilon\delta_{\{(3,1,g,b)\}}\right). \tag{5.4.19}$$

PAVLOV vs. PAVLOV For a pair of PAVLOV players without observation errors, the state $(1, 1)$ is absorbing and accessible from all other states. Thus, $\pi = \delta_{\{(1,1)\}}$ is the unique stationary measure of the Markov chain.

With observation errors, there exists only a single closed communicating class, since $(1, 1)$ is accessible from every state in the state space. Observation errors can lead to defection from mutual cooperation, as long as one player is observed to be cooperating and the other to be defecting. We have to distinguish two separate cases to determine the form of the closed communicating class.

$n = 2$ This case represents the IPD case, in which players will react to an observation (i, j) with $i \neq j$ with mutual defection. Similarly, (i, i) will lead to mutual cooperation in the next stage. The closed communicating class is thus $\{(1, 1), (2, 2)\}$.

$n > 2$ For a larger amount of products, the reaction to an observation of $(1, 2)$ is to play $(3, 2)$. Without observation errors, this will lead to mutual cooperation in the next stage. However, if players observe $(3, 1)$ instead, this will lead to $(2, 4)$ via mutual defection in the next stage. The resulting closed communicating class in this case is $\{(1, 1)\} \cup \{(2, i) \mid i > 2\} \cup \{(i, 2) \mid i > 2\}$.

MATCH vs. MATCH Without errors, the complete state space is partitioned into closed communicating classes $\{(i, j), (j, i)\}$. Each of these closed communicating classes possesses stationary measure $\pi = \frac{1}{2}\left(\delta_{\{(i,j)\}} + \delta_{\{(j,i)\}}\right)$. In the case $i = j$, the closed communicating class simplifies to $\{(i, i)\}$ and the stationary measure is $\pi = \delta_{\{(i,i)\}}$.

With observation errors, MATCH vs. MATCH turns into a random walk on $J_1 \times J_2$, for which the stationary measure is the uniform distribution on the state space.

PAVLOV vs. MATCH While MATCH follows a similar logic as the Tit for Tat-like strategies, it behaves very differently to PAVLOV. Thus, we will take a look at PAVLOV vs. MATCH in more detail.

Without observation errors, we find that the stationary measure is concentrated on separate closed communicating classes. More specifically, we have the following cases:

$n = 2$ There are two closed communicating classes, $\{(1, 1)\}$ and $\{(1, 2), (2, 1), (2, 2)\}$. Thus, the starting distribution is responsible for the resulting stationary distribution: If the players start in the cooperative stance $(1, 1)$, then the process will remain at $(1, 1)$ with probability one. Otherwise, the process will end up in the closed communicating class $\{(1, 2), (2, 1), (2, 2)\}$.

$n > 2$ Similarly to the previous case, we find two closed communicating classes $\{(1, 1)\}$ and $\{(1, 2), (2, 3), (3, 1)\}$. The only difference to the previous case $n = 2$ is that DEFECT is not restricted by the number of classes n.

With observation errors, we find that there is always only a single closed communicating class, with a structure depending on the parameter n.

$n = 2$ In this case, due to observation errors, any action can be considered as cooperative by the competitor. Furthermore, we find that observation errors render all states accessible, so that all states are members of the closed communicating class.

$n > 2$ In this case, the closed communicating class encompasses more states in which MATCH exploits PAVLOV than the other way around. This is a result of the interaction of PAVLOV and MATCH as well as of the influence of observation errors. In particular, the mechanism of PAVLOV will lead to cooperation, after which both strategies will spiral down to mutual defection, at which point the cycle restarts. Observation errors lead to the accessibility of more states than in the case without observation errors for $n > 2$ discussed above. The closed communicating class is $\{(1, j) \mid j \in \{1, \ldots, n\}\} \cup \{(2, j) \mid j \in \{2, \ldots, n\}\} \cup \{(i, 1) \mid i \in \{3, \ldots, n\}\}$.

5.4.3 Erroneous Threshold Products

So far, we have always assumed that both service providers have perfect knowledge of the cooperative threshold product c. In this section, we will explore the consequences of each player S_k having his own version c_k of the cooperative threshold product. This may be caused e.g. by errors during estimation, if the player tries to compute c_k from his observations, or it may be the result of communication mishaps, if the players try to signal the correct value for c_k to one another.

In the preceding sections of this chapter, we have only treated the case $c_1 = c_2 = 1$. However, different values of c_k lead to changes in the strategies. We will use TFT and PAVLOV without errors to highlight typical behavior, since both strategies represent different, but successful, approaches to the IPD and the RM game. TFT has proved successful owing to its nice attitude, forgiveness, retaliation and clarity (as noted by Axelrod (1984)) and the success of TFT was the catalyst for the creation of many other similar strategies in the IPD as well as in the RM game. In fact, its robust variations are intended to maintain these properties despite the presence of observation errors. In contrast to this, PAVLOV adds the possibility to exploit suckers, while giving up robustness against aggressors. Wedekind and Milinski (1996) found that human players of the IPD ended up playing either a variation of TFT or PAVLOV. This highlights the intuitive behavior of these strategies, which strengthens the case for their closer examination.

Both strategies are easily adapted to this case. In this section, COOP amounts to playing c_k. DEFECT means underbidding the competitor, while never offering a more expensive product than c_k, since c_k is taken to be jointly optimal. Recall that in this section, we consider the case without observation errors.

TFT

$$M_{TFT} : V_l \to V_k \tag{5.4.20}$$

$$e_i \mapsto \begin{cases} e_{c_k} & \text{if } i \leq c_k \\ e_{\min(\max(i+1,c_k),n)} & \text{else} \end{cases}, \tag{5.4.21}$$

where we used the notation of Section 5.4.1.

- If $c_1 = c_2 \neq c$, we only need to replace $c = 1$ by c_1 to execute the changes in the stationary distributions. More specifically, there are three closed communicating classes in $\{(c_1, c_1)\}$, $\{(c_1, n), (n, c_1)\}$ and $\{(n, n)\}$. Similarly to the case $c = 1$ in Section 5.4.2, the state of mutual cooperation (c_1, c_1) is accessible only from itself, while (n, n) is accessible only from all states of mutual defection $\{(i, j) \mid i, j > c_1\}$. The third closed communicating class $\{(c_1, n), (n, c_1)\}$ is accessible from the remaining states $\{(c_1, j) \mid j > c_1\} \cup \{(i, c_1) \mid i > c_1\}$. Consequently, a pair of TFT-playing service providers end up each playing c_1 with probability 1 as long as they start with a cooperative stance. The jointly optimal payoff is achieved by playing c, thus mistakes in determining the correct cooperative threshold product lead to a loss in revenue. However, there is no structural change and the Competitive Spiral Down will be avoided just the same, as long as both players do not use $c_1 = c_2 = n$.
- If $c_1 \neq c_2$, we find more drastic effects on the stationary distributions. In this case, the state of a game of two TFT players will converge to the pair of lowest prices (n, n) with probability 1. Thus, even without any observation errors, we find the Competitive Spiral Down effect, so that the game ends up in the single-stage Bertrand-Nash equilibrium.

PAVLOV In the notation of Section 5.4.1, PAVLOV takes the form

$$\overline{M}_{PAVLOV} : V_1 \otimes V_2 \to V_k \tag{5.4.22}$$

$$e_i \otimes e_j \mapsto \begin{cases} e_{c_k} & \text{if } i, j \leq c_k \\ & \text{ or } i, j > c_k \\ e_{\min(\max(j+1,c_k),n)} & \text{else} \end{cases}. \tag{5.4.23}$$

- If $c_1 = c_2 \neq c$, there is no structural change in the stationary distributions. Similarly to TFT, a pair of PAVLOV-playing firms ends up cooperating, although at a different level and therefore with a lower payoff than in the optimal case $c_1 = c_2 = c$.
- If $c_1 \neq c_2$, two PAVLOV players end up alternating between cooperation and defection. This is the same level of cooperative behavior that PAVLOV shows against a pure aggressor playing ALLD.

While knowing the exact value of c might be impossible in practice, slight deviations from the optimal value lead only to slight deviations from the jointly optimal payoff. However, our results indicate that it is paramount for both players to use the same value for c when implementing a repeated game strategy. Otherwise, even the smallest error in c can lead to a complete Competitive Spiral Down.

Consequently, we find that players should not try to estimate the correct value of c based on their observations, as this bears the danger of ending up with different values for different players. Instead, service providers should communicate with their competitors in order to come to an agreement on the value of c. This way, it can be ensured that both use the same value, even if it may be wrong. As we have demonstrated in this section, neither player has an incentive to be dishonest in this situation due to the danger inherent in using different values for the cooperative threshold product.

5.4.4 Example

Since the state space $J_1 \times J_2$ of the RM game Markov chain is n^2-dimensional, the depiction of transition matrices can become infeasible quickly. Nevertheless, we believe that an explicit discussion of the properties of the Markov chain for the strategies presented in Section 5.4.1 can benefit the reader, as long as the examples are kept as simple as possible. For the case $n = 2, c = 1$ of the IPD, Molander (1985) has analyzed transition matrices and stationary measures of GTFT against some typical strategies. However, in Section 5.4.2, we have seen that as soon as there is more than one price below the price level of tacit collusion, the properties of the stationary measure of TFT vs. TFT undergo a radical change in the presence of observation errors. Thus, for illustration purposes we will give examples for typical pairs of strategies for the smallest choice of parameters $n = 3, c = 1$ that satisfies $n > c + 1$.

The matrices of the single-player strategies are the building blocks of the transition matrices of the Markov chain. As discussed in Section 5.2, these can take different forms. For a reactive strategy, the single-player strategy matrix $\widetilde{M}_\sigma$ is a 3×3 matrix, whereas a single-player strategy $\overline{M}_\sigma$ in the form of a general Markov strategy requires a 9×3 matrix. With reputation, as is necessary for CTFT, a single-player strategy is described by a 36×3 matrix even in this simple example. Since this renders the example not very instructive, we will not present matchups involving CTFT here, even though the analysis in this case can be carried out similarly to our examples.

The stationary measure vector π can be calculated as a left-eigenvector to the eigenvalue 1, since $\pi M = \pi$, where M is the transition matrix of the Markov process. However, in some cases it can be more simple as well as more instructive to use a stochastic reasoning. As a part of simplifying the analysis, we can identify unreachable states as they are indicated by columns of the transition matrix containing only zeros. We will call the restriction of a Markov chain to the state space without the unreachable states the reduced Markov chain. Due to these states being unreachable, they cannot influence the stationary measure of the process. Therefore, we can carry out the calculation of the stationary measure vector for the reduced Markov chain and then lift the result to the original state space.

Independently of the strategies under examination, we need the matrix E of the observation error with probability ε

$$E = \begin{pmatrix} 1-\varepsilon & \varepsilon & 0 \\ \varepsilon & 1-2\varepsilon & \varepsilon \\ 0 & \varepsilon & 1-\varepsilon \end{pmatrix}, \tag{5.4.24}$$

as well as the basis change matrix B

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.4.25}$$

As benchmarks, we will use the non-cooperative best response ALLD and cooperative best response ALLC strategies throughout this section. Described in the form of reactive strategies these take the form

$$\widetilde{M}_{ALLD} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.4.26}$$

and

$$\widetilde{M}_{ALLC} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{5.4.27}$$

TFT

Since TFT is a reactive strategy, the single-player strategy TFT yields the 3×3 matrix

$$\widetilde{M}_{TFT} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.4.28}$$

TFT vs. TFT Without observation errors, we get for the transition matrix of the matchup TFT vs. TFT

$$M_{TFT,TFT} = B\left(\widetilde{M}_{TFT} \otimes \widetilde{M}_{TFT}\right) \tag{5.4.29}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.4.30}$$

There are three closed communicating classes, $\{(1,1)\}$, $\{(1,3),(3,1)\}$ and $\{(3,3)\}$. We can compute a stationary distribution π_i for each closed communicating class:

$$\pi_1 = (1,0,0,0,0,0,0,0,0) = e_1 \otimes e_1 \tag{5.4.31}$$

$$\pi_2 = \frac{1}{2}(0,0,1,0,0,0,1,0,0) = \frac{1}{2}\left(e_1 \otimes e_3 + e_3 \otimes e_1\right) \tag{5.4.32}$$

$$\pi_3 = (0,0,0,0,0,0,0,0,1) = e_3 \otimes e_3 \tag{5.4.33}$$

Here, π_1 belongs to the closed communicating class $\{(1,1)\}$, which is not accessible from any other state. The measure π_2 belongs to $\{(1,3),(3,1)\}$, which is accessible from the states $\{(1,2),(1,3),(2,1),(3,1)\}$, and π_2 to $\{(3,3)\}$, which is accessible from $\{(2,2),(2,3),(3,2),(3,3)\}$. Thus, if and only if the initial move is cooperative, as is usually assumed for TFT, the cooperative measure vector π_1 denotes the long-term outcome of the game. Otherwise, the game can end up either in the single-stage Bertrand equilibrium of the lowest prices $(3,3)$ or a mix of $(1,3)$ and $(3,1)$, where players alternate cooperating and defecting.

Building on $M_{TFT,TFT}$, we can construct the transition matrix with errors

$$M^{\varepsilon}_{TFT,TFT} = \begin{pmatrix} (1-\varepsilon)^2 & 0 & (1-\varepsilon)\,\varepsilon & 0 & 0 & 0 & (1-\varepsilon)\,\varepsilon & 0 & \varepsilon^2 \\ (1-\varepsilon)\,\varepsilon & 0 & \varepsilon^2 & 0 & 0 & 0 & (1-\varepsilon)^2 & 0 & (1-\varepsilon)\,\varepsilon \\ 0 & 0 & 0 & 0 & 0 & 0 & (1-\varepsilon) & 0 & \varepsilon \\ (1-\varepsilon)\,\varepsilon & 0 & (1-\varepsilon)^2 & 0 & 0 & 0 & \varepsilon^2 & 0 & (1-\varepsilon)\,\varepsilon \\ \varepsilon^2 & 0 & (1-\varepsilon)\,\varepsilon & 0 & 0 & 0 & (1-\varepsilon)\,\varepsilon & 0 & (1-\varepsilon)^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & 0 & (1-\varepsilon) \\ 0 & 0 & (1-\varepsilon) & 0 & 0 & 0 & 0 & 0 & \varepsilon \\ 0 & 0 & \varepsilon & 0 & 0 & 0 & 0 & 0 & (1-\varepsilon) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.4.34}$$

Similarly to $M_{TFT,TFT}$, displayed in Equation 5.4.29, the states $(1,2)$, $(2,1)$, $(2,2)$, $(2,3)$ and $(3,2)$ are unreachable. However, in contrast to the matrix in Equation 5.4.29, the only closed communicating class is $\{(3,3)\}$. Since for $\varepsilon > 0$, the state $(3,3)$ is accessible from all other states and is absorbing, the unique stationary distribution vector of this Markov chain $\pi = (0,0,0,0,0,0,0,0,1)^t = e_3 \otimes e_3$. Therefore, as we have shown in Section 5.4.2, any positive error probability will lead to a limiting behavior of both players always choosing the lowest price.

TFT vs. ALLD Without errors, the transition matrix of the game with $\sigma_1 = TFT$ and $\sigma_2 = ALLD$ takes the form

$$M_{TFT,ALLD} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.4.35}$$

The unique stationary measure vector $\pi = (0,0,0,0,0,0,0,0,1)^t = e_3 \otimes e_3$ demonstrates that this matchup will end up with both competitors always choosing the lowest price.

With observation errors, the transition matrix

$$M^{\varepsilon}_{TFT,ALLD} = \begin{pmatrix} 0 & (1-\varepsilon)^2 & (1-\varepsilon)\,\varepsilon & 0 & 0 & 0 & 0 & (1-\varepsilon)\,\varepsilon & \varepsilon^2 \\ 0 & (1-\varepsilon)\,\varepsilon & \varepsilon^2 & 0 & 0 & 0 & 0 & (1-\varepsilon)^2 & (1-\varepsilon)\,\varepsilon \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\varepsilon) & \varepsilon \\ 0 & (1-\varepsilon)\,\varepsilon & (1-\varepsilon)^2 & 0 & 0 & 0 & 0 & \varepsilon^2 & (1-\varepsilon)\,\varepsilon \\ 0 & \varepsilon^2 & (1-\varepsilon)\,\varepsilon & 0 & 0 & 0 & 0 & (1-\varepsilon)\,\varepsilon & (1-\varepsilon)^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & (1-\varepsilon) \\ 0 & 0 & (1-\varepsilon) & 0 & 0 & 0 & 0 & 0 & \varepsilon \\ 0 & 0 & \varepsilon & 0 & 0 & 0 & 0 & 0 & (1-\varepsilon) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.4.36}$$

can be analyzed analogously to the symmetric case TFT vs. TFT in Equation 5.4.34. Since state $(3,3)$ is absorbing and accessible from all other states, as long as $\varepsilon > 0$, the unique stationary measure vector is $\pi = (0,0,0,0,0,0,0,0,1)^t = e_3 \otimes e_3$. In other words, observation errors do not influence the stationary measure of TFT vs. ALLD.

GTFT

For GTFT with generosity parameter γ, the single-player matrix takes the form

$$\widetilde{M}_{GTFT} = \begin{pmatrix} 1 & 0 & 0 \\ \gamma & 0 & 1-\gamma \\ \gamma & 0 & 1-\gamma \end{pmatrix}, \tag{5.4.37}$$

when formulated as a reactive strategy.

GTFT vs. GTFT Without errors the transition matrix

$$M_{GTFT,GTFT} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma & 0 & 0 & 0 & 0 & 0 & 1-\gamma & 0 & 0 \\ \gamma & 0 & 0 & 0 & 0 & 0 & 1-\gamma & 0 & 0 \\ \gamma & 0 & 1-\gamma & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma^2 & 0 & (1-\gamma)\gamma & 0 & 0 & 0 & (1-\gamma)\gamma & 0 & (1-\gamma)^2 \\ \gamma^2 & 0 & (1-\gamma)\gamma & 0 & 0 & 0 & (1-\gamma)\gamma & 0 & (1-\gamma)^2 \\ \gamma & 0 & 1-\gamma & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma^2 & 0 & (1-\gamma)\gamma & 0 & 0 & 0 & (1-\gamma)\gamma & 0 & (1-\gamma)^2 \\ \gamma^2 & 0 & (1-\gamma)\gamma & 0 & 0 & 0 & (1-\gamma)\gamma & 0 & (1-\gamma)^2 \end{pmatrix} \tag{5.4.38}$$

of the symmetric matchup GTFT vs. GTFT allows for a simple calculation of the stationary measure analogous to the case TFT vs. TFT with observation errors. Since for $\gamma > 0$, the state $(1,1)$ is absorbing and accessible from all other states, the unique stationary measure $\pi = (1,0,0,0,0,0,0,0,0)^t = e_1 \otimes e_1$ represents almost surely cooperation.

As in the other examples, $M_{GTFT,GTFT}$ can be used to construct the transition matrix in the presence of observation errors $M^{\varepsilon}_{GTFT,GTFT}$. The transition matrix $M^{\varepsilon}_{GTFT,GTFT}$ is too big to be displayed here. However, all but four columns of $M^{\varepsilon}_{GTFT,GTFT}$ contain only zeros. We can use this to simplify the matrix by reducing it to reachable states of the Markov chain, since columns containing only zeros indicate unreachable states. Recall that the state (i,j) corresponds to the basis element $e_i \otimes e_j$ of the tensor product $V_1 \otimes V_2$. Thus, restricting the transition matrix to the space generated by $\{e_i \otimes e_j \mid i \in I_1, j \in I_2\}$ corresponds to finding the transition matrix of the Markov chain restricted to the state space $\{(i,j) \mid i \in I_1, j \in I_2\}$. Removing the zero-columns corresponding to the unreachable states $\{(1,1),(1,3),(3,1),(3,3)\}$, we find the reduced transition matrix as

$$\begin{pmatrix} (1-\varepsilon+\gamma\varepsilon)^2 & (1-\gamma)\varepsilon(1-\varepsilon+\gamma\varepsilon) & (1-\gamma)\varepsilon(1-\varepsilon+\gamma\varepsilon) & (1-\gamma)^2\varepsilon^2 \\ \gamma(1-\varepsilon+\gamma\varepsilon) & (1-\gamma)\gamma\varepsilon & (1-\gamma)(1-\varepsilon+\gamma\varepsilon) & (1-\gamma)^2\varepsilon \\ \gamma(1-\varepsilon+\gamma\varepsilon) & (1-\gamma)(1-\varepsilon+\gamma\varepsilon) & (1-\gamma)\gamma\varepsilon & (1-\gamma)^2\varepsilon \\ \gamma^2 & (1-\gamma)\gamma & (1-\gamma)\gamma & (1-\gamma)^2 \end{pmatrix}. \tag{5.4.39}$$

Since for $\gamma, \varepsilon > 0$, every element of this matrix is positive, there is only a single communicating class and the Markov chain is irreducible. Thus, the eigenvector to the eigenvalue 1 yields the unique stationary measure vector

$$\pi_{red} = \left(\frac{1}{\gamma + (1-\gamma)\varepsilon}\right)^2 \left(\gamma^2, \gamma(1-\gamma)\varepsilon, \gamma(1-\gamma)\varepsilon, (1-\gamma)^2\varepsilon^2\right)^t \tag{5.4.40}$$

$$= \left(\frac{1}{\gamma + (1-\gamma)\varepsilon}\right)^2 \Big(\gamma^2 e_1 \otimes e_1 + \gamma(1-\gamma)\varepsilon e_1 \otimes e_3 \tag{5.4.41}$$

$$+ \gamma(1-\gamma)\varepsilon e_3 \otimes e_1 + (1-\gamma)^2\varepsilon^2 e_3 \otimes e_3\Big) \tag{5.4.42}$$

$$= \left(\frac{1}{\gamma + (1-\gamma)\varepsilon}\right)^2 (\gamma e_1 + (1-\gamma)\varepsilon e_3) \otimes (\gamma e_1 + (1-\gamma)\varepsilon e_3) \tag{5.4.43}$$

of the reduced Markov chain. The representation of the reduced stationary distribution in Equation 5.4.43 as a tensor product provides a simple interpretation of the stationary measure as a result of both players playing a mixed strategy. In the tensor product $(\gamma e_1 + (1-\gamma)\varepsilon e_3) \otimes (\gamma e_1 + (1-\gamma)\varepsilon e_3)$, the first term $(\gamma e_1 + (1-\gamma)\varepsilon e_3)$ denotes the first player's stationary strategy, while the second term denotes the second player's stationary strategy. The first term $\left(\frac{1}{\gamma+(1-\gamma)\varepsilon}\right)^2$ in Equation 5.4.43 is a normalizing factor. Thus, each player ends up playing a mixture of the highest and lowest price, with the probability of each action determined by the generosity parameter γ and the observation error probability ε. In particular, we find that without observation errors, i.e. $\varepsilon = 0$, the reduced stationary distribution $\pi_{red} = e_1 \otimes e_1$ is always cooperative.

Since unreachable states cannot influence the stationary distribution, we can use π_{red} to find the stationary distribution π of the Markov chain on the original state space:

$$\pi = \left(\frac{1}{\gamma + (1-\gamma)\varepsilon}\right)^2 (\gamma e_1 + (1-\gamma)\varepsilon e_3) \otimes (\gamma e_1 + (1-\gamma)\varepsilon e_3) \tag{5.4.44}$$

$$= \left(\frac{1}{\gamma + (1-\gamma)\varepsilon}\right)^2 \left(\gamma^2, 0, \gamma(1-\gamma)\varepsilon, 0, 0, 0, \gamma(1-\gamma)\varepsilon, 0, (1-\gamma)^2\varepsilon^2\right)^t \tag{5.4.45}$$

GTFT vs. ALLD Without observation errors, the transition matrix of GTFT against the aggressive strategy ALLD

$$M_{GTFT,ALLD} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 & 0 & 0 & 0 & 1-\gamma & 0 \\ 0 & \gamma & 0 & 0 & 0 & 0 & 0 & 1-\gamma & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma & 0 & 0 & 0 & 0 & 0 & 1-\gamma \\ 0 & 0 & \gamma & 0 & 0 & 0 & 0 & 0 & 1-\gamma \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma & 0 & 0 & 0 & 0 & 0 & 1-\gamma \\ 0 & 0 & \gamma & 0 & 0 & 0 & 0 & 0 & 1-\gamma \end{pmatrix} \tag{5.4.46}$$

reveals the unreachable states $\{(1,1),(2,1),(2,2),(2,3),(3,1)\}$. Restricting the transition matrix to the space generated by $\{e_1 \otimes e_2, e_1 \otimes e_3, e_3 \otimes e_2, e_3 \otimes e_3\}$ yields the transition matrix

$$\begin{pmatrix} \gamma & 0 & 1-\gamma & 0 \\ \gamma & 0 & 1-\gamma & 0 \\ 0 & \gamma & 0 & 1-\gamma \\ 0 & \gamma & 0 & 1-\gamma \end{pmatrix} \tag{5.4.47}$$

of the Markov process restricted to the reachable states $\{(1,2),(1,3),(3,2),(3,3)\}$. This reduced Markov process is irreducible, and its unique stationary measure vector of the reduced Markov chain is

$$\pi_{red} = \left(\gamma^2, \gamma(1-\gamma), \gamma(1-\gamma), (1-\gamma)^2\right)^t, \tag{5.4.48}$$

which leads to the unique stationary measure vector of the original Markov chain

$$\pi = \left(0, \gamma^2, \gamma(1-\gamma), \gamma(1-\gamma), 0, 0, 0, 0, (1-\gamma)^2\right)^t. \tag{5.4.49}$$

As expected, the stationary measure for the special case $\gamma = 0$ is identical to the stationary measure of TFT vs. ALLD.

Introducing observation errors produces the matrix $M^{\varepsilon}_{GTFT,ALLD}$, which has the same unreachable states as $M_{GTFT,ALLD}$. Thus, it can be restricted to the reduced state space $\{(1,2),(1,3),(3,2),(3,3)\}$ to yield the reduced transition matrix

$$\begin{pmatrix} (1-\varepsilon)(\gamma+\varepsilon-\gamma\varepsilon) & \varepsilon(\gamma+\varepsilon-\gamma\varepsilon) & (1-\gamma)(1-\varepsilon)^2 & (1-\gamma)(1-\varepsilon)\varepsilon \\ \gamma(1-\varepsilon) & \gamma\varepsilon & (1-\gamma)(1-\varepsilon) & (1-\gamma)\varepsilon \\ 0 & (\gamma+\varepsilon-\gamma\varepsilon) & 0 & (1-\gamma)(1-\varepsilon) \\ 0 & \gamma & 0 & (1-\gamma) \end{pmatrix} \tag{5.4.50}$$

operating on the space created by $\{e_1 \otimes e_2, e_1 \otimes e_3, e_3 \otimes e_2, e_3 \otimes e_3\}$. This irreducible Markov chain yields the reduced stationary measure vector

$$\pi_{red} = \frac{1}{C} \begin{pmatrix} \gamma^2 (1-\varepsilon) \\ \gamma(1-\gamma-\varepsilon+2\gamma\varepsilon+\varepsilon^2-\gamma\varepsilon^2) \\ \gamma (1-\varepsilon)(1-\gamma)(\varepsilon^2-\varepsilon+1) \\ (1-\gamma)(\varepsilon^2-\varepsilon+1)(1-\gamma-\varepsilon+2\gamma\varepsilon+\varepsilon^2-\gamma\varepsilon^2) \end{pmatrix}, \tag{5.4.51}$$

where C is some normalizing constant, and finally the stationary measure vector of the complete Markov process

$$\pi = \frac{1}{C} \begin{pmatrix} 0 \\ \gamma^2 (1-\varepsilon) \\ \gamma(1-\gamma-\varepsilon+2\gamma\varepsilon+\varepsilon^2-\gamma\varepsilon^2) \\ 0 \\ 0 \\ 0 \\ 0 \\ \gamma (1-\varepsilon)(1-\gamma)(\varepsilon^2-\varepsilon+1) \\ (1-\gamma)(\varepsilon^2-\varepsilon+1)(1-\gamma-\varepsilon+2\gamma\varepsilon+\varepsilon^2-\gamma\varepsilon^2) \end{pmatrix}. \tag{5.4.52}$$

PAVLOV

Since PAVLOV depends on both players' last actions, it can not be displayed as a reactive strategy. Instead, we have to resort to the more extensive general Markov strategy form to find the single-player strategy matrix

$$\overline{M}_{PAVLOV} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{5.4.53}$$

Consequently, any strategy interacting with PAVLOV has to take the general Markov strategy form as well. The general Markov strategy single-player strategy matrices of

ALLD and ALLC are

$$\overline{M}_{ALLD} = 1_3 \otimes \widetilde{M}_{ALLD} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.4.54}$$

and

$$\overline{M}_{ALLC} = 1_3 \otimes \widetilde{M}_{ALLC} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{5.4.55}$$

PAVLOV vs. PAVLOV In the symmetric matchup without observation errors, the transition matrix

$$M_{PAVLOV,PAVLOV} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.4.56}$$

shows that the only accessible states are $\{(1,1),(2,3),(3,2)\}$, since the columns referring to all other states are populated by zeros. The reduced Markov chain is irreducible and possesses the unique stationary measure vector $\pi_{red} = e_1 \otimes e_1$, leading to the unique stationary measure vector $\pi = e_1 \otimes e_1 = (1,0,0,0,0,0,0,0,0)^t$ of the original Markov chain. Thus, similarly to GTFT, the mechanism of the PAVLOV strategy leads to almost sure cooperation in the symmetric matchup without observation errors independent of the initial distribution. Recall that TFT vs. TFT did not possess this property.

As before, we want to examine robustness of PAVLOV with respect to observation errors. The transition matrix

$$M^{\varepsilon}_{PAVLOV,PAVLOV} = \begin{pmatrix} 1-2\varepsilon+2\varepsilon^2 & 0 & 0 & 0 & 0 & (1-\varepsilon)\,\varepsilon & 0 & (1-\varepsilon)\,\varepsilon & 0 \\ 2\,(1-\varepsilon)\,\varepsilon & 0 & 0 & 0 & 0 & \varepsilon^2 & 0 & (1-\varepsilon)^2 & 0 \\ \varepsilon & 0 & 0 & 0 & 0 & 0 & 0 & (1-\varepsilon) & 0 \\ 2\,(1-\varepsilon)\,\varepsilon & 0 & 0 & 0 & 0 & (1-\varepsilon)^2 & 0 & \varepsilon^2 & 0 \\ 1-2\varepsilon+2\varepsilon^2 & 0 & 0 & 0 & 0 & (1-\varepsilon)\,\varepsilon & 0 & (1-\varepsilon)\,\varepsilon & 0 \\ (1-\varepsilon) & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & 0 \\ \varepsilon & 0 & 0 & 0 & 0 & (1-\varepsilon) & 0 & 0 & 0 \\ (1-\varepsilon) & 0 & 0 & 0 & 0 & \varepsilon & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.4.57}$$

shows that the addition of observation errors does not change the set of reachable states. The exclusion of unreachable states renders the reduced Markov chain characterized by the transition matrix

$$\begin{pmatrix} 1-2\varepsilon+2\varepsilon^2 & (1-\varepsilon)\,\varepsilon & (1-\varepsilon)\,\varepsilon \\ (1-\varepsilon) & 0 & \varepsilon \\ (1-\varepsilon) & \varepsilon & 0 \end{pmatrix} \tag{5.4.58}$$

irreducible. This leads to the unique stationary measure vector

$$\pi = \frac{1}{1+2\varepsilon}(1,0,0,0,0,\varepsilon,0,\varepsilon,0)^t \tag{5.4.59}$$

$$= \frac{1}{1+2\varepsilon}\left(e_1 \otimes e_1 + \varepsilon e_2 \otimes e_3 + \varepsilon e_3 \otimes e_2\right). \tag{5.4.60}$$

Therefore, observation errors lead to a disturbance of the almost sure cooperation. However, in contrast to TFT vs. TFT, the change in behavior is rather small.

PAVLOV vs. ALLD Without errors, the game of a PAVLOV player against a ALLD-playing competitor produces the transition matrix

$$M_{PAVLOV,ALLD} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{5.4.61}$$

This Markov chain possesses the stationary measure vector

$$\pi = \frac{1}{2}(0,0,1,0,0,0,0,1,0)^t \tag{5.4.62}$$
$$= \frac{1}{2}(e_1 \otimes e_3 + e_3 \otimes e_2). \tag{5.4.63}$$

Thus, the game will end up with a mix of the states $(1,3)$ and $(3,2)$, as PAVLOV alternates between cooperating and defecting.

With observation errors, the transition matrix

$$M^{\varepsilon}_{PAVLOV,ALLD} = \begin{pmatrix} 0 & (1-\varepsilon)^2 & \varepsilon^2 & 0 & 0 & (1-\varepsilon)\varepsilon & 0 & (1-\varepsilon)\varepsilon & 0 \\ 0 & (1-\varepsilon)\varepsilon & (1-\varepsilon)\varepsilon & 0 & 0 & \varepsilon^2 & 0 & (1-\varepsilon)^2 & 0 \\ 0 & 0 & \varepsilon & 0 & 0 & 0 & 0 & (1-\varepsilon) & 0 \\ 0 & (1-\varepsilon)\varepsilon & (1-\varepsilon)\varepsilon & 0 & 0 & (1-\varepsilon)^2 & 0 & \varepsilon^2 & 0 \\ 0 & \varepsilon^2 & (1-\varepsilon)^2 & 0 & 0 & (1-\varepsilon)\varepsilon & 0 & (1-\varepsilon)\varepsilon & 0 \\ 0 & 0 & (1-\varepsilon) & 0 & 0 & 0 & 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon & 0 & 0 & (1-\varepsilon) & 0 & 0 & 0 \\ 0 & 0 & (1-\varepsilon) & 0 & 0 & \varepsilon & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.4.64}$$

leads to the unique stationary measure vector

$$\pi = \frac{1}{2}\left(0,0,1,0,0,\frac{\varepsilon}{\varepsilon+1},0,\frac{1}{\varepsilon+1},0\right)^t \tag{5.4.65}$$
$$= \frac{1}{2}\left(e_1 \otimes e_3 + \frac{\varepsilon}{\varepsilon+1}e_2 \otimes e_3 + \frac{1}{\varepsilon+1}e_3 \otimes e_2\right). \tag{5.4.66}$$

PAVLOV vs. ALLC Without observation errors,

$$M_{PAVLOV,ALLC} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.4.67}$$

is the transition matrix of the game with $\sigma_1 =$ PAVLOV and $\sigma_2 =$ ALLC. The only accessible states of this Markov chain are $\{(1,1),(2,1),(3,1)\}$, with two of them being absorbing states. Thus, we have two closed communicating classes generated by the absorbing states $(1,1)$ and $(2,1)$. Consequently, there exist two stationary measure vectors:

$$\pi_1 = (1,0,0,0,0,0,0,0,0)^t \tag{5.4.68}$$
$$= e_1 \otimes e_1 \tag{5.4.69}$$
$$\pi_2 = (0,0,0,1,0,0,0,0,0)^t \tag{5.4.70}$$
$$= e_2 \otimes e_1 \tag{5.4.71}$$

Which stationary measure the process ends up in depends on the starting distribution. If the process starts in $\{(1,1),(2,2),(2,3),(3,2),(3,3)\}$, it will converge to π_1, if it starts in $\{(1,2),(1,3),(2,1),(3,1)\}$, it will converge to π_2. Thus, a PAVLOV player may end up cooperating with or exploiting a cooperative competitor. At least without observation errors, the starting move is solely responsible for which of these paths the game will take.

With observation errors, the transition matrix

$$M^{\varepsilon}_{PAVLOV,ALLC} = \begin{pmatrix} 2\varepsilon^2-2\varepsilon+1 & 0 & 0 & (1-\varepsilon)\varepsilon & 0 & 0 & (1-\varepsilon)\varepsilon & 0 & 0 \\ 2(1-\varepsilon)\varepsilon & 0 & 0 & \varepsilon^2 & 0 & 0 & (1-\varepsilon)^2 & 0 & 0 \\ \varepsilon & 0 & 0 & 0 & 0 & 0 & (1-\varepsilon) & 0 & 0 \\ 2(1-\varepsilon)\varepsilon & 0 & 0 & (1-\varepsilon)^2 & 0 & 0 & \varepsilon^2 & 0 & 0 \\ 2\varepsilon^2-2\varepsilon+1 & 0 & 0 & (1-\varepsilon)\varepsilon & 0 & 0 & (1-\varepsilon)\varepsilon & 0 & 0 \\ (1-\varepsilon) & 0 & 0 & 0 & 0 & 0 & \varepsilon & 0 & 0 \\ \varepsilon & 0 & 0 & (1-\varepsilon) & 0 & 0 & 0 & 0 & 0 \\ (1-\varepsilon) & 0 & 0 & \varepsilon & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.4.72}$$

indicates no change to the reachable states. However, for any $\varepsilon > 0$, the reduced transition matrix

$$\begin{pmatrix} 2\varepsilon^2-2\varepsilon+1 & (1-\varepsilon)\varepsilon & (1-\varepsilon)\varepsilon \\ 2(1-\varepsilon)\varepsilon & (1-\varepsilon)^2 & \varepsilon^2 \\ \varepsilon & (1-\varepsilon) & 0 \end{pmatrix} \tag{5.4.73}$$

contains only a single closed communicating class and is thus irreducible. Consequently this Markov chain possesses a unique stationary measure vector

$$\pi = \frac{1}{C}\left(1,0,0,\frac{\varepsilon^2-3\varepsilon+2}{\varepsilon^2-2\varepsilon+2},0,0,\frac{2\varepsilon-2\varepsilon^2}{\varepsilon^2-2\varepsilon+2},0,0\right)^t \tag{5.4.74}$$
$$= \frac{1}{C}\left(e_1\otimes e_1 + \frac{\varepsilon^2-3\varepsilon+2}{\varepsilon^2-2\varepsilon+2}e_2\otimes e_1 + \frac{2\varepsilon-2\varepsilon^2}{\varepsilon^2-2\varepsilon+2}e_3\otimes e_1\right), \tag{5.4.75}$$

where C is a normalizing constant. Note that for PAVLOV vsALLC, observation errors lead to the existence of a unique stationary measure with a positive probability for all reachable states. The level of cooperation that a PAVLOV player will show towards a cooperating competitor depends on the probability of observation errors.

5.5 Evolution of Markov Strategies

In this section, we introduce dynamics over time into the Markov version of the repeated RM game. Instead of always choosing the same repeated-game strategy to determine the move at each stage, we allow players to switch to more (or—although improbable—less) successful strategies. This reflects the nature of competition between service providers, in which businesses may adapt their behavior to stay competitive.

This kind of experiment is also useful to analyze interactions of each strategy with all the others, not only one specific opponent. Since m single-player strategies amount to $\binom{m}{2} + m$ different combinations, a separate analysis of each combination becomes infeasible or at least impractical quickly.

For this purpose, we devise a round robin tournament between a fixed amount of players. Similarly to the ecological simulation in the IPD by Axelrod (1984) and Wu and Axelrod (1995) discussed in the literature review in Section 2.2.2, we let each player choose a strategy from a fixed pool of possible repeated-game strategies. After each round, a randomly designated sample of players may reevaluate their strategy. The probability to change to a strategy σ is proportional to the average payoff of σ in the previous round.

As mentioned in Section 2.2.2, in recent publications on the IPD, the ecological approach of Axelrod (1984) has been replaced by the method of Nowak and Sigmund (1992), in which the strategy set is not finite anymore. Instead, each strategy defined by any transition matrix is possible. Nowak and Sigmund argue that this approach matches the random nature of evolution in biological contexts, where mutations can happen aimlessly in every possible direction. However, we feel that this is less true in a business context, since decision makers can be assumed to act rational. Therefore, a business is less likely to risk its revenue by choosing a strategy at random. It is more plausible that decision makers rely on a smaller set of promising strategies, which are well understood and have been tested successfully.

From a technical point of view, the biological approach of Nowak and Sigmund (1992) encounters the problem that "the search space is far too big to perform reliable statistics" (Hauert & Schuster, 1997), in the N-player m-memory IPD. Since in our case, the number of prices n may be arbitrarily high instead of $n = 2$ in the IPD, this is a problem we would encounter as well. This is not only a technical obstacle, but also reinforces our view that this kind of simulation is not appropriate in the business context: We expect businesses to be reluctant to test every possible strategy if it takes that many tries to find a successful combination of transition probabilities.

This is why we prefer Axelrod's finite strategy approach. However, in contrast to the tournaments of Axelrod (1984) and Wu and Axelrod (1995), we do not play out each round of the game. Instead, we use the stationary distributions of the Markov strategies to determine the average payoff.

5.5.1 Experimental Setup

We want to fully exploit the range of payoff configurations that are allowed according to Inequalities 5.3.3 – 5.3.7. To this end, we use the fixed valuation demand model presented in Section 5.3.3, which allows us the best control over the payoffs for the game. The prices chosen in this model have to satisfy

$$2f(j+1) > f(j) > f(j+1) \tag{5.5.1}$$

in order to fulfill the assumptions 5.3.3 – 5.3.7. We will simulate both extremes allowed by the set of Inequalities 5.5.1:

$$f(j) = \begin{cases} f(j+1)+1 & \text{(high temptation)} \\ 2f(j+1)-1 & \text{(low temptation)} \end{cases} \tag{5.5.2}$$

The high temptation case allows for narrow underpricing, which is very profitable. In the low temptation case, underpricing is more hurtful for the aggressor and barely more profitable than keeping the same price level as the competitor. Note that, covering the same price range, more prices lead to a higher temptation.

On a similar note, Boerlijst et al. (1997b) analyzed the effect of the level of temptation in the IPD and found significant differences in the behavior depending on the level of temptation. In particular, they found different sets of Evolutionary Stable Strategies (ESSs) for differing levels of temptation, e.g. PAVLOV is an ESS in the low temptation case, but not in the high temptation case.

We use $n = 10$ different prices, with the cooperative price being set at $c = 1$ as stated in the beginning of this chapter. Choosing the same amount of prices for high and for low temptation leads to a different price range for different levels of temptation, and consequently to different payoffs for the respective games. However, since in this evolutionary setting we measure the success of a strategy at any stage by its proportion of the population, the value of the payoffs of the different game does not need to be the same.

We use all of the strategies mentioned in 5.4.1, i.e. ALLD, ALLC, TFT, GTFT, CTFT, PAVLOV and MATCH. Since in this chapter, the strategy UNDER is identical to ALLD, we do not need to include UNDER explicitly in our analysis. Finding a reasonable starting distribution for these strategies is no easy task. In fact, Nowak and Sigmund (1992) argue that "the main problem here is to find plausible values from which to start" for the finite strategy approach employed in this section. However, the same problem occurs for infinite strategy approach, where the question for a sensible starting distribution is transferred to a larger strategy space. In our simulations, we start with a balanced distribution of aggressive and cooperative strategies. Since ALLD is purely aggressive and the other strategies at least partially cooperative, half of the players start with ALLD and the other half is assigned one of the cooperative strategies in a uniform way. We keep the total population size fixed at 1200, so that 100 players start with each cooperative strategy. This high population size ia motivated by the law of large numbers in order to reduce the weight of singular events and come closer to the expected result.

The dynamics in the evolution are introduced by mutation events similarly to the ecological simulation by Axelrod (1984). After each generation, each player's strategy in the next generation is determined randomly, with the choice probabilities for each strategy proportional to its payoff in the previous generation. Thus, more successful strategies are more likely to be adopted, although a player may also choose to use a less successful strategy or not to change his strategy at all. In order to reduce the influence of outliers, the payoff in each generation should reflect a long-term average of interactions. Fortunately, we do not have to carry out repeated interactions of each combination of players to find the average payoff for each possible matchup in a given generation. Instead, in each round—corresponding to a generation—we can save a lot of computation and rely on the stationary measure for each possible combination of strategies. For a fixed set of parameters, we carry out 100 independent simulations, where each simulation is run for 1000 generations.

5.5.2 Simulation Results

Finding the optimal generosity for a given environment is a difficulty inherent to the strategy GTFT, whereas none of the other considered strategies depends on external parameters. As an additional difficulty, the optimal generosity γ for GTFT depends on the mix of competitor strategies, which is changing over time. In order to find the best generosity for the environment described in Section 5.5.1, we have conducted experiments for different γ in an otherwise unchanged environment as described in the previous section. Figure 5.1 displays the proportion of the total population using GTFT over the level of generosity γ. This is done both for a high and for a low level of temptation as described in Equation 5.5.2.

We find that the level of temptation—and thus the pricing structure—strongly affects the success of GTFT. In the low-temptation case, GTFT always becomes extinct except for generosity levels close to 1, which effectively lead to employing the cooperative strategy ALLC. In contrast to this, given a high temptation, we find that with the right choice of γ, GTFT can be very successful. However, since it is hard to find the correct level of generosity, we will consider the generosity levels of $\gamma \in \{0.2, 0.4, 0.6\}$ separately in the following. This corresponds to choosing too low, ideal and too high levels of generosity in the high-temptation case.

In Figures 5.2 – 5.4, we display the proportion of the total population belonging to each strategy over the elapsed generations. We employ a logarithmic scale for the x-axis to

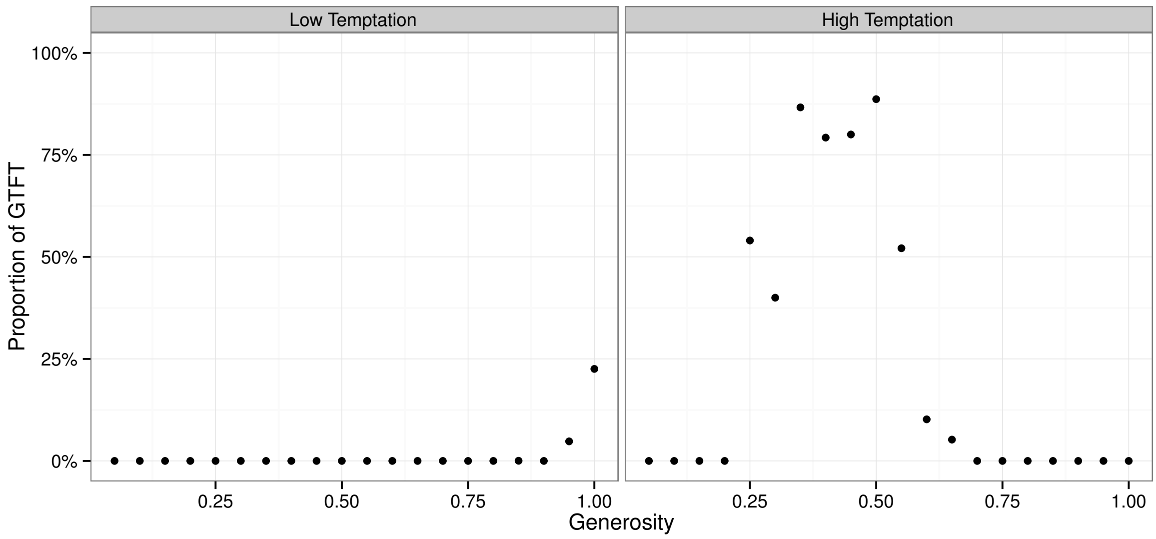

Figure 5.1: Evolutionary success of GTFT for different γ and $\varepsilon = 0.1$

highlight short-term as well as long-term behavior.

Figure 5.2 depicts the mean of all 100 runs in the low-temptation case. We find that although early generations show considerable growth of GTFT, in the end only two strategies, PAVLOV and ALLC, survive. However, we observe that in each run, actually only one of these strategies persists: Either PAVLOV or pure cooperation in the form of ALLC dominates in the long run. In Figure 5.3, we have presented the averaged results of the game conditioned on the final outcome for $\gamma = 0.2$. For different generosities γ, the development of the game given a fixed outcome was similar, although the probability of the outcomes changed (see Figure 5.4). In this scenario, the low temptation makes cooperation attractive, which leads to a rapid decline of ALLD. Of the remaining strategies, ALLC is the most robust against observation errors, since it is in fact independent of its competitor's prices. Thus, along with the decline of ALLD, ALLC becomes the most successful strategy and starts displacing the other strategies. However, if PAVLOV is not extinct or sufficiently decimated, it can build from a considerably outnumbered position to complete a turnaround that sees the strategy ending up in a dominant position. After the environment has grown almost exclusively cooperative, PAVLOV is far more profitable than any other strategy thanks to its error-induced exploitation of ALLC and cooperation in symmetric case. Note however, that over large parts of the time scale PAVLOV is not successful and only starts growing—although rapidly—after the aggressors have vanished after many generations. This is similar to observations of the behavior of PAVLOV in the IPD by Nowak and Sigmund (1993).

In the high-temptation case, displayed in Figure 5.4, we find a larger set of possible

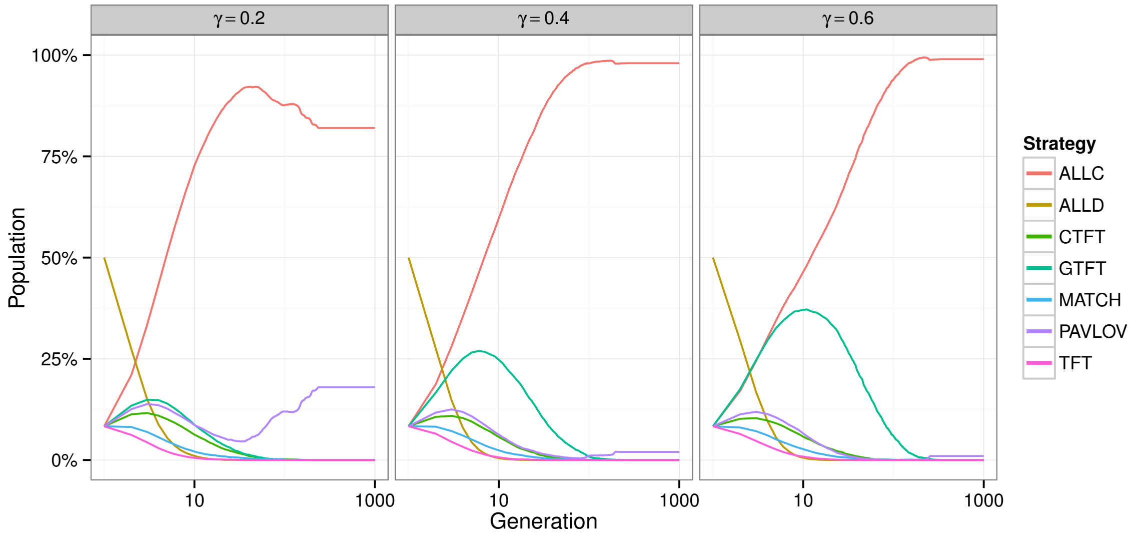

Figure 5.2: Average evolution with low temptation and $\varepsilon = 0.1$

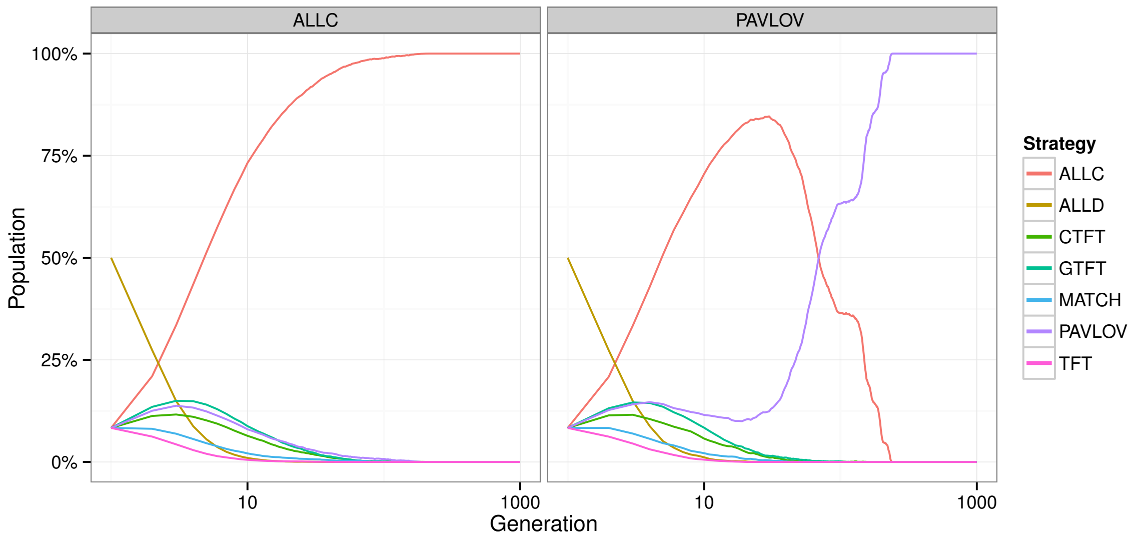

Figure 5.3: Average evolution per outcome with low temptation and $\gamma = 0.2, \varepsilon = 0.1$

surviving strategies, $\{CTFT, GTFT, PAVLOV, MATCH\}$. These can be attributed

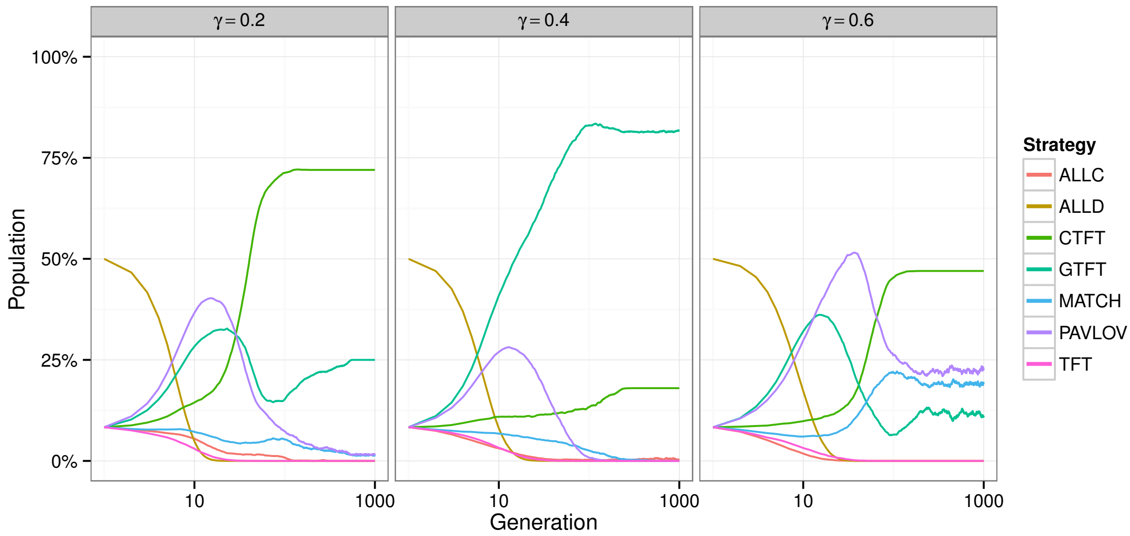

Figure 5.4: Average evolution with high temptation and $\varepsilon = 0.1$

to a higher number of final outcomes as well, which we have documented for the case $\gamma = 0.2$ in Figure 5.5. In this setting, either CTFT or GTFT dominate in the long-term, or the market is shared between PAVLOV and MATCH. Due to the high temptation, cooperation is not as attractive as in the low-temptation case. Therefore, robustness against exploitation from aggressors is more important, leading to the quick extinction of ALLC. As we found in Section 5.4.2, the Tit for Tat variations excel in this regard, whereas PAVLOV tends to cooperate too much with aggressors. Of course, TFT is not robust against errors, but we find CTFT and GTFT to be frequently successful. In rare cases, we find PAVLOV profit from an improbable development, leading to a cooperative environment. As discussed in the low-temptation case, this is where PAVLOV succeeds. However, due to PAVLOV's weakness against MATCH (see Section 5.4.2), the distribution of PAVLOV and MATCH keeps oscillating, although around a stable level.

5.6 Zero-Determinant Strategies

The analysis of the RM game in Chapter 5 has been built on the similarities between the RM game and the IPD. A central feature of the IPD is the non-existence of a single best strategy. No strategy dominates against every possible competition, so that no strategy is successful in every possible scenario. In particular, given any strategy σ_1 presented in Section 5.4.1, it is impossible to predict S_1's or S_2's payoff without knowing the strategy

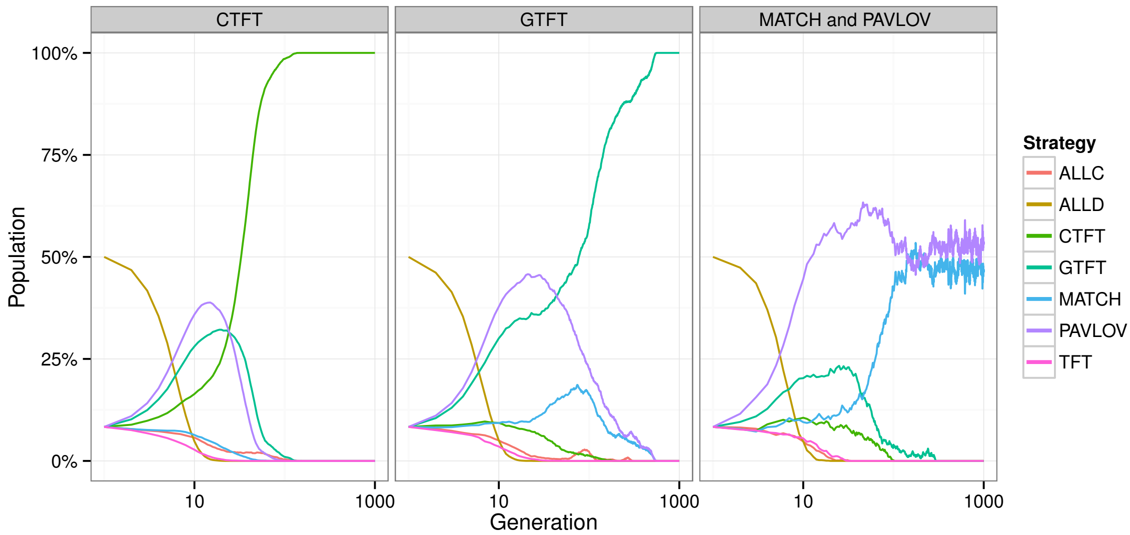

Figure 5.5: Average evolution per outcome with high temptation and $\gamma = 0.2, \varepsilon = 0.1$

σ_2 of S_2. Neither is it possible to establish a relationship for the payoff between a given strategy σ and a competitor strategy, until the choice of competitor strategy is known. Consequently, each player settles for a strategy in the hope that it will perform well in the given competitive environment. The question for the best strategy is tackled by introducing a time dynamics into players' strategies in Section 5.5, where players are allowed to learn from previous encounters and change their strategy appropriately. This way, strategies have to perform in changing environments and a strategy's success in an evolutionary sense is defined by its ability to outperform other strategies within this changing mixture of competitor strategies. Our results in Sections 5.4.2 and 5.5 indicate that the RM game exhibits the same dependence on competitor strategies that prevents predictions about the payoffs without prior knowledge of both players' strategies.

However, as discussed in the literature review in Section 2.2.2, it has been shown that the IPD allows for the existence of so-called Zero-Determinant Strategies that are able to control the relationship of payoffs. First, Boerlijst et al. (1997a) demonstrated the existence of Equalizing Strategies that ensure that both players face the same expected payoff. Independently of Boerlijst et al.'s work, Press and Dyson (2012) proved a vast generalization of this, using an elegant mathematical representation of the payoff in the IPD. Additionally to the existence of Equalizing Strategies, Press and Dyson found that there are strategies in the IPD that demand and get an unfair share of the total payoff. The authors called these strategies Extortionate Strategies. Although Extortionate Strategies work in any setting, they are particularly effective against evolutionary players. If a player using an Extortionate Strategy demands a fixed share of the payoff, the evolution of the competitor will automatically lead to a maximization of profits for the extorting player.

In this section, we will generalize Press and Dyson's analysis of Extortionate Strategies from the IPD case $n = 2, c = 1$ to the case $n \geq 2, c = 1$, where n is the number of products and c denotes the jointly optimal product. In Section 5.6.1, we will introduce the concept of a Zero-Determinant Strategy into the RM game. We will rely on our analysis of the structure of the RM game gained in Section 5.2 to generalize the approach used by Press and Dyson in their analysis of the IPD. In Section 5.6.2, we will show that the generalization of the IPD's payoffs introduced in Section 5.3 guarantees the existence of Extortionate Strategies in the RM game for arbitrary n. Furthermore, we will examine the influence of observation errors on the existence of Extortionate Strategies. We will show that an arbitrarily small error probability leads to the invalidation of Press and Dyson's result even in the simple IPD case with classic payoffs. In Section 5.6.3, we will provide the reader with an example, where we present the steps of our analysis in an explicit way for the case $n = 3$ and $c = 1$.

5.6.1 Prerequisites

In this section, we introduce Zero-Determinant Strategies into the RM game. We focus on general Makov strategies on the state space $J_1 \times J_2$ and ignore the concept of reputation. We start our analysis with $\varepsilon = 0$, i.e. without the possibility of observation errors.

As outlined in Section 5.2.2, we can construct any possible transition matrix M_{σ_1,σ_2} of the Markov RM game from the single-player strategies $\overline{M}_{\sigma_1}$ and $\overline{M}_{\sigma_2}$. If a stationary distribution π exists for the Markov process defined by the transition matrix M_{σ_1,σ_2}, it satisfies

$$\pi^t M_{\sigma_1,\sigma_2} = \pi^t. \tag{5.6.1}$$

Introducing

$$M'_{\sigma_1,\sigma_2} = M_{\sigma_1,\sigma_2} - I_{n^2} \tag{5.6.2}$$

to save some notation, we note that Equation 5.6.1 is equivalent to

$$\pi^t M'_{\sigma_1,\sigma_2} = 0. \tag{5.6.3}$$

The key to Press and Dyson's result in the IPD case was an observation about the relation of the stationary distribution of M_{σ_1,σ_2} and the adjugate matrix of M'_{σ_1,σ_2}, which we want to use in our more general case as well. The adjugate or classical adjoint of a quadratic matrix A is denoted by $\operatorname{adj}(A)$. Its entries are the signed minors of A:

$$\operatorname{adj}(A)_{i,j} = (-1)^{i+j} \det(m(A, j, i)), \tag{5.6.4}$$

where $m(A, i, j)$ is the matrix produced by removing the i-th row and the j-th column of A.

Applying the property 5.1.23 of the adjugate matrix to M'_{σ_1,σ_2} yields

$$\operatorname{adj}(M'_{\sigma_1,\sigma_2})M'_{\sigma_1,\sigma_2} = \det(M'_{\sigma_1,\sigma_2})I_{n^2} = 0, \tag{5.6.5}$$

which implies that every row of $\operatorname{adj}(M'_{\sigma_1,\sigma_2})$ is proportional to the stationary distribution π. Thus, for any vector $v \in V_1 \otimes V_2$ and any row $i \in \{1, \ldots, n^2\}$

$$\pi^t \cdot v = \frac{\sum_{j=1}^{n^2} \operatorname{adj}(M'_{\sigma_1,\sigma_2})_{i,j} \cdot v_j}{\sum_{j=1}^{n^2} \operatorname{adj}(M'_{\sigma_1,\sigma_2})_{i,j}} \tag{5.6.6}$$

$$= \frac{\sum_{j=1}^{n^2} (-1)^j \det(m(M'_{\sigma_1,\sigma_2}, j, i)) \cdot v_j}{\sum_{j=1}^{n^2} (-1)^j \det(m(M'_{\sigma_1,\sigma_2}, j, i))} \tag{5.6.7}$$

$$= \frac{\det(W(M'_{\sigma_1,\sigma_2}, v, i))}{\det(W(M'_{\sigma_1,\sigma_2}, 1, i))}, \tag{5.6.8}$$

where $W(M'_{\sigma_1,\sigma_2}, v, i)$ is equal to M'_{σ_1,σ_2} with the i-th column replaced by v. Without loss of generality, we choose the last column $i = n^2$ for the rest of this section. To save some notation, we will write $W(M'_{\sigma_1,\sigma_2}, v) = W(M'_{\sigma_1,\sigma_2}, v, n)$.

The fact that entries of $\operatorname{adj}(M'_{\sigma_1,\sigma_2})$ are determinants of submatrices of M'_{σ_1,σ_2} enables us to simplify M'_{σ_1,σ_2} without changing the n^2-th row of $\operatorname{adj}(M'_{\sigma_1,\sigma_2})$. More specifically, this means that adding any of the $\{1, \ldots, n^2 - 1\}$-th columns of M'_{σ_1,σ_2} to another does not change the value of the entries in the n^2-th row of $\operatorname{adj}(M'_{\sigma_1,\sigma_2})$, which is the only important row for our calculations in Equation 5.6.6.

Recall that the probability of state (i, j) is represented by the entry corresponding to the basis vector $e_i \otimes e_j$, which is the $(n(i-1)+j)$-th basis vector with our choice of basis 5.2.3. We can use this observation to establish the following relationships between the columns of the transition matrix M_{σ_1,σ_2} and the single-player strategy matrices $\overline{M}_{\sigma_1}$ and $\overline{M}_{\sigma_2}$: For the first player, we have

$$\sum_{t=1}^{n} (M_{\sigma_1,\sigma_2})_{n(i-1)+j,n(r-1)+t} = \sum_{t=1}^{n} \mathbb{P}(X_{s+1} = (r,t) \mid X_s = (i,j)) \tag{5.6.9}$$

$$= \mathbb{P}(X_{s+1}\big|_{V_1} = r \mid X_s = (i,j)) \tag{5.6.10}$$

$$= (\overline{M}_{\sigma_1})_{n(i-1)+j,r} \tag{5.6.11}$$

and analogously for the second player

$$\sum_{r=1}^{n}(M_{\sigma_1,\sigma_2})_{n(i-1)+j,n(r-1)+t} = \sum_{r=1}^{n}\mathbb{P}(X_{s+1}=(r,t)\mid X_s=(i,j)) \tag{5.6.12}$$

$$= \mathbb{P}(X_{s+1}\big|_{V_2}=t\mid X_s=(i,j)) \tag{5.6.13}$$

$$= (\overline{M}_{\sigma_2})_{n(i-1)+j,t}. \tag{5.6.14}$$

In other words, there are n columns in M_{σ_1,σ_2} that sum to the r-th column in $\overline{M}_{\sigma_1}$ as well as n columns that sum to the t-th column in $\overline{M}_{\sigma_2}$.

We want to exploit this for all $r,t=1,\dots,n$ to transform the matrix M'_{σ_1,σ_2}. Formally, we can add the i-th to the j-th column of a matrix M by multiplying the matrix from the right with

$$E^{i,j} = \left(I_{n^2} + (\delta_{k,i}\delta_{l,j})_{k,l=1,\dots,n}\right), \tag{5.6.15}$$

where $(\delta_{k,i}\delta_{l,j})_{k,l=1,\dots,n}$ is the $n\times n$ matrix, for which each element is a product of Kronecker deltas $\delta_{i,j}$ defined as

$$\delta_{i,j} = \begin{cases}1 & \text{if } i=j\\ 0 & \text{if } i\neq j.\end{cases}. \tag{5.6.16}$$

Consequently, $(\delta_{k,i}\delta_{l,j})_{k,l=1,\dots,n}$ denotes the matrix with only zeros and a single one in the i-th row and j-th column. Using this notation, we can modify M'_{σ_1,σ_2} to get

$$M''_{\sigma_1,\sigma_2} = M'_{\sigma_1,\sigma_2}\cdot\prod_{i=1}^{n}\prod_{j=1}^{n-1}\left(E^{n(i-1)+j,n(i-1)+n}\right)\cdot\prod_{i=1}^{n}\prod_{j=1}^{n-1}\left(E^{n(j-1)+i,n(n-1)+i}\right). \tag{5.6.17}$$

The matrix M''_{σ_1,σ_2} takes a convenient form for the analysis of the existence of Zero-Determinant Strategies. The columns $\{n(r-1)+n\mid r\in\{1,\dots,n\}\}$ of the matrix M''_{σ_1,σ_2} are independent of the strategy of the second player, while the columns $\{n(n-1)+t\mid t\in\{1,\dots,n\}\}$ are independent of the strategy of the first player. Since the last column fulfills both conditions, it is constant. More precisely, due to Equations 5.6.9 – 5.6.11, the $(n(r-1)+n)$-th column of M''_{σ_1,σ_2} corresponds to the r-th column of $\overline{M}_{\sigma_1}$ minus a Kronecker delta inherited from subtracting the identity matrix in Equation 5.6.2.

$$(M''_{\sigma_1,\sigma_2})_{n(i-1)+j,n(r-1)+n} = \sum_{t=1}^{n}(M'_{\sigma_1,\sigma_2})_{n(i-1)+j,n(r-1)+t} \tag{5.6.18}$$

$$= \sum_{t=1}^{n}(M_{\sigma_1,\sigma_2}-I_{n^2})_{n(i-1)+j,n(r-1)+t} \tag{5.6.19}$$

$$= (\overline{M}_{\sigma_1})_{n(i-1)+j,r}-\delta_{i,r}. \tag{5.6.20}$$

Similarly, Equations 5.6.12 – 5.6.14 imply that the $(n(n-1)+t)$-th column corresponds to the t-th column of $\overline{M}_{\sigma_2}$ except for a Kronecker delta from the identity matrix in Equation 5.6.2.

$$(M''_{\sigma_1,\sigma_2})_{n(n-1)+j,n(n-1)+t} = \sum_{r=1}^{n}(M'_{\sigma_1,\sigma_2})_{n(i-1)+j,n(r-1)+t} \tag{5.6.21}$$

$$= \sum_{r=1}^{n}(M_{\sigma_1,\sigma_2} - I_{n^2})_{n(i-1)+j,n(r-1)+t} \tag{5.6.22}$$

$$= (\overline{M}_{\sigma_2})_{n(i-1)+j,t} - \delta_{j,t}. \tag{5.6.23}$$

Recall that this does not change the n^2-th row of its adjugate matrix:

$$\Big(\mathrm{adj}(M''_{\sigma_1,\sigma_2})\Big)_{n^2,r} = \Big(\mathrm{adj}(M'_{\sigma_1,\sigma_2})\Big)_{n^2,r} \quad \forall r = 1,\ldots,n^2 \tag{5.6.24}$$

Consequently, Equation 5.6.6 yields

$$\pi^t \cdot v = \frac{\det(W(M''_{\sigma_1,\sigma_2}, v))}{\det(W(M''_{\sigma_1,\sigma_2}, 1))}. \tag{5.6.25}$$

As mentioned in Equation 5.3.8, this can be used to calculate the stationary payoff for each player:

$$s_k = \pi^t \cdot y_k = \frac{\det(W(M''_{\sigma_1,\sigma_2}, y_k))}{\det(W(M''_{\sigma_1,\sigma_2}, 1))} \tag{5.6.26}$$

As the determinant is linear in every column, we get for an affine transformation of payoffs

$$\zeta s_1 + \eta s_2 + \theta = \pi^t \cdot (\zeta y_1 + \eta y_2 + \theta 1_{n^2}) \tag{5.6.27}$$

$$= \frac{\det(W(M''_{\sigma_1,\sigma_2}, \zeta y_1 + \eta y_2 + \theta 1_{n^2}))}{\det(W(M''_{\sigma_1,\sigma_2}, 1))}, \tag{5.6.28}$$

where $1_{n^2} = \sum_{i=1}^{n}\sum_{j=1}^{n} e_i \otimes e_j$ is the vector with all entries equal to 1.

This determinant vanishes if the last column is linearly dependent on the other columns. In this case, the payoffs of the players are related via

$$\zeta s_1 + \eta s_2 + \theta = 0. \tag{5.6.29}$$

If one player chooses a strategy, so that the determinant vanishes independently of the other player's strategy, thus enforcing the linear relationship 5.6.29, this strategy is called a Zero-Determinant Strategy (ZD).

The existence of ZDs depends on the linear relation as well as on the payoffs of the game. For example, already in the case $n = 2$, Press and Dyson (2012) have shown that there exists no ZD for the first player's choice $\eta = 0$. This means that no player can unilaterally set his own score. In the RM game the payoffs take the form

$$y_1 = \sum_{i=1}^{n}\sum_{j=1}^{n} Y_1(i,j) e_i \otimes e_j \tag{5.6.30}$$

$$y_2 = B \cdot y_1 = \sum_{i=1}^{n}\sum_{j=1}^{n} Y_1(j,i) e_i \otimes e_j, \tag{5.6.31}$$

where B is the permutation matrix that reverses the roles of the players:

$$B : V_1 \times V_2 \to V_1 \times V_2 \tag{5.6.32}$$

$$e_i \otimes e_j \mapsto e_j \otimes e_i \tag{5.6.33}$$

Depending on the values chosen in 5.6.29, there are different kinds of ZDs. Prior to the work of Press and Dyson, Boerlijst et al. (1997a) analyzed the case $\zeta = 0 \implies s_2 = \frac{\theta}{\eta}$, calling this an Equalizing Strategy, before the term ZD was established. An Equalizing Strategy enables a player to unilaterally set the other player's payoff.

Press and Dyson (2012) included the equalizing case in their research, embedded in their more general approach. Another interesting choice of parameters leads to

$$\left.\begin{aligned} \zeta &= 1 \\ \eta &= -\chi \\ \theta &= (\chi - 1) Y_1(n,n) \end{aligned}\right\} \implies (s_1 - Y_1(n,n)) = \chi\left(s_2 - Y_1(n,n)\right), \tag{5.6.34}$$

which we call an Extortionate Strategy. An Extortionate Strategy enables one player to collect a fixed multiple of the other player's payoff margin, where the payoff margin is the difference between the player's payoff and the payoff for the mutual choice of the lowest product.

Extortionate Strategies are particularly effective against players that try to adapt their strategy over time in order to maximize their payoff, e.g. by implementing the evolutionary learning approach presented in Section 5.5. In an environment where the other players follow such a learning scheme, a player using an Extortionate Strategy will be dragged towards its maximum payoff, all the while receiving an unfair share of the total payoff.

5.6.2 Existence of Extortionate Strategies

In the following, we will only consider the case $\chi \geq 1$, since playing an Extortionate Strategy with $\chi < 1$ means that a player is enforcing a worse outcome for himself than for the competitor. In this case, Press and Dyson proved the following existence result in their work on ZDs:

Proposition 1 (Press and Dyson (2012))**.** *In the IPD without observation errors, there exists a unique Extortionate Strategy for every* $\chi \geq 1$*.*

As we will show in Proposition 2, we can prove a similar result for the RM game. Since choosing a product $j < c$ with a higher price than the jointly optimal product c is irrational, we will ignore this possibility and focus on the case $c = 1$ as discussed in this chapter's introduction.

Proposition 2. *In the Markov RM game with* $n \geq 3$*,* $c = 1$ *and* $\varepsilon = 0$*, there exist infinitely many Extortionate Strategies for every* $\chi \geq 1$*.*

Proof. Without loss of generality, we take the perspective of player S_1. Our goal is to construct a single-player strategy $\overline{M}_{\sigma_1}$, so that the determinant of $W(M''_{\sigma_1,\sigma_2}, v)$ in Equation 5.6.25 vanishes for the vector $v = y_1 - \chi y_2 + (\chi - 1)Y_1(n,n)1_{n^2}$, since this enforces the extortionate relationship of both players' payoff in Equation 5.6.34.

For this purpose, we reexamine the construction of $W(M''_{\sigma_1,\sigma_2}, v)$. As Equation 5.6.20 shows, the construction of M''_{σ_1,σ_2} yields n columns that are independent of the competitor S_2's strategy, including the last column. Since in $W(M''_{\sigma_1,\sigma_2}, v)$, the last column of M''_{σ_1,σ_2} is replaced by the vector v, there remain $n-1$ columns in $W(M''_{\sigma_1,\sigma_2}, v)$ under the sole control of player S_1. Thus, in order to enforce the linear relationship 5.6.29, it is sufficient for player S_1 to find a strategy $\overline{M}_{\sigma_1}$, for which parameters $\phi_r \in \mathbb{R}, r = 1, \ldots, n-1$, exist, so that

$$\sum_{r=1}^{n-1} \phi_r \left(\left(\overline{M}_{\sigma_1} \right)_{n(i-1)+j,r} - \delta_{r,i} \right) = v_{n(i-1)+j} \tag{5.6.35}$$

for $i, j = 1, \ldots, n$.

We will show that if we choose

$$\left(\overline{M}_{\sigma_1} \right)_{n(i-1)+j,r} = \frac{1}{2} \qquad \text{if } (r = i) \wedge r \notin \{1, n\} \tag{5.6.36}$$

$$\left(\overline{M}_{\sigma_1} \right)_{n(i-1)+j,r} = 0 \qquad \text{if } r \notin \{1, i, n\} \tag{5.6.37}$$

$$\phi_{n-1} > 6\chi \max\{Y_1(i,j) \mid 1 \leq i, j \leq n\} \tag{5.6.38}$$

$$\phi_r = 2\phi_{r+1} \qquad \text{if } r < n, \tag{5.6.39}$$

then, for each ϕ_{n-1} satisfying Inequality 5.6.38, Equation 5.6.35 yields a strategy $\overline{M}_{\sigma_1}$ that satisfies the linear relationship 5.6.29. For fixed ϕ_{n-1}, this strategy is unique. Since there are infinitely many possible choices for ϕ_{n-1}, this results in an infinite number of Extortionate Strategies.

Due to Equation 5.6.37, each row of $\overline{M}_{\sigma_1}$ contains at most three non-zero entries. Since $\overline{M}_{\sigma_1}$ is a stochastic matrix, so that the entries of each row sum to one, the elements in the last column of $\overline{M}_{\sigma_1}$ can be computed from the other non-zero elements in that row:

$$\left(\overline{M}_{\sigma_1}\right)_{n(i-1)+j,n} = 1 - \left(\left(\overline{M}_{\sigma_1}\right)_{n(i-1)+j,1} + \left(\overline{M}_{\sigma_1}\right)_{n(i-1)+j,i}\right) \quad , i \notin \{1,n\} \tag{5.6.40}$$

$$\left(\overline{M}_{\sigma_1}\right)_{n(i-1)+j,n} = 1 - \left(\overline{M}_{\sigma_1}\right)_{n(i-1)+j,1} \quad , i \in \{1,n\} \tag{5.6.41}$$

Since $\left(\overline{M}_{\sigma_1}\right)_{n(i-1)+j,i}$ is given by Equation 5.6.36 for $i \notin \{1,n\}$, Equations 5.6.40 – 5.6.41 show that we only need to find appropriate values for $\left(\overline{M}_{\sigma_1}\right)_{n(i-1)+j,1}$ in order to unambiguously define a general Markov strategy matrix $\overline{M}_{\sigma_1}$ for player S_1. The entries for the first column of $\overline{M}_{\sigma_1}$ need to satisfy Equation 5.6.35 as well as

$$0 \leq \left(\overline{M}_{\sigma_1}\right)_{n(i-1)+j,1} \leq 1 \qquad \text{if } i \in \{1,n\} \tag{5.6.42}$$

$$0 \leq \left(\overline{M}_{\sigma_1}\right)_{n(i-1)+j,1} \leq \frac{1}{2} \qquad \text{if } i \notin \{1,n\}. \tag{5.6.43}$$

We will prove the existence of such a strategy using the payoff structure of the RM game. In the RM game, the payoff is described by Equations 5.3.3 – 5.3.7, reproduced here for the convenience of the reader:

$$Y_1(i+1,i) > Y_1(i,i) \tag{5.6.44}$$

$$Y_1(i+m,i) > Y_1(i+m,i+m) \tag{5.6.45}$$

$$Y_1(i,i) > Y_1(i+m,i+m) \quad \forall i,m \geq 1 \tag{5.6.46}$$

$$Y_1(i,i+m) < Y_1(n,n) \tag{5.6.47}$$

$$2Y_1(i,i) > Y_1(i,i+m) + Y_1(i+m,i) \quad \forall i,m \geq 1 \tag{5.6.48}$$

As a consequence of these inequalities, we find

$$Y_1(i,i+m) < Y_1(n,n) < Y_1(i+m,i+m) < Y_1(i+m,i) \quad \forall m \geq 1, i \geq c \tag{5.6.49}$$

and therefore

$$Y_1(i,j) < Y_1(j,i) \iff i < j. \tag{5.6.50}$$

Using the payoffs for player S_1 in Equation 5.6.30, we can write

$$v_{n(i-1)+j} = Y_1(i,j) - \chi Y_1(j,i) + (\chi - 1) Y_1(n,n) \tag{5.6.51}$$

for the elements of the extortion vector v and use Inequality 5.6.50 to find

$$v_{n(i-1)+j} < (\chi-1)(Y_1(n,n) - Y_1(j,i)) \quad , i < j, \tag{5.6.52}$$

$$v_{n(i-1)+j} = (\chi-1)(Y_1(n,n) - Y_1(i,i)) \quad , i = j, \tag{5.6.53}$$

$$v_{n(i-1)+j} > (\chi-1)(Y_1(n,n) - Y_1(j,i)) \quad , i > j. \tag{5.6.54}$$

Thus, we can characterize the sign of the elements of v depending on both players' actions

$$v_{n(i-1)+j} \begin{cases} < 0 \text{ if } (i = j < n) \vee (i < j) \\ = 0 \text{ if } i = j = n \\ > 0 \text{ if } i > j \end{cases} . \tag{5.6.55}$$

Note that Equations 5.6.38 and 5.6.51 imply

$$\phi_{n-1} > 2 \max |v_{n(i-1)+j}|, \tag{5.6.56}$$

which will be useful to ensure that the entries of the first column of $\overline{M}_{\sigma_1}$ stay in the desired boundaries.

In the following, we will execute a case-by-case analysis, taking advantage of Equation 5.6.55 to determine the appropriate sign of the elements of the extortion vector v for each case.

- $i = 1$

$$\phi_1 \left(\left(\overline{M}_{\sigma_1} \right)_{j,1} - 1 \right) + \sum_{r=2}^{n-1} \phi_r \left(\overline{M}_{\sigma_1} \right)_{j,r} = v_j \tag{5.6.57}$$

$$\implies \left(\overline{M}_{\sigma_1} \right)_{j,1} = 1 + \frac{v_j}{\phi_1} \tag{5.6.58}$$

  Because of Equation 5.6.55, we know that $v_j < 0$. With Equations 5.6.56 and 5.6.39, it follows that $\left(\overline{M}_{\sigma_1} \right)_{j,1} \in (0, 1)$.

- $i = n$

$$\sum_{r=1}^{n-1} \phi_r \left(\overline{M}_{\sigma_1} \right)_{n(n-1)+j,r} = v_{n(n-1)+j} \tag{5.6.59}$$

$$\implies \left(\overline{M}_{\sigma_1} \right)_{n(n-1)+j,1} = \frac{v_{n(n-1)+j}}{\phi_1} \tag{5.6.60}$$

  1. $i > j$
     Equation 5.6.55 shows that $v_{n(n-1)+j} > 0$. Combining this with Equations 5.6.56 and 5.6.39, we know that $\left(\overline{M}_{\sigma_1} \right)_{n(n-1)+j,1} \in (0, 1)$.

  2. $i = j$
     Because of Equation 5.6.55, we have $v_{n^2} = 0 \implies \left(\overline{M}_{\sigma_1} \right)_{n^2,1} = 0$.

- $1 < i < n$

$$\phi_i \left(\left(\overline{M}_{\sigma_1} \right)_{n(i-1)+j,i} - 1 \right) + \sum_{\substack{r=1 \\ r \neq i}}^{n-1} \phi_r \left(\overline{M}_{\sigma_1} \right)_{n(i-1)+j,r} = v_{n(i-1)+j} \tag{5.6.61}$$

$$\implies \phi_i \left(-\frac{1}{2} \right) + \phi_1 \left(\overline{M}_{\sigma_1} \right)_{n(i-1)+j,1} = v_{n(i-1)+j} \tag{5.6.62}$$

$$\implies \left(\overline{M}_{\sigma_1} \right)_{n(i-1)+j,1} = \frac{v_{n(i-1)+j}}{\phi_1} + \frac{\phi_i}{2\phi_1} \tag{5.6.63}$$

1. $i \leq j$
   Because of Equation 5.6.55, we know that $v_{n(i-1)+j} < 0$. Then, Equation 5.6.39 yields $\left(\overline{M}_{\sigma_1} \right)_{n(n-1)+j,1} < \frac{1}{2}$, while Equations 5.6.56 and 5.6.39 show that $\left(\overline{M}_{\sigma_1} \right)_{n(n-1)+j,1} > 0$. Consequently, we have $\left(\overline{M}_{\sigma_1} \right)_{n(n-1)+j,1} \in (0, \frac{1}{2})$.
2. $i > j$
   Equation 5.6.56 implies $\phi_i > 2v_{n(i-1)+j}$, Equation 5.6.39 implies $\phi_1 \geq 2\phi_i$, therefore $\left(\overline{M}_{\sigma_1} \right)_{n(n-1)+j,1} < \frac{1}{2}$. Since Equation 5.6.55 implies that all the terms on the right hand side are positive, we get the desired result $\left(\overline{M}_{\sigma_1} \right)_{n(n-1)+j,1} \in (0, \frac{1}{2})$.

□

Corollary 1. *In the Markov RM game with $c = 1$ and $\varepsilon = 0$, there exists at least one Extortionate Strategy for every $n \geq 2$ and $\chi \geq 1$.*

Proof. As outlined in Section 5.3, for $n = 2, c = 1$ the RM game reduces to the IPD. Since this case has been covered by Press and Dyson (2012), we only need Proposition 2 to prove this corollary. □

With the introduction of observation errors, the result of Proposition 2 is no longer valid, i.e. we can no longer guarantee the existence of extortionate strategies. This is no coincidence, since in the special case of the IPD with so-called conventional payoff values (see e.g. Axelrod & Hamilton, 1981), we can even show that there cannot exist any extortionate strategies.

Proposition 3. *For any observation error probability $\varepsilon \in (0, \frac{1}{2})$, there cannot exist any Extortionate Strategy for $\chi \geq 1$ in the IPD with conventional payoff values $T = 5, R = 3, P = 1, S = 0$.*

Proof. Without loss of generality, we take the perspective of the first player. First, we have to construct the transition matrix $M^{\varepsilon}_{\sigma_1,\sigma_2}$ of the game with observation errors as described in Section 5.2.2. The general Markov strategies in the IPD take the form

$$\overline{M}_{\sigma_1} = \begin{pmatrix} p_1 & 1-p_1 \\ p_2 & 1-p_2 \\ p_3 & 1-p_3 \\ p_4 & 1-p_4 \end{pmatrix}, \overline{M}_{\sigma_2} = \begin{pmatrix} q_1 & 1-q_1 \\ q_2 & 1-q_2 \\ q_3 & 1-q_3 \\ q_4 & 1-q_4 \end{pmatrix}. \tag{5.6.64}$$

Observation errors are introduced for each player via the error matrix

$$E = \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}, \tag{5.6.65}$$

and the matrix Q provides both players with the input

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.6.66}$$

The basis permutation matrix

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.6.67}$$

reverses the role of the two players, so that the second player can work with a standard general Markov strategy matrix.

This allows us to produce the transition matrix

$$M^{\varepsilon}_{\sigma_1,\sigma_2} = (E \otimes E)\, Q \left(\overline{M}_{\sigma_1} \otimes \left(B\overline{M}_{\sigma_2}\right)\right). \tag{5.6.68}$$

Similarly to the explanation in the case without observation errors in the beginning of this section, we can use the transition matrix as the starting point of the analysis of zero-determinant strategies. The matrix $M^{\varepsilon\prime\prime}_{\sigma_1,\sigma_2}$ is created by adding the first column to the second and third columns of $M^{\varepsilon\prime}_{\sigma_1,\sigma_2} = M^{\varepsilon}_{\sigma_1,\sigma_2} - I_4$.

In the IPD with conventional payoffs, the payoff vectors are

$$y_1 = (R, S, T, P)^t = (3, 0, 5, 1)^t \tag{5.6.69}$$
$$y_2 = (R, T, S, P)^t = (3, 5, 0, 1)^t, \tag{5.6.70}$$

where x^t denotes the transpose of the vector x. In order to find an extortionate strategy, the first player S_1 needs to find a strategy with values $p_i, i = 1, \ldots, 4$, so that the extortion vector

$$v = y_1 - \chi y_2 + P(\chi - 1)1_4 \tag{5.6.71}$$

$$= \begin{pmatrix} 2\,(1-\chi) \\ -4\chi - 1 \\ \chi + 4 \\ 0 \end{pmatrix} \tag{5.6.72}$$

is linearly dependent on the columns of $M^{\varepsilon''}_{\sigma_1,\sigma_2}$. In this case, the determinant of the matrix $W(M^{\varepsilon''}_{\sigma_1,\sigma_2}, v)$ vanishes and the extortionate linear dependence between the payoffs of both players is enforced. However, due to the observation errors, $W(M^{\varepsilon''}_{\sigma_1,\sigma_2}, v)$ is too big to be displayed. Fortunately, we need only the second column of $M^{\varepsilon''}_{\sigma_1,\sigma_2}$, since in the case $n = 2$, this is the only column independent of the second player's strategy. Thus, the first player needs to solve the system of linear equations given by

$$\phi_1 \begin{pmatrix} p_4\varepsilon^2 - p_3\varepsilon^2 - p_2\varepsilon^2 + p_1\varepsilon^2 + p_3\varepsilon + p_2\varepsilon - 2p_1\varepsilon + p_1 - 1 \\ -\left(p_4\varepsilon^2 - p_3\varepsilon^2 - p_2\varepsilon^2 + p_1\varepsilon^2 - p_4\varepsilon + 2p_2\varepsilon - p_1\varepsilon - p_2 + 1\right) \\ -\left(p_4\varepsilon^2 - p_3\varepsilon^2 - p_2\varepsilon^2 + p_1\varepsilon^2 - p_4\varepsilon + 2p_3\varepsilon - p_1\varepsilon - p_3\right) \\ p_4\varepsilon^2 - p_3\varepsilon^2 - p_2\varepsilon^2 + p_1\varepsilon^2 - 2p_4\varepsilon + p_3\varepsilon + p_2\varepsilon + p_4 \end{pmatrix} = \begin{pmatrix} 2\,(1-\chi) \\ -4\chi - 1 \\ \chi + 4 \\ 0 \end{pmatrix} \tag{5.6.73}$$

for a $\phi_1 \neq 0$. Note that the left side is just a multiple of the second column of $M^{\varepsilon''}_{\sigma_1,\sigma_2}$.

The solution of this system of linear equations is given by

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = \frac{1}{(1-2\varepsilon)^2\phi_1} \begin{pmatrix} (1-\varepsilon)(\phi_1(1-2\varepsilon) + (\chi-1)(5\varepsilon-2)) \\ \phi_1(1-\varepsilon)(1-2\varepsilon) - (\chi-1)5\varepsilon^2 + 10\chi\varepsilon - 4\chi - 1 \\ \varepsilon\phi_1(2\varepsilon-1) - (\chi-1)5\varepsilon^2 - 10\varepsilon + \chi + 4 \\ \varepsilon(\phi_1(2\varepsilon-1) + (\chi-1)(3-5\varepsilon)) \end{pmatrix}. \tag{5.6.74}$$

In order to represent a strategy of the game, all entries in the solution vector need to be probabilities. Therefore, we have to check that $p_i \in [0, 1]$ for $i = 1, \ldots, 4$.

Starting the analysis with p_4, we find the necessary and sufficient conditions for ϕ_1

$$\phi_1 \geq \frac{\varepsilon(\chi-1)(3-5\varepsilon)}{(1-\varepsilon)(1-2\varepsilon)} \tag{5.6.75}$$

$$\text{and } \phi_1 \leq \frac{(\chi-1)(3-5\varepsilon)}{1-2\varepsilon}. \tag{5.6.76}$$

Since Equation 5.6.75 implies $\phi_1 \geq 0$ and we know that $\phi_1 \neq 0$, the denominator $\frac{1}{(1-2\varepsilon)^2\phi_1}$ is positive. Consequently, the numerator of any p_i must be non-negative. Using the fact that $\chi \geq 1$ and $\varepsilon < \frac{1}{2}$, we find for the numerator of p_2 that

$$\phi_1 \geq \frac{(\chi-1)5\varepsilon^2 - 10\chi\varepsilon + 4\chi + 1}{(1-2\varepsilon)(1-\varepsilon)} \tag{5.6.77}$$

$$> \frac{(\chi-1)(5\varepsilon^2 - 10\varepsilon + 4)}{(1-2\varepsilon)(1-\varepsilon)} \tag{5.6.78}$$

$$> \frac{(\chi-1)(3-5\varepsilon)}{1-2\varepsilon}. \tag{5.6.79}$$

Equations 5.6.76 and 5.6.79 form a contradiction. □

5.6.3 Example

In their work on the IPD, Press and Dyson (2012) covered the case $n = 2, c = 1$. Our results in Section 5.6.2 demonstrate that the structure of the Extortionate Strategies changes with the amount of actions of each player n. While for any $\chi > 1$, there is a unique Extortionate Strategy in the IPD case of $n = 2$, there are infinitely many Extortionate Strategies for $n \geq 3$.

Similarly to the example in Section 5.4.4, we provide the reader with an example for $n = 3$ and $c = 1$, since this is the smallest choice of parameters, for which any $\chi > 1$ allows for more than one Extortionate Strategies. For $n = 2$, we refer the reader to the work of Press and Dyson, while for $n > 3$, the depiction of the strategy matrices becomes infeasible without adding to the insight for the reader.

For $n = 3$, any general Markov strategy for the first player S_1 can be written as

$$\overline{M}_{\sigma_1} = \begin{pmatrix} p_1 & p_2 & 1-p_1-p_2 \\ p_3 & p_4 & 1-p_3-p_4 \\ p_5 & p_6 & 1-p_5-p_6 \\ p_7 & p_8 & 1-p_7-p_8 \\ p_9 & p_{10} & 1-p_9-p_{10} \\ p_{11} & p_{12} & 1-p_{11}-p_{12} \\ p_{13} & p_{14} & 1-p_{13}-p_{14} \\ p_{15} & p_{16} & 1-p_{15}-p_{16} \\ p_{17} & p_{18} & 1-p_{17}-p_{18} \end{pmatrix}, \tag{5.6.80}$$

where the entries of each row are non-negative and sum to one. Analogously, a general

Markov strategy for player S_2 takes the form

$$\overline{M}_{\sigma_2} = \begin{pmatrix} q_1 & q_2 & 1-q_1-q_2 \\ q_3 & q_4 & 1-q_3-q_4 \\ q_5 & q_6 & 1-q_5-q_6 \\ q_7 & q_8 & 1-q_7-q_8 \\ q_9 & q_{10} & 1-q_9-q_{10} \\ q_{11} & q_{12} & 1-q_{11}-q_{12} \\ q_{13} & q_{14} & 1-q_{13}-q_{14} \\ q_{15} & q_{16} & 1-q_{15}-q_{16} \\ q_{17} & q_{18} & 1-q_{17}-q_{18} \end{pmatrix}, \tag{5.6.81}$$

where the entries of each row are non-negative and sum to one.

The basis change matrix B is

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{5.6.82}$$

while the matrix representation of the forking map

$$Q : V_1 \otimes V_2 \to V_1 \otimes V_2 \otimes V_1 \otimes V_2 \tag{5.6.83}$$

$$e_i \otimes e_j \mapsto e_i \otimes e_j \otimes e_i \otimes e_j \tag{5.6.84}$$

is a 9×81 matrix, which cannot be displayed here. Although the dimensions of the matrix

$$M'_{\sigma_1,\sigma_2} = M_{\sigma_1,\sigma_2} - I_9 \tag{5.6.85}$$

$$= Q \cdot \left(\overline{M}_{\sigma_1} \otimes \left(B \cdot \overline{M}_{\sigma_2}\right)\right) - I_9 \tag{5.6.86}$$

are only 9×9, the complexity of its entries render it too big to be displayed here as well.

However, the simplified matrix M''_{σ_1,σ_2} takes the simple form

$$M''_{\sigma_1,\sigma_2} = M'_{\sigma_1,\sigma_2} \cdot \prod_{i=1}^{n}\prod_{j=1}^{n-1}\left(E^{n(i-1)+j,n(i-1)+n}\right) \cdot \prod_{i=1}^{n}\prod_{j=1}^{n-1}\left(E^{n(j-1)+i,n(n-1)+i}\right) = \tag{5.6.87}$$

$$\begin{pmatrix} p_1q_1-1 & p_1q_2 & p_1-1 & q_1p_2 & p_2q_2 & p_2 & q_1-1 & q_2 & 0 \\ p_3q_7 & p_3q_8-1 & p_3-1 & p_4q_7 & p_4q_8 & p_4 & q_7 & q_8-1 & 0 \\ p_5q_{13} & p_5q_{14} & p_5-1 & p_6q_{13} & p_6q_{14} & p_6 & q_{13} & q_{14} & 0 \\ q_3p_7 & q_4p_7 & p_7 & q_3p_8-1 & q_4p_8 & p_8-1 & q_3-1 & q_4 & 0 \\ p_9q_9 & p_9q_{10} & p_9 & q_9p_{10} & p_{10}q_{10}-1 & p_{10}-1 & q_9 & q_{10}-1 & 0 \\ p_{11}q_{15} & p_{11}q_{16} & p_{11} & p_{12}q_{15} & p_{12}q_{16} & p_{12}-1 & q_{15} & q_{16} & 0 \\ q_5p_{13} & q_6p_{13} & p_{13} & q_5p_{14} & q_6p_{14} & p_{14} & q_5-1 & q_6 & 0 \\ q_{11}p_{15} & q_{12}p_{15} & p_{15} & q_{11}p_{16} & q_{12}p_{16} & p_{16} & q_{11} & q_{12}-1 & 0 \\ p_{17}q_{17} & p_{17}q_{18} & p_{17} & q_{17}p_{18} & p_{18}q_{18} & p_{18} & q_{17} & q_{18} & 0 \end{pmatrix}. \tag{5.6.88}$$

Note that there are only p_i values in the third and sixth columns, and only q_i values in the seventh and eighth columns, so that each of these columns is entirely dependent on a single player. Replacing the last column of M''_{σ_1,σ_2} with a vector v leads to the matrix

$$W(M''_{\sigma_1,\sigma_2}, v) = \tag{5.6.89}$$

$$\begin{pmatrix} p_1q_1-1 & p_1q_2 & p_1-1 & q_1p_2 & p_2q_2 & p_2 & q_1-1 & q_2 & v_1 \\ p_3q_7 & p_3q_8-1 & p_3-1 & p_4q_7 & p_4q_8 & p_4 & q_7 & q_8-1 & v_2 \\ p_5q_{13} & p_5q_{14} & p_5-1 & p_6q_{13} & p_6q_{14} & p_6 & q_{13} & q_{14} & v_3 \\ q_3p_7 & q_4p_7 & p_7 & q_3p_8-1 & q_4p_8 & p_8-1 & q_3-1 & q_4 & v_4 \\ p_9q_9 & p_9q_{10} & p_9 & q_9p_{10} & p_{10}q_{10}-1 & p_{10}-1 & q_9 & q_{10}-1 & v_5 \\ p_{11}q_{15} & p_{11}q_{16} & p_{11} & p_{12}q_{15} & p_{12}q_{16} & p_{12}-1 & q_{15} & q_{16} & v_6 \\ q_5p_{13} & q_6p_{13} & p_{13} & q_5p_{14} & q_6p_{14} & p_{14} & q_5-1 & q_6 & v_7 \\ q_{11}p_{15} & q_{12}p_{15} & p_{15} & q_{11}p_{16} & q_{12}p_{16} & p_{16} & q_{11} & q_{12}-1 & v_8 \\ p_{17}q_{17} & p_{17}q_{18} & p_{17} & q_{17}p_{18} & p_{18}q_{18} & p_{18} & q_{17} & q_{18} & v_9 \end{pmatrix}. \tag{5.6.90}$$

As noted in Equation 5.6.25, the expected payoff of the game can be calculated with the help of the determinant of $W(M''_{\sigma_1,\sigma_2}, v)$ for appropriately chosen v. In order to enforce the linear relationship in Equation 5.6.34 that guarantees the player S_1 an unfair share of the payoff, we need to define the extortion vector v corresponding to the extortionate linear relationship of payoffs.

$$v = y_1 - \chi y_2 + (\chi - 1) Y_1(n,n) 1_{n^2} \tag{5.6.91}$$

$$= \begin{pmatrix} Y_1(1,1) \\ Y_1(1,2) \\ Y_1(1,3) \\ Y_1(2,1) \\ Y_1(2,2) \\ Y_1(2,3) \\ Y_1(3,1) \\ Y_1(3,2) \\ 0 \end{pmatrix} - \chi \begin{pmatrix} Y_1(1,1) \\ Y_1(2,1) \\ Y_1(3,1) \\ Y_1(1,2) \\ Y_1(2,2) \\ Y_1(3,2) \\ Y_1(1,3) \\ Y_1(2,3) \\ 0 \end{pmatrix} + (\chi - 1) \begin{pmatrix} Y_1(n,n) \\ Y_1(n,n) \\ Y_1(n,n) \\ Y_1(n,n) \\ Y_1(n,n) \\ Y_1(n,n) \\ Y_1(n,n) \\ Y_1(n,n) \\ 0 \end{pmatrix} \tag{5.6.92}$$

Then, we need to find a strategy matrix $\overline{M}_{\sigma_1}$ that makes the determinant $\det(W(M''_{\sigma_1,\sigma_2}, v))$ vanish. Following the construction in Section 5.6.2, we obtain the Extortionate Strategy of player S_1

$$\overline{M}_{\sigma_1} = \frac{1}{4\phi_2} \begin{pmatrix} 4\phi_2 + 2v_1 & 0 & -2v_1 \\ 4\phi_2 + 2v_2 & 0 & -2v_2 \\ 4\phi_2 + 2v_3 & 0 & -2v_3 \\ \phi_2 + 2v_4 & 2\phi_2 & 3\phi_2 - 2v_4 \\ \phi_2 + 2v_5 & 2\phi_2 & 3\phi_2 - 2v_5 \\ \phi_2 + 2v_6 & 2\phi_2 & 3\phi_2 - 2v_6 \\ 2v_7 & 0 & 4\phi_2 - 2v_7 \\ 2v_8 & 0 & 4\phi_2 - 2v_8 \\ 0 & 0 & 4\phi_2 \end{pmatrix} \tag{5.6.93}$$

with $\phi_2 > 6\chi \max\{Y_1(i,j) \mid 1 \leq i,j \leq n\}$.

5.7 Dynamic Pricing

In the preceding sections of Chapter 5, we have described the RM game between two service providers selling a resource at n price points as a Markov process. In this section, we will show that a similar analysis is possible for continuous prices, which transfers the problem to Dynamic Pricing. We will derive the transition kernels in a similar way as we did in Section 5.2 for the discrete-price case. Then, we will show how the strategies described in Section 5.4.1 can be formulated in this framework.

For this purpose, let each service provider S_k offer a single resource at a price $f_k \in F_k = \mathbb{R}^+$ at every stage $s \in \mathbb{N}_0$. Let f_c denote the jointly optimal price in this game and thus the equivalent of the price of the cooperative threshold product in the case with discrete products. At each stage, each player may base his pricing decision on the actions of both players in the previous stage.

The resulting stochastic process is a Markov process with a continuous state space. In such a state space, it is no longer possible to describe the transition probabilities of the process with the help of transition matrices as in the discrete case. As discussed in Section 5.1, a Markov process on a general measurable space $(\Omega, \mathcal{F})$ is described by a transition kernel K, i.e. a map on $(\Omega, \mathcal{F})$, for which

- $K(x, \cdot)$ is a probability measure for all $x \in \Omega$ and
- $K(\cdot, A)$ is $\mathcal{F}$-measurable for all $A \in \mathcal{F}$.

In words, $K(x, A)$ denotes the probability that the state of the Markov process X at stage $s+1$ is included in the set $A \in \mathcal{F}$ given that the state at stage s was $X_s = x$. The transition kernel can be used to find the probability measure μ_{s+1} of the next stage given the current probability measure μ_s as $\mu_{s+1}(A) = \int\limits_{\Omega} K(x, A)\mu_s(dx)$ for $A \in \mathcal{F}$. This describes the evolution of the process.

5.7.1 Construction of the Markov Process

The state space of the process is $F_1 \times F_2$ equipped with the σ-Algebra $\mathcal{F}_1 \otimes \mathcal{F}_2$, where $\mathcal{F}_k = \mathcal{B}(F_k) = \mathcal{B}(\mathbb{R}^+)$ denotes the Borel-σ-Algebra of F_k.

At any stage s, the measure μ_s yields the joint probability distribution of both players' prices, while the marginal measure ν_s^k, $l \neq k$, yields the marginal distribution of player S_k's prices. The marginal measure can be found by integrating over the competitor's prices:

$$\nu_s^1 = \int\limits_{F_2} \mu_s(\cdot, df_2) \tag{5.7.1}$$

$$\nu_s^2 = \int\limits_{F_1} \mu_s(df_1, \cdot) \tag{5.7.2}$$

Similarly to the previous sections, we differentiate between general and reactive Markov strategies. A reactive strategy depends only on the competitor's previous action, whereas a general Markov strategy can depend on both player's previous actions.

General Markov Strategies Any Markov strategy σ_k of player S_k can be represented by a general Markov kernel $\overline{K}_{\sigma_k}$, which can depend on both players' previous actions. Since S_k can only set his own prices, we can only deduct the probability measure of S_k's prices at stage $s+1$ given both players' prices at stage s as

$$\mu_{s+1}(A \times F_2) = \int_{F_1 \times F_2} \overline{K}_{\sigma_k}((f_1, f_2), A \times F_2)\mu_s(df_1 df_2) \quad \text{if } k = 1 \tag{5.7.3}$$

$$\mu_{s+1}(F_1 \times A) = \int_{F_1 \times F_2} \overline{K}_{\sigma_k}((f_1, f_2), F_1 \times A)\mu_s(df_1 df_2) \quad \text{if } k = 2 \tag{5.7.4}$$

for any $A \in \mathcal{F}_k$.

In order to describe the evolution of the game's Markov process, we need to account for the interaction of both players' strategies. We find that the restriction of each player's transition kernel to the Cartesian products $A \times F_2$ and $F_1 \times A$ is sufficient for the creation of a combined transition kernel.

Lemma 1. *The product $\widetilde{K}$ of both players' general Markov strategy transition kernels on the sets $A \times F_2$ and $B \times F_1$ with $A \in F_1, B \in F_2$*

$$\overline{K}_{\sigma_1}((f_1, f_2), A \times F_2)\overline{K}_{\sigma_2}((f_1, f_2), F_1 \times B) =: \widetilde{K}((f_1, f_2), A \times B) \tag{5.7.5}$$

generates another transition kernel describing the effect of both players' strategies.

Proof. We have to check the properties of a transition kernel described in Section 5.1.1 and recapped in the beginning of this section.

- The marginalizations $\overline{K}_{\sigma_1}|_{F_1}(\cdot) = \overline{K}_{\sigma_1}(\cdot \times F_2)$ and $\overline{K}_{\sigma_2}|_{F_2}(\cdot) = \overline{K}_{\sigma_2}(F_1 \times \cdot)$ of $\overline{K}_{\sigma_k}$ to $(F_k, \mathcal{F}_k)$ are marginal probability measures on $(F_k, \mathcal{F}_k)$. Consequently, $\widetilde{K}$ is a probability measure on $(F_1 \times F_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ for every (f_1, f_2), since it is the product measure of $\overline{K}_{\sigma_1}|_{F_1}$ and $\overline{K}_{\sigma_2}|_{F_2}$.
- $\widetilde{K}$ is $\mathcal{F}_1 \otimes \mathcal{F}_2$-measurable for every $A \times B$, since it is a product of two $\mathcal{F}_1 \otimes \mathcal{F}_2$-measurable functions.

□

It is thus only necessary for our purposes to define each player's transition kernel $\overline{K}_{\sigma_k}$ on all sets $A \times F_2$ and $B \times F_1$ with $A \in F_1, B \in F_2$. Therefore, the evolution of the game's Markov process is described by

$$\mu_{s+1}(A \times B) = \int_{F_1 \times F_2} \widetilde{K}((f_1, f_2), A \times B)\mu_s(df_1 df_2) \tag{5.7.6}$$

$$= \int_{F_1 \times F_2} \overline{K}_{\sigma_1}((f_1, f_2), A)\overline{K}_{\sigma_2}((f_1, f_2), B)\mu_s(df_1 df_2) \tag{5.7.7}$$

for $A \in \mathcal{F}_1, B \in \mathcal{F}_2$. Here, we have used the shorthand notation $\overline{K}_{\sigma_k}((f_1, f_2), A)$ meaning $\overline{K}_{\sigma_1}((f_1, f_2), A \times F_2)$ if $k = 1$ or $\overline{K}_{\sigma_2}((f_1, f_2), F_1 \times A)$ if $k = 2$. For brevity's sake we will continue to use this notation throughout this section.

Reactive Strategies As in the previous sections, reactive strategies take a simpler form, since each player can only react to his competitor's price. Thus, reactive strategies use the marginal measures $\nu_s^k(df_k) = \int_{F_l} \mu_s(df_1 df_2)$, $k \neq l$, which describe the probability measure of S_k's prices only. In this case, the transition kernel $K(f_l, A)$ with $A \in \mathcal{F}_k$ is a probability measure on F_k and measurable for f_l. For player S_k, the probability measure of his prices in stage $s+1$ can be calculated using the competitor price $f_l \in F_l$ at stage s as

$$\nu_{s+1}^k(A) = \int_{F_l} K_{\sigma_k}(f_l, A) \nu_s^l(df_l) \tag{5.7.8}$$

for $A \in \mathcal{F}_k$ and $k \neq l$. If both players use reactive strategies, the Markov process of the game evolves via

$$\mu_{s+1}(A \times B) = \int_{F_1 \times F_2} K_{\sigma_1}(f_2, A) K_{\sigma_2}(f_1, B) \nu_s^1(df_1) \nu_s^2(df_2) \tag{5.7.9}$$

for $A \in \mathcal{F}_1, B \in \mathcal{F}_2$.

Observation Errors Since observation errors occur independently of the other player's actions, errors can be defined via a transition kernel E on $(F_k, \mathcal{F}_k)$, so that for $A \in \mathcal{F}_k$, we find for the probability measure of erroneous observations of S_k's prices:

$$\nu_s^{k,\varepsilon}(A) = \int_{F_k} E(f_k, A) \nu_s^k(df_k) \tag{5.7.10}$$

With $A \in \mathcal{F}_1, B \in \mathcal{F}_2$, it follows that the occurrence of observation errors on the space of both players' prices $F_1 \times F_2$ is described by

$$\mu_s^\varepsilon(A \times B) = \int_{F_1 \times F_2} E(f_1, A) E(f_2, B) \mu_s(df_1 df_2). \tag{5.7.11}$$

Similarly to the discrete case, observation errors occur, before both players then apply their strategy simultaneously. Thus, the transition kernel of the Dynamic Pricing process is generated by the composition of error and strategy transition kernels

$$\overline{K}_{\sigma_1}\left((f_1', f_2'), A\right) \overline{K}_{\sigma_2}\left((f_1', f_2'), B\right) E\left(f_1, df_1'\right) E\left(f_2, df_2'\right), \tag{5.7.12}$$

so that the evolution of the game can be described by

$$\mu_{s+1}(A \times B) = \int\limits_{F_1 \times F_2} \overline{K}_{\sigma_1}\left((f_1', f_2'), A\right) \overline{K}_{\sigma_2}\left((f_1', f_2'), B\right) E\left(f_1, df_1'\right) E\left(f_2, df_2'\right) \mu_s(df_1 df_2) \tag{5.7.13}$$

for $A \in \mathcal{F}_1, B \in \mathcal{F}_2$.

Reputation As described in Section 4.2.1, the reputation r_k of service provider S_k takes values in $\{g, b\}$, where $r_k = g$ denotes a good and $r_k = b$ a bad reputation. The reputation evolves according to the transition kernel R on the space $(F_1 \times F_2 \times P_1 \times P_2, \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes \mathcal{P}_1 \otimes \mathcal{P}_2)$, where $\mathcal{P}_k$ is the power set of P_k.

By abuse of notation, we keep the same notation for the strategies, although the state space has changed. We can use a composition of transition kernels to describe the evolution of the probability measure induced by the process without errors:

$$\begin{aligned}\mu_{s+1}(A \times B \times \{(r_1', r_2')\}) = \int\limits_{F_1 \times F_2} & \overline{K}_{\sigma_1}\left((f_1, f_2, r_1', r_2'), A\right) \overline{K}_{\sigma_2}\left((f_1, f_2, r_1', r_2'), B\right) \\ & \cdot R((f_1, f_2, r_1, r_2), (f_1, f_2, r_1', r_2')) \\ & \cdot \mu_s(df_1 df_2 dr_1 dr_2)\end{aligned} \tag{5.7.14}$$

To incorporate observation errors, we only have to modify the input prices for the reputation and for the strategies using the error transition kernels E:

$$\begin{aligned}\mu_{s+1}(A \times B \times \{(r_1', r_2')\}) = \int\limits_{F_1 \times F_2} & \overline{K}_{\sigma_1}\left((f_1', f_2', r_1', r_2'), A\right) \overline{K}_{\sigma_2}\left((f_1', f_2', r_1', r_2'), B\right) \\ & \cdot R((f_1', f_2', r_1, r_2), (f_1', f_2', r_1', r_2')) \\ & \cdot E(f_1, df_1')E(f_2, df_2')\mu_s(df_1 df_2 dr_1 dr_2)\end{aligned} \tag{5.7.15}$$

5.7.2 Strategies

In order to formulate the equivalent strategies to the discrete case, we rely on classification of actions and observations as cooperating or defecting as in Section 4.2.1, so that the strategies can be transferred from the IPD. Similarly to the discrete case, we say that a provider S_k plays DEFECT, if he chooses the single-stage best response $f^\star(f_l)$ to his competitor's price f_l. We say that S_k plays COOP, if he chooses the jointly optimal price f_c. In order to implement the mechanism of IPD strategies, it is also necessary that the observations of players' prices are classified as either cooperative or non-cooperative behavior. Thankfully, we can use the same logic as in the discrete case. A competitor

S_l is observed to play DEFECT, if his—possibly erroneous—price is lower than f_c , otherwise he is interpreted as playing COOP.

With this classification, the strategies from Section 5.4.1 can be written as linear combinations of $\delta_{f_c}(A)$ and $\delta_{f^\star(f_l)}(A)$, where

$$\delta_x(A) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ else} \end{cases} \tag{5.7.16}$$

is the Dirac measure, f_c is the jointly optimal price and $f^\star(f_l)$ is the single-stage revenue-maximizing price given a competitor price f_l. The value of $f^\star(f_l)$ depends on the demand model and can be hard to find. However, we stress that our treatment of the problem focuses on the repeated-game aspect of the Dynamic Pricing problem. In fact, our model may be used in combination with other approaches providing the solution to the single-stage problem, for which we refer the reader to our literature review in Chapter 2.

In the following, we will provide the transition kernels for player S_k for the set of strategies presented in Section 5.4.1.

ALLD For $A \in \mathcal{F}_k$, player S_k follows the reactive strategy ALLD by choosing the single-stage non-cooperative best response

$$K_{ALLD}(f_l, A) = \delta_{f^\star(f_l)}(A) \quad \forall A \in \mathcal{F}_k. \tag{5.7.17}$$

ALLC

$$K_{ALLC}(f_l, A) = \delta_{f_c}(A) \quad \forall A \in \mathcal{F}_k \tag{5.7.18}$$

Note that the strategy of pure cooperation ALLC is independent of the competitor's last action. Nevertheless, we choose to model ALLC as a reactive strategy, since this is the simplest representation of a strategy in our framework.

TFT

$$K_{TFT}(f_l, A) = \delta_{f_c}(A)\delta_{f_l}((0, f_c]) + \delta_{f^\star(f_l)}(A)\delta_{f_l}((f_c, \infty)) \quad \forall A \in \mathcal{F}_k \tag{5.7.19}$$

In this reactive strategy, the first summand denotes playing COOP as reaction to a cooperating competitor, while the second summand means playing DEFECT against a defecting competitor.

GTFT For $A \in \mathcal{F}_k$, player S_k follows the general strategy GTFT by implementing the transition kernel

$$K_{GTFT}(f_l, A) = \delta_{f_c}(A)\delta_{f_l}((0, f_c]) + \Big(\gamma\delta_{f_c}(A) + (1-\gamma)\delta_{f^\star(f_l)}(A)\Big)\,\delta_{f_l}((f_c, \infty)). \quad (5.7.20)$$

Here, the first summand amounts to playing COOP as reaction to competitor doing likewise. The second summand describes the reaction to the competitor playing DEFECT. With generosity probability γ, player S_k will react to the opponent's DEFECT with cooperation.

CTFT This general Markov strategy relies on players' reputation. For $A \in \mathcal{F}_k$, player S_k follows CTFT by choosing the transition kernel

$$\begin{aligned}\overline{K}_{CTFT}((f_1, f_2, r_1, r_2), A) =& \delta_{f_c}(A)\,(1 - \delta_{r_k}(\{b\})\delta_{r_l}(\{g\})) \quad (5.7.21)\\ &+ \delta_{f^\star(f_l)}(A)\delta_{r_k}(\{b\})\delta_{r_l}(\{g\}). \quad (5.7.22)\end{aligned}$$

The first summand stands for cooperation, whereas the second summand denotes playing DEFECT if the player has a bad and his competitor a good reputation.

PAVLOV The transition kernel for the PAVLOV strategy of S_k is

$$\overline{K}_{PAVLOV}((f_1, f_2), A) = \delta_{f_c}(A)\theta(f_1, f_2) + \delta_{f^\star(f_l)}(A)\,(1 - \theta(f_1, f_2)) \quad \forall A \in \mathcal{F}_k, \quad (5.7.23)$$

where we have introduced

$$\theta(f_1, f_2) = \delta_{f_1}((0, f_c])\delta_{f_2}((0, f_c]) + \delta_{f_1}((f_c, \infty))\delta_{f_2}((f_c, \infty)) \quad (5.7.24)$$

as shorthand notation for the case that the players have either both cooperated or both defected at the previous stage. The strategy then reduces to playing COOP if both have acted similarly at the previous stage, as denoted by the first summand, or playing DEFECT otherwise, as denoted by the second summand.

MATCH Player S_k follows the reactive strategy MATCH by using the transition kernel

$$K_{MATCH}(f_l, A) = \delta_{f_l}(A) \quad \forall A \in \mathcal{F}_k. \quad (5.7.25)$$

5.8 Summary

In this chapter, we have studied a simplification of the iterated RM game described in Chapter 4. The main difference to the previous chapter consisted in dropping the capacity constraint, which helped simplify each player's single-stage best response. As shown in the previous chapter, where we investigated competitive effects depending on the capacity constraints, this is equivalent to a focus on the point of maximal competitive effects. Furthermore, we limited each player's memory to a single-stage, so that the game can be described as a Markov process. These modifications enabled us to provide a mathematical analysis of the iterated game.

As outlined in Section 3, we started the analysis of this game in order to find answers to research questions 2 and 5 – 7. Thus, we were interested in the competitive behavior of strategies in the repeated RM game, using different measures for success and highlighting the effect of observation errors. Additionally, we wanted to investigate the existence of Extortionate Strategies.

In the remainder of this section, we will recapitulate our procedure and summarize our findings in the search for answers of this chapter's research questions.

After providing a motivation for the simplified game and outlining the underlying assumptions, we used Section 5.1 to introduce the mathematical background needed for the investigations in this chapter.

In Section 5.2, we examined the interactions of the game's main components such as each player's strategy, observation errors and the evolution of players' reputation. As a result, we found that the transition matrices of the game can be derived using simpler lower-dimensional matrices. Since, in this section, the payoff structure is not yet defined, this result is applicable to any 2-player Markov game with an n-dimensional strategy space.

The connection to the IPD was established in Section 5.3, where we described the payoff structure of the RM game by generalizing the IPD's defining payoff inequalities to the context of an n-dimensional strategy space. In particular, we showed that the IPD is a special case of the Markov RM game, arising if each player is restricted to two products, of which the more expensive one leads to the joint optimum. We showed that the RM game's desired payoff structure occurs as a natural consequence of myopic pricing for a host of demand models.

In Section 5.4, we applied the insights into the structure of the game gained in Section 5.2 to discuss the success of strategies in the iterated RM game in order to answer our research question 1.

Similarly to our procedure in the previous Chapter 4, we first adapted strategies from the IPD to the RM framework in Section 5.4.1. The strategies under consideration varied in their complexity: from repetitions of single-stage strategies as discussed in the literature (Isler & Imhof, 2008, 2010) via requiring only knowledge of the competitor's previous actions through depending on both players' actions to using not only players' previous actions, but also their reputations.

Then, in Section 5.4.2 we presented analytical results about the long-term behavior of selected strategy pairs. In particular, we were interested in the change in behavior induced by observation errors. Without observation errors, we found precise results for the stationary behavior of all considered strategy pairs. However, with observation errors, we have only been able to find the support of the stationary measure on the strategy space for some strategy pairs. For many of the considered strategies, we discovered stark shifts in behavior induced by the possibility of observation errors, that could go as far as change the stationary outcome from fully cooperative to fully non-cooperative for TFT.

In Section 5.4.3, we challenged the assumption that players know the price level for tacit collusion that helps achieve the jointly optimal payoff. We concentrated on the strategies TFT and PAVLOV, since—due to their success (Nowak & Sigmund, 1992, 1993) and prevalence among human players (Wedekind & Milinski, 1996) in the IPD—they served as archetypes for the mechanisms of the other strategies presented in Section 5.4.1. We found that the changes in the stationary measures depended significantly on the type of error associated with the cooperative threshold product. If both players used a wrong, but equal, price level, the effect was by far not as dangerous as if players used differing price levels. In real life, service providers trying to implement a strategy from Section 5.4.1 need to figure out the product level to use as the cooperate threshold product. Our results indicate that players should not try to estimate the correct price level, since this can easily lead to differing cooperative threshold products among the players and consequently, to deteriorating revenues. Instead, players should agree with each other on a common price level, even if this means settling for an imperfect cooperative threshold product. Players can feel safe in this communication, since the ruinous consequences of differing price levels imply that neither player has an incentive to be dishonest.

Section 5.4.4 was used to provide examples to illustrate the structure of the game laid out in Sections 5.2 and 5.4.2. We presented all the necessary ingredients for a study of the game's stationary behavior, including matrix representations of the game's constituent components like the single-player strategies, errors and reputation, as well as the resulting transition matrices. In the following, we performed an analysis of the stationary measures, where we not only reproduced results from Section 5.4.2, but also were able to specify stationary measures with observation errors that had evaded us in Section 5.4.2.

The following Section 5.5 was dedicated to an evolutionary approach to the game. In this section, players could change their strategies over time to adopt more successful strategies. In this section, we evaluated strategies' success in a round-robin tournament, so that each strategy interacts with all other strategies at some point in time, whereas in the previous

Sections 5.4.2 and 5.4.4 we could only examine a small subset of possible matchups. In order to carry out this tournament, we created a simulator that used our results from the previous sections to determine the stationary behavior of interacting strategies. Thus, for a fixed starting distribution and simulation parameters, this approach yielded an optimal strategy among all strategies presented in Section 5.4.1 and consequently, an answer to research question 6. Depending on the parameters chosen as well as the pricing structure, we found different strategies prevailing. If the pricing structure rendered cooperation relatively attractive, the latter stages of the evolutionary game were dominated either by the purely cooperative ALLC or by PAVLOV, which was able to exploit such a cooperative stance. On the other hand, a pricing structure favoring aggressive behavior led to resounding success of the robust Tit for Tat variations GTFT and CTFT, although in rare cases a mixture of PAVLOV and MATCH could prevail.

While the evolutionary approach provided us with optimal strategies in any given environment, Section 5.6 presented us with a class of strategies that are able to exploit evolutionary learning. Building on the work of Press and Dyson (2012), we described Zero-Determinant Strategies, and as a special case Extortionate Strategies, in the context of the RM game. Zero-Determinant Strategies are the class of strategies that allows a player to enforce a linear relationship between both players' payoffs, while Extortionate Strategies specify this to receive an unfair share of the payoffs. As an answer to research question 7, we generalized Press and Dyson's existence result for Extortionate Strategies from the IPD to the Markov RM game. We found that there exists not a unique, but an infinite amount of Extortionate Strategies, as long as each service provider offers more than two products. However, investigating research question 5, we could demonstrate that such a result does not hold anymore, as soon as observation errors are introduced into the framework. We showed that in this case, even the special case of the IPD with conventional payoffs does not allow for the existence of any Extortionate Strategies. Consequently, in a realistic setting with the possibility of observation errors, evolutionary learning as in Section 5.5 can be implemented without having to fear being systematically exploited. This stresses the importance of the strategies introduced in Section 5.4.1, since strategies that are able to beat the system in the sense of Zero-Determinant Strategies do not exist in general.

Finally, in Section 5.7, we concentrated on applying our methods to Dynamic Pricing. We demonstrated that a similar analysis as the one we have carried out in this chapter for the RM case with finite products can be executed with continuous prices. In particular, we provided a derivation of transition kernels similar to Section 5.2 and adapted the strategies from Section 5.4.1.

6 Conclusion

In this chapter, we will complete this thesis by reviewing our contributions and discussing their implications for the future in practice and research. First, in Section 6.1, we will give a summary of this thesis's motivation, targets, solution approaches and results. Then, we will discuss the practical implications of our results in Section 6.2. Finally, we will focus on the limitations of our work and discuss opportunities for future research in Section 6.3.

6.1 Summary

In Chapter 1, we demonstrated the importance of iterated competitive interactions of Revenue Management (RM)-using service providers and outlined the difficulties treating this problem as well as the shortcomings of current approaches. Subsequently, we reviewed the literature on RM, game theory and simulations in Chapter 2. We found that hardly any research had been conducted that treated the problem of RM under competition from a multi-stage perspective. Consequently, we detailed this research gap in Chapter 3 and formulated research questions, which guided our investigations in the following chapters. In the remainder of this section, we will recapitulate the research questions posed in Chapter 3 and discuss our corresponding results.

Research Question 1. *How can strategies of the Iterated Prisoner's Dilemma (IPD) be adapted to the repeated RM game?*

In Chapter 4, we formulated a mathematical model of the repeated interaction of two service providers using RM based on an forecasting engine and an optimizer. We could demonstrate that this repeated game resembled the far simpler, but well-studied IPD. Addressing the research question 1, we developed a heuristic to transfer strategies from the IPD to the repeated RM game based on the work of Isler and Imhof (2008, 2010).

This heuristic allowed us to introduce a host of strategies into the repeated RM game:

- a natural consequence of using a standard RM system: the purely non-cooperative strategy ALLD
- the cooperative strategy: ALLC
- the most successful strategies from the IPD:

  - Tit for Tat (TFT)
  - Generous Tit for Tat (GTFT)
  - Contrite Tit for Tat (CTFT)
  - PAVLOV
- irrational strategies based on widely-used practice of price matching:
  - MATCH
  - UNDER

Research Question 2. *Which strategy leads to a jointly optimal Nash equilibrium in the repeated RM game?*

Simulations using the revenue management simulator REvenue Management Training for Experts (REMATE) using an idealized forecasting method showed that the heuristic yielded several candidates for a strategy forming a jointly optimal Nash equilibrium. Note that we could not prove that a strategy undoubtedly leads to a Nash equilibrium in this chapter, since simulating a strategy against the infinite amount of possible strategies was infeasible. Instead, we relied on testing a set of necessary conditions that rendered a strategy a plausible candidate for a jointly optimal Nash equilibrium. However, we found that the introduction of observation errors had a great effect on the performance of some strategies such as TFT, while others such as GTFT, CTFT and PAVLOV appeared robust against erroneous observations. The PAVLOV strategy could even profit from observation errors, as they helped in exploiting blind cooperation. Nevertheless, even in the presence of observation errors, we found a candidate strategy in CTFT fulfilling all our necessary conditions to lead to a jointly optimal Nash equilibrium.

In this chapter, we could replicate Isler and Imhof's (2008) observation that the effect of competition depends heavily on the severity of the capacity constraints. For severe capacity constraints, the cooperative and non-cooperative solutions are very similar and resemble the monopoly outcome. On the other hand, for vanishing capacity restrictions, the competitive effects completely dominate any classic RM decisions.

Research Question 3. *How closely do standard RM methods approximate the game-theoretic solutions?*

Since Cooper et al. (2009) showed that realistic forecasting methods do not necessarily lead to optimal competitive solutions, we examined whether our results obtained by an idealized forecast could be approximated using a standard forecasting method. For this purpose we described a variety of forecasting techniques, from simple methods, that are widely-used in the industry, to sophisticated approaches, that are not yet used in practice. Similarly to Cooper et al., we found that not all forecasts exhibit a similar behavior to the idealized forecast suggested by Isler and Imhof (2008). However, we found that a forecast using an appropriately complex demand model approximated the results of the idealized forecast reasonably well. Thus, our results stressed the well-known dependence of RM on the forecast quality.

Research Question 4. *Which strategies are best suited to react to simple irrational strategies like price-matching or underpricing?*

Both for idealized and realistic forecasts, we found that the key to success against the irrational strategies MATCH and UNDER was a mixture of cooperation and non-cooperation as well as robustness against observation errors.

Investigating research questions 6 and 7 turned out infeasible by means of simulation using the formulation of the RM game developed in Chapter 4. Thus, in Chapter 5, we studied a simplification of the repeated RM game that was based on the observation that the capacity restriction inhibits the effects of competition in the game. Dropping the capacity constraint and restricting strategies to a one-stage memory enabled the description of the game's progress as a Markov process. The resulting game was a strict generalization of the IPD with one-stage memory strategies, which involves no loss of generality compared to the IPD with arbitrary strategies (Press & Dyson, 2012).

Research Question 5. *What is the effect of observation errors on strategies in the repeated RM game?*

In Chapter 5, we developed a framework to describe general symmetric 2-player Markov games, which we used to analyze the stationary behavior of the simplified RM game. Within this framework, we could investigate the effect of errors on the game via mathematical analysis, which gave us a better understanding of the effect of the game's parameters. In particular, we could show that non-robust strategies collapse even for the lowest probabilities of an observation error. Furthermore, we extended the analysis to demonstrate that errors in the calculation of the optimal cooperative price level only have a low impact on the long-term behavior as long as both players employ the same price level.

Research Question 6. *Which repeated-game strategy is competitively robust in the sense that it fares best against a diverse set of competing strategies?*

Using our findings about the stationary behavior of the strategies in question, we devised an evolutionary simulation similar to e.g. Wu and Axelrod (1995) and Nowak and Sigmund (1993), that helped answer research question 6. Our results showed that the optimal strategy against a diverse set of competitor strategies depends heavily on the demand model and on the players' pricing structure.

Research Question 7. *Do Extortionate Strategies exist in the repeated RM game?*

Continuing our research on this topic, we generalized the results of Press and Dyson (2012), who found that the IPD allows for strategies that can unilaterally exploit evolutionary players. We proved that these strategies exist in the simplified RM game as well. However, we detected the importance of observation errors for the existence of Extortionate Strategies, as we could show that in the presence of errors, such a result does not even hold in the IPD.

6.2 Practical Implications

Currently, state of the art RM systems do not explicitly incorporate competitive effects in their models and thus cannot react adequately to competitors' behavior (Martínez-de Albéniz & Talluri, 2011). Instead, service providers rely on the input of human analysts to deal with the effects of competition (Zeni, 2003). As Isler and Imhof (2008) pointed out, automating these competitive decisions in RM can lead to the Competitive Spiral Down, so that Isler and Imhof concluded:

> Pricing, and specifically the necessary long-term components of the pricing strategy, cannot be automated entirely. When using automated systems along the lines described above, it is important to observe competitor behavior and use overrides of the system according to a defined long-term strategy.

This is based on the lack of research on repeated interactions between RM service providers, as pointed out in Chapter 2. However, we can challenge this statement, since we explored this research gap and presented a range of strategies for the repeated RM game that avoid the negative effects which Isler and Imhof warned about. Thus, creating an automated system using a successful strategy of the repeated game seems possible. On the other hand, our results can also be used as guidance to define a long-term strategy for the human analysts. Due to the simplicity of our solutions, these strategies can help human analysts deal with complex competitive interactions.

As a consequence of our analysis of repeated game strategies using realistic forecasting systems, we stress once more the importance of the forecast for the performance of a RM system. In order to profit from a repeated game strategy, it is necessary to use an appropriate forecast, which produces high-quality estimates for a sufficiently complex demand model. Otherwise, even the most successful repeated game strategy will suffer from the ruinous Competitive Spiral Down or the even worse Spiral Down.

When investigating the interaction with irrational strategies like price matching and underpricing, we found that irrational strategies do not call for different repeated game strategies. Thus, service providers need not to worry specifically about irrational competitors when implementing a repeated game strategy.

Most of the strategies under considerations relied on a heuristic that classified competitor behavior as cooperating if the lowest competitor price did not undercut a fixed threshold price. We found that the effects of an erroneous threshold price can differ greatly: While both firms making the same mistake determining the price level only slightly affects revenues, using different cooperative price levels can lead to the Competitive Spiral Down even for successful strategies of the repeated game. Thus, when determining this cooperative price level, firms should rely on signaling rather than estimating.

Furthermore, we examined the recently discovered class of Extortionate Strategies, which can unilaterally enforce unfair relationships between players in a two-player game (Press & Dyson, 2012). Although we generalized the existence result of Press and Dyson to a RM context, we found that the mere possibility of observation errors renders this result invalid. Since in a realistic environment, observation errors cannot be completely prevented, we conclude that practitioners do not have to worry about Extortionate Strategies.

6.3 Limitations and Future Research

In this thesis, we focused on a limited set of strategies in our analysis of the repeated RM game. Based on the premise that the repeated RM game resembles the IPD, we adapted successful strategies from the IPD and combined this with well-known simple competitive RM strategies like price matching. Although we tried to cover a broad range of strategies and include the most promising approaches, the inclusion of further strategies might bring new results. For example, the Prudent Pavlov strategy by Boerlijst et al. (1997b) mentioned in the literature review in Section 2.2.2 is robust against perception errors, which we have not included in our work. Similarly, the fault-tolerant non-Markov strategies developed by Pelc (2010) possess an interesting robustness against errors. Furthermore, researchers may want to use a different heuristic, which would change the structure of the transferred strategies.

Similarly, we concentrated on a limited set of forecasting techniques when investigating the ability of real-world forecasting techniques to be used in repeated game strategies. While these forecasts were chosen as to represent a large variety of methods popular in practice and research, enlarging the scope of methods could help produce better results. Since our investigations have shown the need for a sufficiently complex demand model, the most promising approaches rely on discrete choice models as discussed in the literature review in Section 2.1.1.

Furthermore, the RM system described in this thesis employed models for forecasting and optimization that relied on a nesting of products by price. However, Gayle (2004) showed that focusing only on price as characteristic might not describe customer behavior accurately.

In order to enhance the clarity of our insights, our model of repeated interaction between service providers was condensed to a symmetric two-player game between two firms selling a single resource. As for every model, this was a simplification that allows for various extensions.

As shown in the IPD, the extension to more than two players can change the dynamics of the game profoundly (Berkemer, 2006).

Furthermore, if firms sell multiple resources, there are more opportunities to enforce cooperative behavior by retaliating on a different resource. As Evans and Kessides (1994) showed, such a behavior is already used in the industry for the practice of pricing, although we do not know of a similar account for the practice of RM.

Challenging the symmetric setup can lead to different results as well. For the single-stage RM game, Martínez-de Albéniz and Talluri (2011) showed that capacity inequalities between the firms induce a specific behavior in the Nash equilibrium, that could not be observed in a symmetric game.

In our model of competition, we have ignored the possibility of entry or exit in a market. However, such effects can be caused or at least influenced by the competitors' RM strategy (Harris, 2007). On the other hand, the entry or exit of competitors has been shown to affect competitive behavior (Eliashberg & Jeuland, 1986; Gorin & Belobaba, 2008).

References

Akçay, Y., Natarajan, H., & Xu, S. (2010). Joint Dynamic Pricing of Multiple Perishable Products Under Consumer Choice. *Management Science*, *56*(8), 1345–1361.

Anderson, C., & Wilson, J. (2003). Wait or buy? The strategic consumer: Pricing and profit implications. *Journal of the Operational Research Society*, *54*(3), 299–306.

Anderson, S. P., & Schneider, Y. (2007, December). *Price Competition with Revenue Management and Costly Consumer Search.* (Working Paper)

Arrow, K., Harris, T., & Marschak, J. (1951). Optimal inventory policy. *Econometrica: Journal of the Econometric Society*, *19*, 250–272.

The Art of Managing Yield (Tech. Rep.). (1987). American Airlines.

Axelrod, R. M. (1980). Effective Choice in the Prisoner's Dilemma. *Journal of Conflict Resolution*, *24*(1), 3–25.

Axelrod, R. M. (1984). *The Evolution of Cooperation.* Basic Books.

Axelrod, R. M., & Hamilton, W. D. (1981). The Evolution of Cooperation. *Science*, *211*(4489), 1390–1396.

Banach, S., & Tarski, A. (1924). Sur la décomposition des ensembles de points en parties respectivement congruentes. *Fundamenta Mathematica*, *6*, 244–277.

Bartke, P. (2013). *Estimating Demand from Sparse Booking Data in Airline Revenue Management.* Unpublished doctoral dissertation, Freie Universität Berlin.

Bartke, P., Cleophas, C., & Zimmermann, B. (2013). Managing Complexity in Airline Revenue Management. *Journal of Revenue and Pricing Management*, *12*(1), 36–45.

Beckmann, M. J. (1958). Decision and team problems in airline reservations. *Econometrica: Journal of the Econometric Society*, *26*, 134–145.

Bellman, R. (1954). Some problems in the theory of dynamic programming. *Econometrica: journal of the Econometric Society*, *22*(1), 37–48.

Belobaba, P. P. (1987). Airline Yield Management: An Overview of Seat Inventory Control. *Transportation Science*, *21*, 63–73.

Belobaba, P. P. (1989). Application of a probabilistic decision model to airline seat inventory control. *Operations Research*, *37*(2), 183–197.

Belobaba, P. P., & Hopperstad, C. (2004). Algorithms for Revenue Management in Unrestricted Fare Markets. In *Meeting of the INFORMS Section on Revenue Management, Massachusetts Institute of Technology, Cambridge, MA.*

Belobaba, P. P., & Wilson, J. L. (1997). Impacts of yield management in competitive airline markets. *Journal of Air Transport Management*, *3*(1), 3–9.

Berkemer, R. (2006, May). *Tit-For-Tat-like conditional strategies – A Generalization for arbitrary large groups.* (Working Paper)

Bertrand, J. (1883). Théorie des richesses: revue de 'Théorie mathématique de la richesse sociale' par Léon Walras et 'Recherches sur les principes mathématiques de la richesse'par Augustin Cournot. *Journal des Savants*, *67*, 499–508.

Bertsekas, D. (1995). *Dynamic programming and optimal control* (Vol. 1) (No. 2). Athena Scientific Belmont, MA.

Bertsimas, D., & De Boer, S. (2005). Simulation-Based Booking Limits for Airline Revenue Management. *Operations Research*, *53*(1), 90–106.

Bertsimas, D., & Perakis, G. (2006). Dynamic Pricing: A Learning Approach. In S. Lawphongpanich, D. W. Hearn, & M. J. Smith (Eds.), *Mathematical and Computational Models for Congestion Charging* (Vol. 101, pp. 45–79). Springer US.

Bertsimas, D., & Popescu, I. (2003). Revenue management in a dynamic network environment. *Transportation Science*, *37*(3), 257–277.

Besbes, O., & Zeevi, A. (2009). Dynamic Pricing Without Knowing the Demand Function: Risk Bounds and Near-Optimal Algorithms. *Operations Research*, *57*(6), 1407–1420.

Bischi, G. I., Chiarella, C., & Kopel, M. (2004). The Long Run Outcomes and Global

Dynamics of a Duopoly Game with Misspecified Demand Functions. *International Game Theory Review*, *6*(3), 343–379.

Boerlijst, M. C., Nowak, M. A., & Sigmund, K. (1997a). Equal Pay for all Prisoners. *American Mathematical Monthly*, *104*(4), 303–305.

Boerlijst, M. C., Nowak, M. A., & Sigmund, K. (1997b). The Logic of Contrition. *Journal of Theoretical Biology*, *185*(3), 281–293.

Boyd, E. A., & Kallesen, R. (2004). The science of revenue management when passengers purchase the lowest available fare. *Journal of Revenue and Pricing Management*, *3*(2), 171–177.

Boyd, R. (1989). Mistakes Allow Evolutionary Stability in the Repeated Prisoner's Dilemma Game. *Journal of Theoretical Biology*, *136*, 47–56.

Boyd, R., & Lorberbaum, J. P. (1987). No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. *Nature*, *327*, 58–59.

Brueckner, J. K., & Spiller, P. T. (1994). Economies of traffic density in the deregulated airline industry. *Journal of Law and Economics*, *37*(2), 379–415.

Brumelle, S., & Walczak, D. (2003). Dynamic airline revenue management with multiple semi-Markov demand. *Operations Research*, *51*(1), 137–148.

Brumelle, S. L., & McGill, J. I. (1993). Airline Seat Allocation with Multiple Nested Fare Classes. *Operations Research*, *41*, 127–137.

Chamberlin, E. H. (1929). Duopoly: Value Where Sellers are Few. *The Quarterly Journal of Economics*, *44*(1), 63–100.

Chattoe, E. (1996). Social Science Microsimulation. In K. G. Troitzsch, U. Mueller, N. Gilbert, & J. E. Doran (Eds.), (p. 78–104). Springer.

Chiang, W.-C., Chen, J. C., & Xu, X. (2007). An overview of research on revenue management: current issues and future research. *International Journal of Revenue Management*, *1*(1), 97–128.

Choi, S., & Mattila, A. S. (2004). Hotel revenue management and its impact on customers' perceptions of fairness. *Journal of Revenue and Pricing Management*, *2*(4), 303–314.

Cirillo, C., & Hetrakul, P. (2011). *Mixed Logit and Latent Class Model for Railway Revenue Management* (Tech. Rep.). Mid-Atlantic Universities Transportation Center.

Cleophas, C. (2009). *Simulation-Based Analysis of Forecast Performance Evaluations for Airline Revenue Management.* Unpublished doctoral dissertation, University of Paderborn.

Cleophas, C., Frank, M., & Kliewer, N. (2009a). Recent Developments in Demand Forecasting for Airline Revenue Management. *International Journal of Revenue Management*, *3*(3), 252–269.

Cleophas, C., Frank, M., & Kliewer, N. (2009b). Simulation-based key performance indicators for evaluating the quality of airline demand forecasting. *Journal of Revenue and Pricing Management*, *8*(4), 330–342.

Cléaz-Savoyen, R. L. (2005). *Airline Revenue Management Methods for Less Restricted Fare Structures.* Unpublished master's thesis, Massachusetts Institute of Technology.

Comfort, W., & Gordon, H. (1961). Vitali's theorem for invariant measures. *Transactions of the American Mathematical Society*, *99*(1), 83–90.

Cooper, W. L., de Mello, T. H., & Kleywegt, A. J. (2006). Models of the Spiral-Down Effect in Revenue Management. *Operations Research*, *54*(5), 968—987.

Cooper, W. L., de Mello, T. H., & Kleywegt, A. J. (2009, November). *Learning and Pricing with Models that do not explicitly incorporate competition.* (Working Paper)

Cournot, A. (1838). *Recherches sur les principes mathématiques de la théorie des richesses.* L. Hachette.

Cross, R. G., Higbie, J. A., & Cross, Z. N. (2011). Milestones in the application of analytical pricing and revenue management. *Journal of Revenue and Pricing Management*, *10*(1), 8–18.

Currie, C. S. M., Cheng, R. C. H., & Smith, H. K. (2008). Dynamic pricing of airline tickets with competition. *Journal of the Operational Research Society*, *59*, 1026–1037.

Curry, R. (1990). Optimal Airline Seat Allocation with Fare Classes Nested by Origins and Destinations. *Transportation Science*, *24*, 193–204.

Dana, J. D., Jr. (2008). New Directions in Revenue Management Research. *Production and Operations Management*, *17*(4), 399–401.

Decision Window Path Preference Model (DWM) (Tech. Rep.). (1994). The Boeing Airplane Company.

d'Huart, O., & Belobaba, P. P. (2012). A model of competitive airline revenue management interactions. *Journal of Revenue and Pricing Management*, *11*(1), 109–124.

Dijkstra, E. W. (1959). A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, *1*(1), 269–271.

Dockner, E., & Jørgensen, S. (1988). Optimal Pricing Strategies for New Products in Dynamic Oligopolies. *Marketing Science*, *7*(4), 315–334.

Dresher, M. (1981). *The mathematics of games of strategy: theory and applications*. Dover Publications.

Edgeworth, F. (1888). The mathematical theory of banking. *Journal of the Royal Statistical Society*, *51*(1), 113–127.

Edgeworth, F. (1897). La teoria pura del monopolio. *Giornale degli Economisti*, *40*, 13–31. (English translation: The pure theory of monopoly. In Papers Relating to Political Economy, Vol. I, edited by F. Edgeworth. London: Macmillan, 1925)

Eliashberg, J., & Jeuland, A. P. (1986). The Impact of Competitive Entry in a Developing Market upon Dynamic Pricing Strategies. *Marketing Science*, *5*(1), 20–36.

Evans, W. N., & Kessides, I. N. (1994). Living by the "Golden Rule": Multimarket Contact in the U.S. Airline Industry. *The Quarterly Journal of Economics*, *109*(2), 341–366.

Farrell, J., & Ware, R. (1989). Evolutionary Stability in the Repeated Prisoner's Dilemma. *Theoretical Population Biology*, *36*(2), 161–166.

Feuerstein, S. (2005). Collusion in Industrial Economics—A Survey. *Journal of Industry, Competition and Trade*, *5*(3-4), 163–198.

Fiig, T., Isler, K., Hopperstad, C., & Belobaba, P. (2009). Optimization of Mixed Fare Structures: Theory and Applications. *Journal of Revenue and Pricing Management*, *9*(1), 152–170.

Fischer, T., & Kamerschen, D. R. (2003). Measuring Competition in the U.S. Airline Industry Using the Rosse-Panzar Test and Cross-Sectional Regression. *Journal of Applied Economics*, *6*(1), 73–93.

Flood, M. (1958). Some experimental games. *Management Science*, *5*(1), 5–26.

Foxlin, E. (1996). Inertial head-tracker sensor fusion by a complementary separate-bias Kalman filter. In *Virtual Reality Annual International Symposium, 1996., Proceedings of the IEEE 1996* (pp. 185–194).

Frank, M., Friedemann, M., & Schröder, A. (2007a). Methodenanalyse im Revenue Management. *Internationales Verkehrswesen*, *59*(4), 140–143.

Frank, M., Friedemann, M., & Schröder, A. (2007b). Principles for simulations in revenue management. *Journal of Revenue and Pricing Management*, *7*(1), 7–16.

Frean, M. R. (1994). The Prisoner's Dilemma without Synchrony. *Royal Society of London Proceedings Series B*, *257*, 75–79.

Friedman, J. W. (1971). A Non-cooperative Equilibrium for Supergames. *The Review of Economic Studies*, *38*(113), 1–12.

Fudenberg, D., Levine, D., & Maskin, E. (1994). The Folk Theorem with Imperfect Public Information. *Econometrica*, *62*(5), 997–1039.

Fudenberg, D., & Levine, D. K. (1988). Open-Loop and Closed-Loop Equilibria in Dynamic Games with Many Players. *Journal of Economic Theory*, *44*(1), 1–18.

Fudenberg, D., & Maskin, E. (1986). The Folk Theorem in Repeated Games with Discounting or Incomplete Information. *Econometrica*, *54*(3), 533–554.

Gallego, G., & Hu, M. (2007). *Dynamic Pricing of Perishable Assets under Competition.* (Working Paper)

Gallego, G., Li, L., & Ratliff, R. (2009). Choice-based EMSR methods for single-leg revenue management with demand dependencies. *Journal of Revenue and Pricing Management*, *8*(2), 207–240.

Gallego, G., & van Ryzin, G. J. (1994). Optimal Dynamic Pricing of Inventories with Stochastic Demand over Finite Horizons. *Management Science*, *40*(8), 999–1020.

Garcia, A., & Smith, R. (1996, November). *Approximating Markov Perfect Equilibria* *for Nonstationary Dynamic Games* (Tech. Rep.). Department of Industrial and Operations Engineering, University of Michigan.

Gayle, P. G. (2004). Does price matter? Price and Non-price Competition in the Airline Industry. In *North American Summer Meeting of the Econometric Society.*

Geraghty, M. K., & Johnson, E. (1997). Revenue Management Saves National Car Rental. *Interfaces*, *27*(1), 107–127.

Gerlach, M., Cleophas, C., & Kliewer, N. (2013). Airline Codeshare Alliances. *Business & Information Systems Engineering*, *2*, 1–11.

Glover, F., Glover, R., Lorenzo, J., & McMillan, C. (1982). The passenger-mix problem in the scheduled airlines. *Interfaces*, *12*(3), 73–80.

Gorin, T., & Belobaba, P. P. (2004). Revenue management performance in a low-fare airline environment: Insights from the Passenger Origin-Destination Simulator. *Journal of Revenue and Pricing Management*, *3*(3), 215–236. (Special Issue Papers)

Gorin, T., & Belobaba, P. P. (2008). Assessing predation in airline markets with low-fare competition. *Transportation Research Part A: Policy and Practice*, *42*(5), 784–798.

Gotts, N., Polhill, J., & Law, A. (2003). Agent-Based Simulation in the Study of Social Dilemmas. *Artificial Intelligence Review*, *19*(1), 3–92.

Grewal, M. S., Henderson, V. D., & Miyasako, R. S. (1991). Application of Kalman filtering to the calibration and alignment of inertial navigation systems. *Automatic Control, IEEE Transactions on*, *36*(1), 3–13.

Grüne-Yanoff, T. (2009). Learning from Minimal Economic Models. *Erkenntnis*, *70*(1), 81–99.

Guo, P., Xiao, B., & Li, J. (2012). Unconstraining Methods in Revenue Management Systems: Research Overview and Prospects. *Advances in Operations Research*, *2012*, 1–23.

Harris, F. H. d. (2007). Large-scale entry deterrence of a low-cost competitor: an early success of airline revenue management. *International Journal of Revenue Management*, *1*(1), 5–27.

Hauert, C., & Schuster, H. (1997). Effects of increasing the number of players and memory size in the iterated Prisoner's Dilemma: a numerical approach. *Royal*

Society of London Proceedings Series B, *264*(1381), 513–519.

Hayes, K. J., & Ross, L. B. (1998). Is Airline Price Dispersion the Result of Careful Planning or Competitive Forces? *Review of Industrial Organization*, *13*(5), 523–541.

Hess, J. D., & Gerstner, E. (1991). Price-matching Policies: An Empirical Case. *Managerial and Decision Economics*, *12*(4), 305–315.

Hopperstad, C. (1995). *Potential Application of the Boeing Passenger Origin-Destination Simulator (PODS) to Yield Management System Planning* (Tech. Rep.). Boeing Commercial Airplane Group.

Huang, D., Zhou, H., & Zhao, Q.-H. (2010). A competitive multiple-product newsboy problem with partial product substitution. *Omega*, *39*, 302–312.

Hungerford, T. W. (1974). Algebra. In *Graduate Texts in Mathematics* (Vol. 73). Springer.

Imhof, L. A., Fudenberg, D., & Nowak, M. A. (2007). Tit-for-tat or Win-stay, Lose-shift? *Journal of Theoretical Biology*, *247*(3), 574–580.

Infare Solutions. (2013, August). http://www.infare.com/. Retrieved from http://www.infare.com/

Iosifescu, M., & Grigorescu, S. (1990). *Dependence with complete connections and its applications* (Vol. 96). Cambridge University Press.

Isler, K., & Imhof, H. (2008). A game theoretic model for airline revenue management and competitive pricing. *Journal of Revenue and Pricing Management*, *7*(4), 384–396.

Isler, K., & Imhof, H. (2010). Airline Revenue Management and Pricing in Competitive Markets. In *Symposium On Challenges in Commodity-Energy pricing and Revenue management (Soccer), Academy of Sciences and Humanities, Heidelberg.*

Jiang, H., Netessine, S., & Savin, S. (2011). Robust Newsvendor Competition under Asymmetric Information. *Operations Research*, *1*, 254–261.

Kallenberg, O. (2002). *Foundations of Modern Probability.* Springer Verlag.

Kalman, R. E., et al. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of basic Engineering*, *82*(1), 35–45.

Kalnins, A., Froeb, L., & Tschantz, S. (2010, September). *Mergers Increase Output When Firms Compete by Managing Revenue.* (Working Paper)

Khouja, M. (1999). The single-period (news-vendor) problem: literature review and suggestions for future research. *Omega*, *27*(5), 537–553.

Kimms, A., & Müller-Bungart, M. (2007). Revenue management for broadcasting commercials: the channel's problem of selecting and scheduling the advertisements to be aired. *International Journal of Revenue Management*, *1*(1), 28–44.

Kincaid, W. M., & Darling, D. A. (1963). An Inventory Pricing Problem. *Journal of Mathematical Analysis and Applications*, *7*, 183–208.

Kraines, D. P., & Kraines, V. Y. (2000). Natural Selection of Memory-one Strategies for the Iterated Prisoner's Dilemma. *Journal of Theoretical Biology*, *203*(4), 335–355.

Kwoka, J., & Shumilkina, E. (2010). The price effect of eliminating potetntial competition: Evidence from an airline merger. *The Journal of Industrial Economics*, *58*(4), 767–793.

Kwon, C., Friesz, T. L., Mookherjee, R., Yao, T., & Feng, B. (2009). Non-cooperative competition among revenue maximizing service providers with demand learning. *European Journal of Operational Research*, *197*(3), 981–996.

Ladany, S. P., & Arbel, A. (1991). Optimal cruise-liner passenger cabin pricing policy. *European Journal of Operational Research*, *55*(2), 136–147.

Law, A. M., & Kelton, D. W. (2000). *Simulation Modeling and Analysis.* McGraw Hill Boston, MA.

Ledvina, A. (2011). *Differential Games in Oligopolistic Markets.* Unpublished doctoral dissertation, Princeton University.

Ledvina, A., & Sircar, R. (2011). Dynamic Bertrand Oligopoly. *Applied Mathematics & Optimization*, *63*(1), 11–44.

Lee, T. C., & Hersh, M. (1993). A Model for Dynamic Airline Seat Inventory Control with Multiple Seat Bookings. *Transportation Science*, *27*(3), 252–265.

Lin, K. Y., & Sibdari, S. Y. (2009). Dynamic price competition with discrete customer choices. *European Journal of Operational Research*, *197*, 969–980.

Lippman, S. A., & McCardle, K. F. (1997). The Competitive Newsboy. *Operations Research*, *45* (1), 54–65.

Littlewood, K. (1972). Forecasting and control of passenger bookings. In *Proceedings of the Twelfth Annual AGIFORS Symposium.*

Loan, C. F. V. (2000). The ubiquitous Kronecker product. *Journal of Computational and Applied Mathematics*, *123* (1), 85–100.

Lua, W. F. (2007). *Matching of Lowest Fare Seat Availability in Airline Revenue Management Systems.* Unpublished master's thesis, Massachusetts Institute of Technology.

Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, *16* (4), 451–476.

Marcus, B., & Anderson, C. K. (2008). Revenue management for low-cost providers. *European Journal of Operational Research*, *188*, 258–272.

Martínez-de Albéniz, V., & Talluri, K. (2011). Dynamic Price Competition with Fixed Capacities. *Management Science*, *57* (6), 1078–1093.

Maskin, E., & Tirole, J. (2001, October). Markov Perfect Equilibrium: I. Observable Actions. *Journal of Economic Theory*, *100* (2), 191–219.

Maynard Smith, J. (1982). *Evolution and the Theory of Games.* Cambridge University Press.

Maynard Smith, J., & Price, G. (1973). The Logic of Animal Conflict. *Nature*, *246*, 15.

McGill, J. I., & van Ryzin, G. J. (1999). Revenue Management: Research Overview and Prospects. *Transportation Science*, *33* (2), 233–256.

Molander, P. (1985). The Optimal Level of Generosity in a Selfish, Uncertain Environment. *Journal of Conflict Resolution*, *29* (4), 611–618.

Mookherjee, R., & Friesz, T. L. (2008). Pricing, Allocation, and Overbooking in Dynamic Service Network Competition When Demand is Uncertain. *Production and Operations Management*, *17* (4), 455–474.

Morrison, S. A., Winston, C., Bailey, E. E., & Carlton, D. W. (1996). Causes and Consequences of Airline Fare Wars. *Brookings Papers on Economic Activity.*

Microeconomics, *1996*, 85–131.

Mukhopadhyay, S., Samaddar, S., & Colville, G. (2007). Improving Revenue Management Decision Making for Airlines by Evaluating Analyst-Adjusted Passenger Demand Forecasts. *Decision Sciences*, *38*(2), 309–327.

Nash, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, *36*(1), 48–49.

Neill, D. B. (2001). Optimality Under Noise: Higher Memory Strategies for the Alternating Prisoner:s Dilemma. *Journal of Theoretical Biology*, *211*, 159–180.

Netessine, S., & Shumsky, R. (2002). Introduction to the theory and practice of yield management. *INFORMS Transactions on Education*, *3*(1), 34–44.

Netessine, S., & Shumsky, R. A. (2005). Revenue Management Games: Horizontal and Vertical Competition. *Management Science*, *51*(5), 813–831.

Nomani, A. Q. (1990). Fare Warning: How Airlines Trade Price Plans. *Wall Street Journal*, *9 October*.

Nowak, M., & Sigmund, K. (1990). The Evolution of Stochastic Strategies in the Prisoner's Dilemma. *Acta Applicandae Mathematicae*, *20*(3), 247–265.

Nowak, M., & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, *355*(6357), 250–253.

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, *364*(6432), 56–58.

Orlov, E. (2011). How does the internet influence price dispersion? Evidence from the airline industry. *The Journal of Industrial Economics*, *59*(1), 21–37.

Örmeci, E., & Burnetas, A. (2004). Admission control with batch arrivals. *Operations Research Letters*, *32*(5), 448–454.

Osborne, M., & Rubinstein, A. (1994). *A Course in Game Theory.* MIT press.

Pelc, A. (2010). Fault-tolerant strategies in the Iterated Prisoner's Dilemma. *Information Processing Letters*, *110*(10), 389–395.

Perakis, G., & Sood, A. (2006). Competitive Multi-period Pricing for Perishable Products:

A Robust Optimization Approach. *Mathematical Programming, Series B*, *107*, 295–335.

Pinsky, M., & Karlin, S. (2010). *An Introduction to Stochastic Modeling*. Academic press.

Press, W., & Dyson, F. (2012). Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, *109*(26), 10409–10413.

Pölt, S. (2011). The rise and fall of RM. *Journal of Revenue and Pricing Management*, *10*(1), 23–25.

Rajamani, M. R. (2007). *Data-based Techniques to Improve State Estimation in Model Predictive Control*. Unpublished doctoral dissertation, University of Wisconsin-Madison.

Ratliff, R., & Vinod, B. (2005). Airline pricing and revenue management: A future outlook. *Journal of Revenue and Pricing Management*, *4*(3), 302–307.

Reyes, M. H. (2006). *Hybrid Forecasting for Airline Revenue Management in Semi-Restricted Fare Structures*. Unpublished master's thesis, Massachusetts Institute of Technology.

Rubinstein, A. (1979). Equilibrium in Supergames with the Overtaking Criterion. *Journal of Economic Theory*, *21*(1), 1–9.

Shen, Z., & Su, X. (2007). Customer Behavior Modeling in Revenue Management and Auctions: A Review and New Research Opportunities. *Production and Operations Management*, *16*(6), 713–728.

Silver, E., Pyke, D., Peterson, R., et al. (1998). *Inventory Management and Production Planning and Scheduling* (Vol. 3). Wiley (New York).

Simon, C. (2007). *Dynamic Pricing with Demand Learning under Competition*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Smith, B. C., Leimkuhler, J. F., & Darrow, R. M. (1992). Yield Management at American Airlines. *Interfaces*, *22*, 8–31.

Smith, B. C., & Penn, C. W. (1988). Analysis of Alternative Origin-Destination Control Strategies. In *Proceedings of the Twenty Eighth Annual AGIFORS Symposium*

(Vol. 28).

Stavins, J. (2001). Price discrimination in the airline market: The effect of market concentration. *Review of Economics and Statistics*, *83*(1), 200–202.

Stromberg, K. (1979). The Banach-Tarski paradox. *The American Mathematical Monthly*, *86*(3), 151–161.

Subramanian, J., Stidham, S., & Lautenbacher, C. J. (1999). Airline Yield Management with Overbooking, Cancellations, and No-Shows. *Transportation Science*, *33*(2), 147–167.

Sugden, R. (1986). *The Economics of Rights, Co-operation, and Welfare.* Blackwell.

Talluri, K. T. (2003, July). *On Equilibria in Duopolies with Finite Strategy Spaces.* (Working Paper)

Talluri, K. T., Castejon, F., Codina, B., & Magaz, J. (2010). Proving the performance of a new revenue management system. *Journal of Revenue and Pricing Management*, *9*(4), 300–312.

Talluri, K. T., & van Ryzin, G. J. (1998). An Analysis of Bid-Price Controls for Network Revenue Management. *Management Science*, *44*(11-Part-1), 1577–1593.

Talluri, K. T., & van Ryzin, G. J. (2004a). Revenue Management Under a General Discrete Choice Model of Consumer Behavior. *Management Science*, *50*(1), 15–33.

Talluri, K. T., & van Ryzin, G. J. (2004b). *The Theory and Practice of Revenue Management.* Kluwer Academic Publishers.

Thomas, B. (1985). On evolutionarily stable sets. *Journal of Mathematical Biology*, *22*(1), 105–115.

Train, K. (2003). *Discrete Choice Methods with Simulation.* Cambridge University Press.

Travel Market Simulator. (2013a, August). http://www.dsim.cc/. Retrieved from http://www.dsim.cc/

Travel Market Simulator. (2013b, August). Retrieved from http://www.linkedin.com/in/airsim

Tucker, A. W. (1980). On Jargon: The Prisoners' Dilemma. *UMAP journal*, *1*(S 101).

Vinod, B. (2005). Retail revenue management and the new paradigm of merchandise optimisation. *Journal of Revenue and Pricing Management*, *3*(4), 358–368.

Vitali, G. (1905). *Sul problema della misura dei gruppi di punti di una retta*. Gamberini e Parmeggiani, Bologna.

von Neumann, J. (1928). Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, *100*(1), 295–320.

Walczak, D., Mardan, S., & Kallesen, R. (2009). Customer choice, fare adjustments and the marginal expected revenue data transformation: A note on using old yield management techniques in the brave new world of pricing. *Journal of Revenue and Pricing Management*, *9*(1), 94–109.

Weatherford, L. R. (2004). EMSR versus EMSU: Revenue or utility? *Journal of Revenue and Pricing Management*, *3*(3), 277–284.

Weatherford, L. R. (2009). Survey results from airlines that use revenue management software systems. *Journal of Revenue and Pricing Management*, *8*(4), 323–329.

Weatherford, L. R., & Kimes, S. E. (2003). A comparison of forecasting methods for hotel revenue management. *International Journal of Forecasting*, *19*, 401–415.

Weatherford, L. R., & Ratliff, R. M. (2010). Review of revenue management methods with dependent demands. *Journal of Revenue and Pricing Management*, *9*(4), 326–340.

Wedekind, C., & Milinski, M. (1996). Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(7), 2686–2689.

Williamson, E. L. (1992). *Airline Network Seat Inventory Control: Methodologies and Revenue Impacts*. Unpublished doctoral dissertation, Flight Transportation Laboratory, MIT, Cambridge, MA.

Winter, T. (2010). Market sensitive revenue optimization: Science and practice. In *GOR Pricing and Revenue Management Working Group Annual Meeting Proceedings. Frankfurt, Germany.*

Wishlinski, B. (2006). Extracting the most benefit from a revenue management solution implementation. *Journal of Revenue and Pricing Management*, *5*(2), 109–117.

Wollmer, R. (1992). An Airline Seat Management Model for a Single Leg Route when Lower Fare Classes Book First. *Operations Research*, *40*, 26–37.

Wu, J., & Axelrod, R. M. (1995). How to Cope with Noise in the Iterated Prisoner's Dilemma. *Journal of Conflict Resolution*, *39*(1), 183–189.

Yao, X. (1996). Evolutionary stability in the n-person iterated prisoner's dilemma. *Biosystems*, *37*(3), 189–197.

Zeni, R. H. (2003). The value of analyst interaction with revenue management systems. *Journal of Revenue and Pricing Management*, *2*(1), 37–46.

Zeni, R. H. (2007). Can we really look to the past to forecast future demand? *Journal of Revenue and Pricing Management*, *6*(4), 312–314.

Zhang, D., & Kallesen, R. (2008). Incorporating competitive price information into revenue management. *Journal of Revenue and Pricing Management*, *7*(1), 17–26.

Zhao, X. (2003). *Competitive supply chain and revenue management : four essays.* Unpublished doctoral dissertation, University of British Columbia, Vancouver, Canada.

Zhao, X., & Atkins, D. R. (2008). Newsvendors Under Simultaneous Price and Inventory Competition. *Manufacturing & Service Operations Management*, *10*(3), 539–546.

Zhou, Y., Fan, M., & Cho, M. (2005). *On the Threshold Purchasing Behavior of Customers Facing Dynamically Priced Perishable Products* (Tech. Rep.). Working paper.

Zimmermann, B., Cleophas, C., & Frank, M. (2011). Universeller Simulator für Revenue Management. *Internationales Verkehrswesen*, *63*(3), 76–84.

Symbols

x^+	max(0,x).
1_n	n-dimensional vector, where every element equals one.
A	Matrix describing the change of basis of players' reputation in the revenue management game Markov chain.
adj(A)	Adjugate of matrix A.
α	Exponential smoothing parameter.
$A \setminus B$	The complement of set B in set A.
B	Matrix for permutation of both players' probability vectors in the revenue management game Markov chain.
B_k	Matrix mapping S_l's probability vector to S_k's probability vector space in the revenue management game Markov chain.
$b(j_k,j_k^{min},j_l^{min})$	Bookings in class j of firm S_k, given that the lowest available product of S_k is j_k^{min} and the lowest available product of S_l is j_l^{min}.
$\mathcal{B}(J)$	Borel-σ-Algebra over the set J.
$A \times B$	Cartesian product of sets A and B.
$c_{elasticity}$	Price elasticity of dependent part of hybrid demand.
$A \circ B$	Composition of maps A and B.
c	Cooperative threshold product used for tacit collusion.
c_p	Parameter used in calculation of distance-dependent part of process covariation during Kalman filtering.
$\Delta U_t(x)$	Bid price $U_t(x) - U_t(x-1)$.
Δ	Distance matrix used in Kalman filtering.
$\delta_{i,j}$	Kronecker delta. The Kronecker delta is 1, if $i = j$, otherwise it is 0.
δ_x	Dirac measure concentrated on x.
d^D	Dependent demand forecast.
d^P	Psychic forecast that takes competitor availability into account.
d^H	Hybrid demand forecast.
d_{base}	Dependent part of hybrid demand for cheapest product.

d^I	Independent demand forecast.
$\tilde{d}(j)$	Marginal demand of product j.
$\hat{d}$	Point estimate of demand.
d	Reference value of demand.
D	True demand depending on both service provider's offer set.
$\det(A)$	Determinant of matrix A.
d^K	Dependent demand estimated using a Kalman filter.
E	Matrix describing observation errors for a single player in the revenue management game Markov chain.
$E(f,A)$	Transition kernel of a single player's observation errors in the revenue management game Markov process with continuous prices.
e_i	The i-th canonical basis vector of $\mathbb{R}^n$.
$E^{i,j}$	Matrix with values of 1 in the diagonal and a single 1 in the i-th row and j-th column, otherwise 0.
ε	Probability of an observation error.
η	Linear dependence parameter in linear relationship of players' payoff enforced by a Zero-Determinant Strategy.
$f(j)$	Price of product j.
f_c	Jointly optimal fare in the revenue management game with continuous prices.
$\tilde{f}(j)$	Marginal fare of product j.
F_k	Price range of S_k in the revenue management game with continuous prices.
$\mathcal{F}_k$	σ-Algebra of S_k's price range in the revenue management game with continuous prices.
$f^\star(f)$	Single-stage revenue-maximizing price given a competitor price f in the revenue management game Markov process with continuous prices.
γ	Probability of generosity of GTFT.
H	Reconstraining matrix used in Kalman filtering.
I_n	n-dimensional identity matrix.
j_k^{min}	Lowest available product of service provider S_k.
$\hat{j}_k^{min}$	Observation of service provider S_k's lowest available product including observation errors.
j	Product.
J	Product set offered by a service provider.
K	Kalman gain.

$\overline{K}_{\sigma_k}(f,A)$	Transition kernel of general Markov strategy in the revenue management game Markov process with continuous prices.
$\overline{M}_{\sigma_k}$	General form of a single-player strategy for strategy σ_k in the revenue management game Markov chain.
$m(A,i,j)$	Matrix constructed by omitting the i-th row and j-th column of matrix A.
M_{σ_k}	Single-player strategy matrix for reactive strategy σ_k in the revenue management game Markov chain.
$\widetilde{M}_{\sigma_k}$	Building block of player S_k's reactive strategy matrix in the revenue management game Markov chain.
$M^{\varepsilon}_{\sigma_1,\sigma_2}$	Transition matrix of two strategies σ_1 and σ_2 in the presence of observation errors in the revenue management game Markov chain.
$M^{\varepsilon,r}_{\sigma_1,\sigma_2}$	Transition matrix of two strategies σ_1 and σ_2 accounting for reputation in the presence of observation errors in the revenue management game Markov chain.
$M^{r}_{\sigma_1,\sigma_2}$	Transition matrix of two strategies σ_1 and σ_2 accounting for reputation without observation errors in the revenue management game Markov chain.
M_{σ_1,σ_2}	Transition matrix of the revenue management game Markov chain for the strategy pair σ_1 and σ_2 without observation errors.
μ_s	Measure on the product of both players' price spaces induced by the revenue management game Markov process with continuous fares at stage s.
$\mathbb{N}^+$	Natural numbers without zero $\{1,2,\dots\}$.
$\mathbb{N}_0$	Natural numbers with zero $\{0,1,\dots\}$.
$x \sim \mathcal{N}(\mu,\Sigma)$	Random variable x distributed following a normal distribution with mean μ and covariance matrix Σ.
ν^k_s	Marginal measure on player S_k's price space induced by the revenue management game Markov process with continuous fares at stage s.
n	Number of Products.
$\varnothing$	The empty set.
ϕ_i	Linear dependence parameter.
π	Stationary distribution.
$\mathcal{P}(J)$	Power Set of the set J.
P_k	Space of probabilities of S_k's reputation, i.e. $P_k = \{(\rho^1_k, \rho^2_k) \in [0,1]^2 : \rho^1_k + \rho^2_k = 1\}$.
$\mathbb{P}$	Probability.

Q	Matrix that duplicates input to both players of the revenue management game Markov chain.
$\mathbb{R}^+$	Non-negative real numbers.
$\mathbb{R}$	Real numbers.
r_k	Reputation of S_k, where $r_k = 0$ represents a good and $r_k = 1$ a bad reputation.
ρ_k	Vector of probability distribution of S_k's reputation, where $\mathbb{P}(r_k = i) = \rho_k^i$.
R_k	Space of possible reputations for S_k in the revenue management game Markov chain.
$R((\cdot,r_1,r_2),(\cdot,r_1',r_2'))$	Transition kernel of both players' reputations in the revenue management game Markov process with continuous prices.
R	Matrix describing the evolution of reputation of both players in the revenue management game Markov chain.
S_k	Service provider participating in the revenue management game.
Σ^b	Booking covariance matrix used in Kalman filtering.
Σ^d	Reference covariance matrix used in Kalman filtering.
Σ^p	Process covariance matrix used in Kalman filtering.
σ_k	Strategy of S_k.
s	Stage of the dynamic revenue management game. Each stage covers a single sales period.
s_k	Expected payoff of player S_k given the stationary distribution of the revenue management game Markov chain.
$A \otimes B$	Tensor product of vector spaces, σ-Algebras or linear maps A and B.
θ	Linear dependence parameter in linear relationship of players' payoff enforced by a Zero-Determinant Strategy.
T	End of booking horizon.
t	Element of the discretization of the booking horizon, called time step.
x^t	Transpose of x.
$U_t(x)$	Value of remaining capacity x at time step t.
$U_t^k(x_k,x_l)$	Value of service provider S_k's remaining capacity x_k given competitor S_l's remaining capacity x_l at time step t.

V_k	Space of probability vectors of S_k's lowest available product $\{(v_k^1, \ldots, v_k^n) \in [0,1]^n : \sum_i v_k^i = 1\}$ in the revenue management game Markov chain.
v_k^j	Probability that S_k's lowest available product is j in the revenue management game Markov chain.
χ	Extortion parameter of Extortionate Strategies.
X_s	State of the revenue management game Markov process at stage s.
y_k	Payoff vector of player S_k for all combinations of lowest available products of both players in the revenue management game Markov chain.
$Y_k(i,j)$	Revenue of player S_k given that the lowest available product of S_1 is i and the lowest available product of S_2 is j in the revenue management game Markov chain.
ζ	Linear dependence parameter in linear relationship of players' payoff enforced by a Zero-Determinant Strategy.

Acronyms

APD	Alternating Prisoner's Dilemma. 21
CTFT	Contrite Tit for Tat. 20, 48–50, 62, 63, 65–67, 71, 77, 78, 103, 106, 110, 123, 127, 149, 152, 154
DAVN	Displacement Adjusted Virtual Nesting. 11
EMSR	Expected Marginal Seat Revenue. 10
EMSRa	Expected Marginal Seat Revenue—Version a. 10
EMSRb	Expected Marginal Seat Revenue—Version b. 10
ESS	Evolutionary Stable Strategy. 16, 18, 20, 123
FCFS	First Come, First Serve. 55, 58
GTFT	Generous Tit for Tat. 19, 48–50, 61–63, 65–67, 71, 77, 78, 102, 105, 106, 110, 114, 116, 118, 123–125, 127, 149, 152, 154, 175
IPD	Iterated Prisoner's Dilemma. 3, 4, 17–21, 29, 31–35, 46–50, 61, 63, 71, 76, 77, 80, 81, 90, 92, 96, 98, 101–103, 105–108, 110, 122, 123, 125, 127–129, 134, 137, 138, 140, 147, 150–153, 155–158
O&D	Origin and Destination. 7, 10
PODS	Passenger Origin-Destination Simulator. 22, 28–30, 34, 52
REMATE	REvenue Management Training for Experts. 28, 29, 34, 36, 51–56, 58, 68, 77, 154
RM	Revenue Management. 1–14, 16, 21–36, 38, 40, 42, 43, 46–54, 58, 59, 63, 67–69, 71, 76, 77, 79–83, 90, 91, 94, 96–98, 101, 102, 104, 105, 108, 110, 122, 127–129, 133–135, 137, 143, 150, 152–158

TF2T	Tit for 2 Tats. 19
TFT	Tit for Tat. 17–21, 48–50, 59, 61, 63, 65–67, 71, 77, 78, 102, 103, 105, 108–114, 116, 118, 119, 123, 127, 148, 151, 154
ZD	Zero-Determinant Strategy. 132, 133

Glossary

ALLC	Strategy for repeated game that plays single-stage cooperative Nash equilibrium strategy every stage. 48, 51, 58, 59, 63, 65, 69–72, 102, 104, 111, 118, 120, 121, 123–125, 127, 148, 152, 153
ALLD	Strategy for repeated game that plays single-stage non-cooperative Nash equilibrium strategy every stage. 48, 51, 58, 59, 61, 62, 65, 66, 69–72, 75, 80, 102, 104, 109, 111, 113, 116, 118, 119, 123–125, 148, 153
Competitive Spiral Down	Systematic deterioration of forecasts and profits due to competitive effects. 26, 27, 31, 34, 35, 48, 51, 58, 59, 61, 62, 65, 66, 68–71, 76–79, 96, 104, 105, 109, 110, 156, 157
Contrite Tit for Tat	Strategy for repeated game executing similar logic to Tit for Tat to players' reputation. 20, 48–50, 62, 63, 65–67, 71, 77, 78, 103, 106, 110, 123, 127, 149, 152, 154, 179
COOP	Strategy for single-stage game that plays cooperative Nash equilibrium strategy. 47–49, 101–103, 108, 147–149
DEFECT	Strategy for single-stage game that plays non-cooperative Nash equilibrium strategy. 47–49, 101–103, 107, 108, 147–149
Dynamic Pricing	Revenue Management using price as control variable. 9, 13, 24, 26, 143, 146, 148, 152
Equalizing Strategy	A Zero-Determinant Strategy, which sets the competitor's payoff to a fixed amount. 20, 128, 133
Expected Marginal Seat Revenue	Heuristic revenue management optimization technique, which generalizes Littlewood's rule to more than two classes. 10, 179

Extortionate Strategy	A Zero-Determinant Strategy, which fixes a linear relationship of both players' profit. 21, 33, 81, 82, 128, 129, 133, 134, 137, 140, 143, 150, 152, 156, 157, 178
General Markov Strategy	Strategy in the revenue management game taking into account both players' previous actions. 91, 103, 110, 117, 129, 135, 138, 140, 144, 145, 149
Generous Tit for Tat	Strategy for repeated game adding a bias to cooperation to the logic of Tit for Tat. 19, 20, 48–50, 61–63, 65–67, 71, 77, 78, 102, 105, 106, 110, 114, 116, 118, 123–125, 127, 149, 152, 154, 175, 179
Iterated Prisoner's Dilemma	Repeated version of a symmetric 2-player game, in which mutual cooperation maximizes the players' combined payoffs, but mutual non-cooperation is the only Nash equilibrium. 3, 4, 17–21, 29, 31–35, 46–50, 61, 63, 71, 76, 77, 80, 81, 90, 92, 96, 98, 101–103, 105–108, 110, 122, 123, 125, 127–129, 134, 137, 138, 140, 147, 150–153, 155–158, 179
JOI	Jointly optimal strategy in the repeated game. 58
LOW	Strategy for repeated game that always makes the lowest price available. 58, 66
MATCH	Strategy for repeated game that copies competitor's last action. 49, 50, 65–67, 75, 78, 103, 107, 108, 123, 127, 149, 152, 154, 155
PAVLOV	Strategy for repeated game relying on a Win-Stay, Lose-Shift logic. 20, 49, 50, 62, 63, 65–67, 71, 72, 77, 78, 103, 104, 107–109, 117–121, 123, 125, 127, 149, 151, 152, 154
PODS	Revenue management simulator developed by Boeing and the MIT. 22, 28–30, 34, 52, 179
Q-Forecasting	Forecasting method using a customer model, in which customers always buy the cheapest available product and the customers' willingness to pay follows an exponential distribution. 8, 38

Reactive Strategy	Strategy in the revenue management game taking into account only competitor's previous actions. 92, 102, 103, 110, 111, 114, 117, 144, 146, 148, 149
REMATE	Revenue management simulator developed by Lufthansa. 28, 29, 34, 36, 51–56, 58, 68, 77, 154, 179
REMIGIUS	Revenue management simulator developed by the university of Clausthal. 28
Spiral Down	Systematic deterioration of forecasts and profits due to incorrect models of customer behavior. 8, 69, 72, 78, 156
Tit for 2 Tats	Strategy for repeated game based on Tit for Tat. 19, 180
Tit for Tat	Strategy for repeated game based on the concept of equivalent retaliation. 17–21, 48–50, 59, 61–63, 65–67, 71, 77, 78, 102, 103, 105, 107–114, 116, 118, 119, 123, 127, 151, 152, 154, 180–183
Unconstraining	Process of reconstructing the true demand process from censored observations. 7
UNDER	Strategy for repeated game that underbids competitor's last action. 50, 65–67, 75, 76, 78, 103, 104, 123, 154, 155
Zero-Determinant Strategy	A strategy in a two-player game allowing for a linear dependence of payoffs. 21, 128, 129, 131–133, 152, 175, 177, 178, 180–182