Dynamic Pricing with Demand Learning under Competition

by

Carine Simon
Ingénieur des Arts et Manufactures, Ecole Centrale Paris
(2004)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2007

Author . . . .
Sloan School of Management
August 16, 2007

Certified by . . . .
Georgia Perakis
Associate Professor
Thesis Supervisor

Certified by . . . .
Soulaymane Kachani
Assistant Professor
Thesis Supervisor

Accepted by . . . .
Dimitris Bertsimas
Co-Director, Operations Research Center

Dynamic Pricing with Demand Learning under Competition

by

Carine Simon

Ingénieur des Arts et Manufactures, Ecole Centrale Paris (2004)

Submitted to the Sloan School of Management
on August 16, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

Abstract

In this thesis, we focus on oligopolistic markets for a single perishable product, where firms compete by setting prices (Bertrand competition) or by allocating quantities (Cournot competition) dynamically over a finite selling horizon. The price-demand relationship is modeled as a parametric function, whose parameters are unknown, but learned through a data driven approach. The market can be either in disequilibrium or in equilibrium.

In disequilibrium, we consider simultaneously two forms of learning for the firm: (i) learning of its optimal pricing (resp. allocation) strategy, given its belief regarding its competitors' strategy; (ii) learning the parameters in the price-demand relationship. In equilibrium, each firm seeks to learn the parameters in the price-demand relationship for itself and its competitors, given that prices (resp. quantities) are in equilibrium.

In this thesis, we first study the dynamic pricing (resp. allocation) problem when the parameters in the price-demand relationship are known. We then address the dynamic pricing (resp. allocation) problem with learning of the parameters in the price-demand relationship. We show that the problem can be formulated as a bilevel program in disequilibrium and as a Mathematical Program with Equilibrium Constraints (MPECs) in equilibrium.

Using results from variational inequalities, bilevel programming and MPECs, we prove that learning the optimal strategies as well as the parameters, is achieved. Furthermore, we design a solution method for efficiently solving the problem. We prove convergence of this method analytically and discuss various insights through a computational study.

Finally, we consider closed-loop strategies in a duopoly market when demand is stochastic. Unlike open-loop policies (such policies are computed once and for all at the beginning of the time horizon), closed loop policies are computed at each time period, so that the firm can take advantage of having observed the past random disturbances in the market. In a closed-loop setting, subgame perfect equilibrium is the relevant notion of equilibrium. We investigate the existence and uniqueness of a subgame perfect equilibrium strategy, as well as approximations of the problem in order to be able to compute such policies more efficiently.

Thesis Supervisor: Georgia Perakis
Title: Associate Professor

Thesis Supervisor: Soulaymane Kachani
Title: Assistant Professor

Acknowledgments

I would like to express my extreme gratitude to my academic and research advisor Prof. Georgia Perakis. I was very fortunate to work with Georgia who provided me with a wealth of opportunities to develop skills in research and teaching. She encouraged me to write papers, submit my research to competitions, present in conferences, and enabled me to develop a taste for teaching. But Georgia's involvement went far beyond my professional development. I am extremely thankful for her continuous support throughout my four years at MIT. It made a huge difference in my life during my PhD, to know that Georgia cared for me on a personal level, and that I could turn to her for advice and support, not only for academic or professional questions, but also for personal problems. While pushing me to always give my best, she was always understanding of the ups and downs a student can feel and always cared about my well-being. Georgia, I am deeply indebted.

I would like to extend my thanks to my research advisor Prof. Soulaymane Kachani. Soulaymane has been a source of inspiration for me since my studies in France at Ecole Centrale Paris. He was a great source of information and advice since my application to graduate school in the United States. As a thesis advisor, he taught me to balance theory and applications, and to give good presentations. Last, but not least, I am very grateful for his help and advice when I decided to go to industry.

I would like to thank the members of my thesis committee, Profs. Dimitris Bertsimas and Steve Graves. Their feedback was extremely invaluable, and dramatically improved the quality of this work.

During my doctoral studies at MIT, I was fortunate to be given teaching opportunities, and discovered a passion for teaching. Profs. Bertsimas, Freund, Graves and Simchi-Levi were all great teachers and mentors, and it was an amazing learning experience, as well as a great pleasure, to be able to see them teach. I was impressed by their ability to combine theory and practice, to deliver clear managerial insights and to cater to audiences as diverse as doctoral students and MBA students can be. Those classes were my best at MIT, and I am grateful to them for that.

What made me choose the Operations Research Center for my doctoral studies, are the amazing people I met when I first visited. I made great friends at the ORC, with whom I shared some hard times - finals, qualifying exams -, and especially very good times: the annual retreat, ski trips, parties, spring breaks, etc. My PhD was no only an academic, but also a human experience. The Operations Research Center was a second family for four years, and I hope to keep in touch with my friends at the ORC. Do visit me in New York! Finally, life at the ORC would not be so smooth, without the help of a very dedicated staff: Andrew Carvahlo, Veronica Mignott, Paulette Mosley and Laura Rose were always there to answer my questions.

This year, I was extremely happy to live with three amazing friends: David, Jonathan and Martin. I would like to thank them for their support in the bad times, and all the fun and laughs in the good times. You made this last year at MIT very special for me, and I feel extremely lucky to have you as my friends.

Last, but not least, I would not be here without the unconditional love and support of my family. I lack the words to express my infinite gratitude to my parents. You

made me everything that I am today! I cannot thank my sister Marjorie enough for all her support, even though I have been a busy sister. Thank you Joshua, for always being at my sides and making this journey with me. I would like to dedicate my thesis to my loved ones: my family, and Joshua.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation: Challenges of Real-Life Problems

In recent years, dynamic pricing and inventory control have drawn increased interest from both practitioners and researchers. A study by McKinsey and Company on the cost structure of Fortune 1000 companies conducted annually shows that implementing good pricing policies yields bigger revenue gains than reductions of variable costs, fixed costs or an increase in sales volumes:

> "Pricing right is the fastest and most effective way for managers to increase profits."

Determining the right price to charge customers, or the right quantity to sell to customers, requires a company to have a wealth of information and data. In particular, the company needs information about customer behavior, its own cost structure, as well as information concerning the competition and the market itself. Furthermore, it requires that prices or quantities can be adjusted in a timely fashion at minimal cost. Until recently, neither was possible. As a result, traditional pricing and inventory control techniques were often static.

The fast development of information technology and the Internet had a dramatic impact on pricing/ replenishment velocity. Thanks to these tools, firms can gather information about customers and competitors; they can also update prices and inventories dynamically at low cost and hence, they allow the sellers to implement dynamic price optimization.

Controlling pricing and inventory control mechanisms has therefore become a crucial tool. This is even more the case in settings where supply is hard to adjust, due to the perishable nature of the products, short selling seasons or long lead times. Early applications of dynamic pricing/ inventory control include industries where short-term supply is hard to change: airlines, cruise boats, hotels, electricity markets. Other industries then realized the benefits of dynamic pricing strategies including retailers in brick-and-mortar stores as well as online. Pricing has become so complex that some industries now outsource their entire pricing and revenue management functions to specialized companies such as Sabre Airlines Solutions for airlines. A number of specialized consultants providing dynamic price optimization solutions were also born,

such as DemandTec, Khimetrics (recently acquired by SAP), Manugistics, ProfitLogic (acquired by Oracle), PROS Revenue Management and Vendavo.
The wealth of information available to sellers has created tremendous opportunities to exploit this data in order to improve decision-making and eventually increase the bottom line. Hence, more and more companies rely on techniques from the fields of data-mining, decision analysis or even artificial intelligence to make predictions and decisions based on the data that is available to them. In particular, it is key for companies to learn the impact of their pricing or inventory strategies on customers. Furthermore, it is crucial for companies to learn the impact of their competitors' strategies on customers, and to incorporate it into their decision process.
Indeed, very few markets are monopolistic. In fact, most developed economies prohibit monopolistic behaviors since they are hurtful to consumers. For instance, the United States adopted the Antitrust Act which legislates against monopolies. On the other hand, models of perfect competition, where neither sellers nor buyers have the power to influence prices, is not a realistic model. Indeed, the assumptions required for a market to be perfectly competitive - atomicity of the buyers and sellers, homogeneity of the products, perfect and complete information, equal access of the firms to technology and resources, absence of barriers to entry - are hard to encounter in real-life settings. Imperfect competition, such as oligopolistic competition, is a model which can be applied to a lot of industries. An oligopolistic market consists of a few firms with a large number of customers, and the firms have the power to influence the price-demand relationship. Due to the Antitrust Act, many industries in developed economies are oligopolistic, e.g consumer goods, cars, airline tickets, wireless communications, power.
The field which studies competitive markets is that of Game Theory, with a focus on equilibrium analysis. In particular, the most polular notion of equilibrium that is studied is that developed by John Nash. A Nash equilibrium is characterized by the fact that no firm has any incentive to unilaterally deviate from the equilibrium strategy. However, Game Theory focusses on agents who exhibit a fully rational behavior. This assumption is generally not met in most real-life applications, especially those involving human interactions and decisions. Furthermore, other critiques of the concept of Nash equilibrium have arisen. For instance, it is not clear how such equilibrium may be reached, or what happens in the case where multiple equilibria exist. As a result of this, there is a need to study markets not only in states of equilibrium, but also in disequilibrium.

1.2 Literature Review

The literature which was useful in the elaboration of this thesis pertains to various fields of Operations Research: the Revenue Management and Pricing literature and the Theory of Learning inspired much of the research. Furthermore, the study of tools from the optimization literature such as variational inequalities and quasi variational inequalities, bilevel programming and Mathematical Programs with Equilibrium Constraints were key to getting insights and making the computational study possible.

1.2.1 Revenue Management and Pricing Literature

There is an extensive literature on the topic of revenue management and pricing, which has gained a lot of interest from academics and researchers in the last several decades. Below, we only list some review papers and books. The recent book by Talluri and van Ryzin [84] provides a review of the theory and practice of Revenue Management from its inception to its latest developments. Review papers include McGill and van Ryzin [65], Bitran and Caldentey [16]. They provide an overview of pricing models in Revenue Management, whereas the survey paper of Weatherford and Bodily [17] concentrates on Revenue Management in the airline industry. We refer the reader to these reviews. In what follows we focus more on the pricing and learning literature.

Elmaghraby and Keskinocak [32] provide a research overview of dynamic pricing. They observe that three main characteristics of the market environment influence the pricing problem: first, whether replenishment of inventory is allowed; second, whether demand arrivals are independent over time; third, whether customers act myopically or strategically.

Dynamic Pricing models with no replenishment, and independent demands over time rely on common assumptions : a market with imperfect competition (e.g monopoly), a finite selling horizon with finite stock and no replenishment. The demand is typically decreasing in price. The goal of the firm in these settings is to maximize the expected profits over the selling horizon.

Gallego and van Ryzin [46] and Feng and Gallego [38] model the demand as a Poisson process with a rate that is decreasing in price. Bitran, Caldentey and Mondschein [14] as well as Bitran and Mondschein [15] consider a demand rate which depends on time and has a known distribution. Lazear [57], Elmaghraby et al. [31] model the demand using reference prices; Lazear's demand model is deterministic whereas Elmaghraby et al. [31] assume that demand is stochastic with a known distribution. Smith and Achabal [80] model demand as depending on price, time and inventory level. Maglaras and Meissner [60] examine two problems: dynamic pricing in a monopoly with imperfect competition and dynamic capacity control with exogenous prices. They show that these problems have a common formulation as a single-resource, single-product pricing problem. In all these papers except for Elmaghraby et al. [31], the customers are assumed to act myopically. Another difference among the aforementioned papers is whether the pricing policy is a discrete-time or a continuous-time policy. Lazear [57] and Elmaghraby et al. [31] study periodic pricing policies where prices are updated at discrete time intervals, whereas Gallego and Van Ryzin [47] and Bitran and Mondschein [15] study continuous-time pricing policies. Some models allow to choose prices in a continuous range, whereas others restrict themselves to a fixed number of price changes, as in Bitran and Mondschein [15], Feng and Gallego [39] and Feng and Xiao [40].

For models of dynamic pricing with inventory replenishment, independent demand

and myopic customers, most of the research focusses on a monopoly market, in a single or multi-product setting. Whereas typical Inventory Management research considers price to be static, and exogenous, the following papers consider joint inventory management and pricing. Federgruen and Heching [37], Zabel [88], and Thowsen [85] address the optimal inventory and pricing of a seller who faces uncertain demand and changes its prices periodically. They find that a base stock list price policy is optimal in a wide range of settings. Rajan et al. [75] focus on changes that occur within an order cycle for a firm selling perishable products. Popescu and Wu [74] study dynamic pricing for customers with repeated interactions: in this setting, customers are sensitive to the pricing history through a reference price. They show that the pricing strategy has long-term implications, in that promotions which increase short-term profits, may decrease future profits. Adida and Perakis [1] propose a nonlinear fluid model for joint dynamic pricing and inventory control with no backorders.

Competition was studied extensively in the traditional Economics literature. The book by Friedman [42] presents the theory of oligopoly, and Vives [87] provides a modern theory of oligopoly using the new tools of Game Theory. Fudenberg and Tirole [43], as well as Maskin and Tirole ([62], [63], [61]) study dynamic oligopoly. Sweezy [83] conjectures a kinked-demand curve in competitive oligopoly. Stigler [82] also focusses on the kinky demand curve and shows the stickiness of prices in an oligopoly. Farahat and Perakis [35] compare total system but also individual seller profits in Cournot and Bertrand oligopoly market settings when demand is an affine function of prices and products are substitutes.

1.2.2 Dynamic Pricing under Competition

Recent work in dynamic pricing considers competitive settings: Dockner and Jørgensen [29] consider optimal pricing strategies in an oligopoly market, but from a marketing perspective. Bernstein and Federgruen [11] built an inventory model for supply chains in an oligopoly, where the decision variables include prices, service levels and inventory control. Kachani and Perakis [51] propose a deterministic fluid model for dynamic pricing in a capacitated, make-to-stock manufacturing system. Perakis and Sood [72] propose a dynamic pricing model and study Nash equilibria in an oligopoly of a single perishable product, while Nguyen and Perakis [69] consider multiple perishable products with shared capacity.
In closed-loop, competitive settings, the notion of Nash equilibrium may be too weak. For instance, the *ultimatum game* is well-known to have Nash equilibria which rely on incredible threats (see for instance Fudenberg and Tirole [43]). As a result, the analysis of dynamic games focusses on a refinement of the concept of Nash equilibrium, called *subgame-perfect equilibrium*. For dynamic games with a Markovian property, subgame-perfect equilibria are called *Markov-perfect equilibria*. Maskin and Tirole [64] review the concept of Markov-Perfect equilibrium in the context of games with observable actions. Dynamic games and subgame-perfect equilibria in the context of market dynamics have been studied in the Economics literature. Several papers address the problem of industry dynamics using subgame perfect equilibria. In par-

ticular, the paper by Ericson and Pakes [33] and its model have been the subject of numerous papers and extensions. Dynamic games have been much less studied in the Revenue Management and Dynamic Pricing literature. Some models of dynamic games, such as in [27] still focus on Nash equilibrium strategies. The paper by Gallego and Hu [45] models the problem presented in Gallego and Van Ryzin [46] in an oligopolistic setting. They investigate a special type of closed-loop policies, and state that a Markov-Perfect equilibrium can be found in principle by solving the system of Hamilton-Jacobi-Bellman equations. However, they do not establish existence of Markov-perfect equilibrium strategies, and do not attempt to characterize them. Krawczyk and Tidball [55] model a dynamic game in seasonal water allocation, and consider open-loop feedback Nash equilibrium strategies. This is the strategies arising in equilibrium, when firms use open-loop feedback strategies. Finally, Jun and Vives [50] study a dynamic infinite-horizon duopoly model with costs of adjustments. Their focus is on stationary policies. They compare open-loop and Markov-perfect equilibrium strategies. Finally, in a recent working paper, Levin, McGill and Nediak [58] prove existence and uniqueness of the subgame-perfect equilibrium in oligopolistic market facing strategic consumers.

1.2.3 Dynamic Pricing with Learning

Traditionally, most publications in revenue management used to postulate a demand model a priori. Hoever, recently, researchers have become more interested in the problem of learning demand. There are two main types of approaches. The bulk of the literature on dynamic pricing with learning takes a Bayesian approach to learning: demand is modeled as a stochastic function, for which the firm is assumed to know the distribution. Only the parameters of this distribution need to be learned. The second approach is a nonparametric, data-driven approach.

Papers considering a Bayesian approach to learning are the following: In their paper [20], Carvahlo and Puterman consider a loglinear demand model with unknown regression coefficients, whereas in [21], they consider a binomial model of demand. Aviv and Pazgal [5] propose a Markov-modulated demand model. The demand is given by a partially observed Markov process, with Bayesian adjustment of the parameters. Aviv and Pazgal [6], and more recently, Araman and Caldentey [3], as well as Farias and van Roy [36], consider a continuous-time model, where demand has an a prior known distribution but the market size parameter is unknown. They show that there is a trade-off between a low price, which yields a loss in revenue, and a high price which lowers the probability of purchase, and slows learning. Petruzzi and Dada [73] consider a model with joint pricing and restocking decisions. Demand has an additive or multiplicative stochastic component, whose distribution is periodically updated through Bayesian updating. Lobo and Boyd [18] present a model with linear demand with additive gaussian stochasticity, and unknown demand parameters who undergo Bayesian learning. They justify price variations in the market by the rational learning behavior of the firms. Balvers and Cosimano [8] focus on a monopoly with stochastic linear demand with unknown intercept and slope. The slope is assumed to have a persistent effect, and thus prompts learning. They define the speed of learning, which is controlled by the firm since it depends on the price and show that learning implies muted responses to changes in demand or market price. Rustichini and Wolinski [79] study a monopoly which faces an uncertain demand, and learns about it through its pricing experience. The demand curve facing the monopoly is not constant and differs from the informed monopoly's policy. They show that even when the rate at which the demand varies is negligible, the stationary probability that the monopoly's policy deviates from the full information counterpart is non negligible. Mirman, Samuelson and Urbano [66] examine a Cournot monopoly in a two-period horizon, facing a stochastic price curve with unknown parameter. They develop conditions under which the firm will find it optimal to adjust its initial price or quantity away from their myopic level in order to increase informativeness of observed market outcomes and thus increase future expected profits. Afeche and Ata [2] propose a Bayesian learning approach to revenue management in a queueing system. Finally Bertsimas and Perakis [24] address dynamic pricing in a monopoly and a duopoly, where demand is modeled by a parametric function whose parameters are a priori unknown, but learned over time through least squares estimation. It is to our knowledge the only paper addressing dynamic pricing with learning in a non monopolistic environment.

The following papers address learning in a nonparametric way. Larson, Olson and Sharma [56] analyze a stochastic inventory control problem where the demand distribution is not known, but learned through a Bayesian approach. But unlike other Baeysian approach, theirs is nonparametric: prior information is given by a Dirichlet process. Besbes and Zeevi [13] propose a blind nonparametric approach to learning where the firm relies on price experimentation to obtain an empirical distribution for the demand function.
Finally, learning is also addressed in the statistics and decision theory literatures: Kalyanam [52] proposes a model that draws on Bayesian estimation, inference and decision theory to learn uncertain demand. Sutton and Barto [9] provide an introduction to reinforcement learning, with wide application areas. Learning also arises in stochastic processes when the parameters are unknown, as in Easley and Kiefer [30].

1.2.4 Variational Inequalities, Quasi Variational Inequalities, Bilevel Programming and MPECs

In this paper, we use variational inequalities (VIs) and quasi variational inequalities (QVIs) to model market equilibria. The monograph by Facchinei and Pang [34] contains the most recent developments concerning VIs. For an introduction to variational inequalities and their applications, we refer the interested reader to the book by Kinderlehrer and Stampacchia [53], which has become a classic in the field. Harker and Pang review the theory, algorithms and applications of finite-dimensional VIs together with complementary problems [49]. Nagurney [68] provides background on VIs and some applications in Economics and Transportation. Bensoussan and Lions

[10] apply variational inequalities to stochastic control.
QVIs are more general than VIs because instead of considering a feasible space, it considers a point-to-set mapping of feasible solutions. This is, for instance, the case in generalized games where the strategy space of each player depends on the other players' strategies. Unlike variational inequalities which have been studied extensively, and for which numerous results have been established concerning existence, uniqueness of solutions, sensitivity analysis, and efficient solutions algorithms, research on QVIs is still in its infancy. From the theoretical standpoint, Chan and Pang [22], Baiocchi and Capelo [7], Tian and Zhou [86], Cubiotti [26] have made the most recent contributions to the field. Harker [48] relates QVIs to generalized Nash games. In oligopolistic competition, QVIs have been introduced by Cardell, Hitt and Hogan [19], by Fukushima and Pang [44] in power markets, by Perakis and Sood [72] and finally by Nguyen and Perakis [69] in a dynamic pricing setting.
Mathematical Problems with Equilibrium Constraints (MPECs) are optimization problems with constraints resulting from an equilibrium problem. They are a generalization of the concept of bilevel problems, where instead of a lower-level optimization problem, there is a lower-level equilibrium problem, possibly formulated as a VI or QVI. MPECs are difficult, non-convex and non-smooth optimization problems. The monograph by Luo, Pang and Ralph [59] provides a good review of the main results concerning MPECs. Bilevel programs were originally developed to model Stackelberg games: the lower-level optimization problem corresponds to the follower's problem, and the upper-level to the leader. Hence, they are useful in the study of oligopoly, as in Murphy, Sherali, Soyster[67], or in problems with staged decisions such as optimal product design in Choi, Desarbo and Harker [23] where the agent decides the product positioning and the optimal price, as well as quality control in services as in Armstrong [4]. The recent survey by Marcotte, Colson and Savard [25] provides an overview of the field of bilevel programming. Just like bilevel problems, MPECs have benefited from advances in nonlinear programming, sensitivity analysis and implicit programming. Fiacco and McCormick [41] use the implicit function theorem to obtain existence of a locally unique Frechet-differentiable solution function to the lower level. Robinson ([76] and [77]) uses generalized equations and fixed-point theorems to study parametric VIs and nonlinear programs. Kojima [54] uses degree theory to establish strong stability of stationary points of parametric nonlinear programs. Finally, in their book dedicated to MPECs, Luo, Pang and Ralph [59] relax some of the conditions under which earlier results were established. Due to their wide applicability, solutions methods and algorithms for MPECs have received a lot of coverage. Nonetheless the computation of global solutions remains elusive, if not impossible.

1.3 Goals and Contributions of the Thesis

The main goal of this thesis is to study learning in competitive markets, using past price or inventory data. Firms may use historical market data to learn the price-demand relationship and learn their optimal pricing or allocation strategy. In stable environments, firms may rely on historical data from prior selling horizons. However,

this historical data may not capture fluctuations in the market, as often occurs in real life. Markets are directly affected by exogenous factors such as prices of commodities, oil prices, changes in interest rates and exchange rates. For instance, transportation industries such as airlines or trucking companies are directly affected by variations in oil prices; consumer goods industries such as packaged foods companies are affected by variations in commodity prices, whereas multinational companies carrying out business abroad must hedge against variations in foreign exchange rates. Furthermore, many markets have a seasonal component. Hence, selling horizons are not identical repetitions of the same game, but have features which make them unique. That is why firms need to incorporate the most up-to-date information, as soon as it becomes available.
We consider two types of learning that occur in the market:

- When pricing (resp. allocation) strategies are at equilibrium in the market, then we isolate the learning of the price-demand relationship, using both historical data from past selling horizons, as well as data from the current selling horizon that becomes available as time goes by. The market is said to be in equilibrium when firms have full knowledge of their competitors' strategy. All the firms act rationally, hence, set their prices (resp. allocations) as best response to their competitors' prices (resp. allocations). Therefore, the strategies used by all firms are the Nash equilibrium strategies.

- When the market has not reached the state of equilibrium, two types of learning take place concurrently: firms not only seek to learn the price-demand relationship, but also the optimal pricing (resp. allocation) strategies for themselves and their competitors. When the market is not in equilibrium, the firms have not learned their competitors' strategies yet. They form a belief regarding their competitors' strategies based on observing their past behaviors, and set their prices (resp. allocation) to be the best response to their belief.

Furthermore, we study learning in two types of markets. In a Bertrand competitive setting, firms compete via prices, that is, at each time period, all firms set prices for the next time period simultaneously. In a Cournot competitive setting, firms compete via allocations, that is, at each time period, all firms set the quantities they allocate to the next time period simultaneously.
It is important to emphasize that Bertrand and Cournot competition are not equivalent. They indeed yield different outcomes. Vives [87] proved that in oligopoly markets for non-differentiated products, the Nash equilibrium prices under Cournot competition are higher than under Bertrand competition. Nothing can be said concerning the equilibrium quantities, unless the firms are identical: in this case, the Nash equilibrium quantities in the Cournot oligopoly are higher than the Bertrand equilibrium quantity. Hence, these two models are not equivalent. They also find applications in different kinds of markets: for instance, the Bertrand model is more relevant in markets where capacity can be changed easily. This is, for instance, the case in consumer goods industries. Farahat and Perakis [35] generalize the analysis of the differences between Bertrand and Cournot oligopolies to the comparisons of

the profits. They prove that when the demand function is affine, Cournot profits are almost always higher than Bertrand profits and characterize when this relationship does not hold. Hence, there is a clear need to study the two types of competition separately.

A further difference between the Cournot and Bertrand model is their application. The Cournot model of competition is well adapted to industries where capacity is difficult to adjust, such as power markets, the airlines, hotel, equipment rentals and cruise lines industries. In the power industry, storage of capacity is impossible, whereas in the travel and rental industries, capacity is considered to be fixed and a sunk cost once the capacity investment (e.g. planes, cars and ships) is made.

However, the bulk of the literature on revenue management in competitive environment focusses on the study of the state of equilibrium. One of the criticisms of game theory and the notion of Nash equilibrium is that it assumes full rationality of the agents whose behavior is modeled. This implies that all agents have the same capability to analyze the game and compute equilibria. However, if only one of the agents does not follow the equilibrium strategy, then the other players are not guaranteed superior payoffs by using the Nash equilibrium strategies. In this thesis, we tackle bounded rationality of firms, by considering competitive markets in disequilibrium, as well as in equilibrium.

Another contribution of this thesis is to model time-varying consumer behavior. Indeed, in many industries for perishable products, such as consumer goods, fashion items, travel or entertainment, different types of customers buy at different instants in the selling horizon. For instance, in goods that perish due to obsolescence, early adopters who are generally not price sensitive, purchase early. On the other hand, followers wait for the technology to become more obsolete and take advantage of cheaper prices. Similarly, fashion-sensitive consumers purchase at the beginning of the season, whereas price-sensitive customers look for promotional sales and bargains that mostly occur towards the end of the season. In the travel or entertainment industries, the opposite behavior is observed: price-sensitive customers tend to book early, before the inventory of lower-priced fares or tickets is depleted, whereas less price-sensitive customers can afford to wait. In this work, the parametric model of price-demand relationship allows for time-varying sensitivities.

However, unlike most models in the literature, our model does not assume that the price-demand relationship is known to the firms a priori. We only assume that we know the parametric family it belongs to. Indeed, only recently has the problem of learning the price-demand relationship started to be addressed by researchers. We propose in our research a data-driven approach to learning the price-demand relationship. The firms use historical data, as well as the most recent market information, in order to update its estimation of the price-demand relationship. The firms do so in an online manner: the dynamic pricing (resp. allocation) problem and the learning of the price-demand relationship are performed concurrently. As a result of this, the problem has a bilevel programming or MPEC (Mathematical Problem with Equilibrium Constraints) formulation.

In summary, in this thesis, we make the following contributions:

- Capacitated markets: we study markets in which firms have a limited capacity which they must allocate over the selling horizon, with no replenishment or back orders. Capacity constrained problems are usually harder to analyze than non-capacitated ones. In particular, under Bertrand competition, capacity constraints make the problem significantly harder, since they turn the market equilibrium into a generalized Nash equilibrium problem.

- Competitive markets: we explicitly incorporate the impact of competitors' actions in our revenue management model. However, we do not assume that all the firms are necessarily fully rational and that the market is operating in a state of equilibrium. Indeed, we study both the state of disequilibrium where a firm's strategy is the best response to its belief regarding its competitors' strategy, and the state of equilibrium, where all firms use Nash equilibrium strategies.

- Unknown price-demand relationship: we assume that only the parametric form of the price-demand relationship is known, but model how firms learn the parameters over time, using an online data-driven approach. Online learning is required when the firms operate in the fluctuating market environments. Our approach can be applied with various forms of demand functions, and is therefore fairly general. Furthermore, it captures time-varying customer behavior by allowing the parameters in the model to vary with time.

- Joint pricing (resp. allocation) and learning: we address the two problems jointly, and towards this end, we use concepts from the areas of bilevel programming and MPECs. Furthermore, our approach is the first one to the best of our knowledge which addresses two types of learning concomitantly, that is, learning the equilibrium strategies, while learning the price-demand relationship. We prove that in the long run, the two types of learning are achieved.

1.4 Outline of the Thesis

The thesis is organized as follows: In Chapter 1, we give the motivation behind the problem and review the literature that is relevant to the thesis. In Chapter 2, we introduce the problem of dynamic pricing with demand learning under Bertrand competition, and dynamic allocation with learning under Cournot competition. In each case, we introduce the problem in both disequilibrium and equilibrium state. In Chapter 3, we focus on the dynamic pricing (resp. allocation) problem for known price-demand relationship, i.e when the parameters of the demand (resp. price) function are known, both in disequilibrium and equilibrium. We show existence and uniqueness of solution to the best response and equilibrium problems, and perform sensitivity analysis on the solutions, when the value of the parameters changes. In Chapter 4, we focus on the joint dynamic pricing (resp. allocation) with learning, and show how the joint problem can be formulated: as a bilevel problem, or equivalently a mixed integer program in disequilibrium, and as an MPEC, or equivalently a mixed integer program in equilibrium. In Chapter 5, we establish a key result in the thesis and prove it: learning of the price-demand relationship and of the equilibrium strategies occurs in the long run. Chapter 6 is the algorithmic and computational part of the thesis. It addresses the problem of finding algorithms to solve the problem, and implements the method on data. Chapter 7 investigates closed-loop policies in a stochastic setting. It establishes existence and uniqueness of Markov-perfect strategies, and approximate solutions. Finally, in Chapter 8, we present our conclusions on this research, and elaborate on further avenues of research related to this thesis. Appendix A contains the notation used throughout the thesis, whereas Appendix B contains all figures and graphs. The bibliography of the thesis is a list of the literature referred to in this thesis.

Chapter 2

Model Formulation

In this chapter, we introduce the notation and terminology that will be used throughout the thesis. We present the dynamic control with learning model for the Bertrand and for the Cournot competitive markets in two steps. First, we model the price-demand relationship as a parametric function.

Under Bertrand competition, we model demand for each firm, at each time period as a parametric function of the prices of all the firms in the market. The parameters aim at quantifying the impact of the firm's quality, as well as the impact of its own and its competitors' prices on its own demand. The latter parameters are therefore called price sensitivities. Under Cournot competition, we model a firm's price at each time period as a parametric function of the quantities of all the firms in the market. The parameters quantify the impact of the firm's quality, and of its own and its competitors' quantities on its own price.

Second, we present the dynamic control problem for known price-demand relationship. Under Bertrand competition, we formulate the dynamic pricing problem for known demand function. Each firm aims at maximizing its total revenue over the selling horizon, given its competitors' pricing strategy. This problem is called the Bertrand best-response problem, and the optimal prices are called best-response prices. We formulate the best-response problem as a nonlinear optimization problem. Under Cournot competition, we formulate the dynamic allocation problem for known price function. The Cournot best-response problem is that of each firm maximizing its total revenue over the selling horizon, given its competitors' allocation strategy. The solution to this problem are the best-response quantities or allocations.

We then focus on the market as a whole and address questions regarding the state of the market, i.e whether the market is in a state of equilibrium or disequilibrium. The notion of equilibrium used here is that of Nash equilibrium, whereby no one firm has incentive to unilaterally deviate from the equilibrium strategy.

In disequilibrium, each firm's strategy is the best-response to its belief regarding its competitors' strategy. This is motivated by the game theoretical concept of *rationalizable strategies*. In game theory, a strategy is called rationalizable if it is the best-response to some belief regarding competitors' strategies. Rationalizable strategies are related to strictly dominated strategies, in the sense that a rational firm will never use a strictly dominated strategy. Hence, the set of rationalizable strategies is

contained in the set of strategies which survive iterated strict dominance In fact, the two notions are equivalent in duopolies.
In equilibrium, each firm, on top of acting optimally, knows that its competitors also act optimally, therefore each firm's strategy is the best-response to the optimal strategy of their competitors. Hence, the Nash equilibrium is the set of strategies simultaneously solving the best-response problems of each firm. We model the market equilibrium via a quasi-variational inequality in the Bertrand market, and as a variational inequality in the Cournot market. In subsequent chapters, we address the issues of existence and uniqueness of equilibrium in more details.

2.1 Notation and Terminology

Throughout this thesis, we use the following notation:

- N: number of firms competing in the market;
- T: selling horizon after which all unused capacity is lost;
- $h = -H, \ldots, 0$: superscript indicating the selling horizon: $h = 0$ denotes the current selling horizon, and $h < 0$ denote historical selling horizons.
- $C_i, i = 1, \ldots, N$: finite inventory of each firm over the selling horizon; we call this inventory the total capacity of the firm.
- For a vector x with components $x_i^t,\ i = 1, \ldots, N,\ t = 1, \ldots, T$, we denote by $\mathbf{x_i} = (x_i(1), \ldots, x_i(T))$ the subvector corresponding to firm i, and by $\mathbf{x(t)} = (x_1(t), \ldots, x_N(t))$ the subvector corresponding to time period t.
- $\widehat{p_i}^0(t)$: price set by firm i at period t in the current selling horizon; $\widehat{p_i}^h(t)$: historical price of firm i for period t.
- $\widehat{q_i}^0(t)$: quantity set by firm i at period t in the current selling horizon; $\widehat{q_i}^h(t)$: historical quantity of firm i for period t.
- $\overline{\mathbf{p_{-i}}}$: price vector corresponding to firm i's belief at time t regarding its competitors' pricing strategy.
- $\overline{\mathbf{q_{-i}}}$: allocation vector corresponding to firm i's belief at time t regarding its competitors' allocation strategy.

2.2 Problem in Bertrand Competition

Under Bertrand competition, firms compete by setting prices. At each time period of the selling horizon, each firm updates its pricing decision for future periods. The quantities sold at each time period by each firm are functions of their own and their competitors' prices, as prescribed by the demand function.
We make the following informational assumptions:

- At each time period t in the selling horizon, each firm has observed all the past prices in the market, that is the prices that were set by all the firms in the selling horizon, in periods 1 to $t-1$. For instance, at time $t = 10$ of the selling horizon, firms have observed prices for periods 1 to 9.

- Additionally, the firms have collected *historical prices*, i.e prices that were used in past selling horizons. For instance, consider a New-York- San Francisco flight departing Monday, November 27, 2006. Then historical prices are prices for New York - San Francisco flights departing on Mondays in Fall 2006.

- Firms know the total capacity of their competitors, however, they cannot observe the remaining capacity of their competitors within the selling horizon. In other words, they do not know how much their competitors sell at each period in the selling horizon.

In disequilibrium state, each firm wants to set its prices optimally, given its belief concerning its competitors' pricing policy. This belief is based on the historical data collected by each firm concerning their competitors. In other words, each firm wants to find the vector of prices which maximize its total revenue over the selling horizon, and such that the total quantity sold over the horizon does not exceed capacity. Two forms of learning occur in this state:

- Learning of the demand function, or in other words, learning of the price sensitivities: the parameters of the demand function are initially unknown to the firm, which would like to learn them over the selling horizon. Indeed, as time elapses, the firm observes more information which can be used to achieve more accurate knowledge of the demand function. In particular, at each time period of the horizon, each firm observes the prices set by itself and its competitors in the previous period, and can incorporate this additional piece of information into its estimation of the price sensitivities.

- Learning of the market equilibrium strategies: when the market is in disequilibrium, each firm's strategy consists in setting prices that are best-response to its belief regarding its competitors' strategy. The firms form a belief of their competitors' strategy by observing past prices. Hence, after each additional price observed, firms update their belief, and compute their best-response to this new belief. This strategy is very much akin to learning strategies in Game Theory such as fictitious play or tatonnement. Eventually, through this dynamic strategic updating process, the firms' strategies converge to equilibrium strategies.

In what follows, we first specify the demand model. Then we focus on the dynamic pricing with learning in disequilibrium state. Finally, we address the dynamic pricing with learning in equilibrium.

2.2.1 The Demand Model

Demand is modeled as a parametric function of prices. For instance, demand may be a linear, loglinear, exponential function of prices. Let $q_i(p_i(t), p_{-i}(t), \beta_i(t))$ be the demand of firm i at time t. We assume that demands are independent over time, that is, demand in period t solely depends on the prices set in that period. For each time period t, we can write the price-demand relationship for all the companies as a mapping $\mathbf{q}(\mathbf{p}(\mathbf{t}), \beta(\mathbf{t}))$.
We discuss the demand function in further details in Chapter 3, where we analyze the best-response problem. In particular, we define the feasible set of parameters, such that the demand function satisfies certain properties.
For now, we denote by $\mathcal{B}_i$ the set of feasible parameters for firm i.

2.2.2 The Dynamic Pricing Problem in Disequilibrium

In this section, we focus on the problem faced by the firm at the beginning of selling period $t \in \{2, \ldots, T\}$ when the market is in disequilibrium. In this state, each seller computes its best-response to its beliefs concerning its competitors' strategy.
To provide a better understanding of the process each firm may follow to determine its optimal pricing strategy and learn the form of the demand function, we decompose the problem facing each firm at each time period of the horizon into three steps. However, in reality, these three steps are performed simultaneously by the firm. In Step 1, assuming the price sensitivities are known, each firm would like to find the best-response prices, for its belief of its competitors' strategy. In Step 2, each firm would like to estimate the price sensitivities. Finally, in Step 3, each firm seeks to determine its optimal pricing strategy for future periods.

Step 1: Computation of the Best-Response Prices

In Step 1, we assume that the price sensitivities β are known. In particular, in the disequilibrium state, we assume that each firm knows its own price sensitivities $\beta_\mathbf{i}$.
Each firm i seeks to find its best-response prices. As each firm is a revenue maximizer, then for fixed prices of its competitors $\overline{\mathbf{p_{-i}}}$, its best-response prices are those which maximize its total expected revenue $\pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ over the selling horizon. Each firm therefore solves a best-response problem $\mathcal{BR}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$:

$$\max_{\mathbf{p_i}} \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \quad = \sum_{t=1}^{T} p_i(t) q_i(p_i(t), \overline{p_{-i}(t)}, \beta_i(t)) \tag{2.1}$$

$$\text{s.t} \quad \sum_{t=1}^{T} q_i(p_i(t), \overline{p_{-i}(t)}, \beta_i(t)) \le C_i \tag{2.2}$$

$$0 \le p_i(t) \le p_i^{max} \tag{2.3}$$

The objective function (2.1) is the total expected revenue of the firm over the selling horizon for fixed pricing strategy $\overline{\mathbf{p_{-i}}}$ of i's competitors. Constraint (2.2) incorporates the fact that the problem we are addressing is capacitated, that is, the total quantity sold by the firm cannot exceed its overall capacity. Constraint (2.3) is the set of

feasible prices.
The solution to the best-response problem is the best-response policy of firm i as a function of its competitors' strategy, denoted $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$.

Notice that two cases may arise concerning the capacity constraint:

- The constraint is active or binding: in other words, capacity is a scarce resource. Since the capacity constraint links the demand in different time periods together, the best-response price in any period t will depend in this case on the parameters of all time periods $\beta_i(1), \ldots, \beta_i(T)$.

- The constraint is inactive or non binding: in other words, capacity is abundant. In this case, the best-response problem is separable into T instantaneous non capacitated best-response problems, one for each time period. The best-response price in each period t then depends only on the parameters at time t,

In reality, firm i solves its best-response problem for its belief regarding its competitors' strategy. Hence, $\overline{\mathbf{p_{-i}}}$ is firm i's belief concerning its competitors's strategies. A firm might form its belief regarding its competitors' strategy by observing its competitors' historical prices in previous selling horizons, as well as its competitors past prices in previous time period of the current selling horizon. At time t in the selling horizon, a firm has observed prices that were set by its competitors in past periods of the current horizon, and may use these as its belief regarding competitors' past prices. However, it has not observed future prices, and must therefore rely on historical prices from previous selling horizon. Examples of such beliefs are as follows:

1. $$\overline{p_{-i}}(s) = \begin{cases} \widehat{p_{-i}}^0(s) & \text{if } s \leq t-1 \\ \frac{1}{H}\sum_{h=-H}^{-1} \widehat{p_{-i}}^h(s) & \text{if } s \geq t \end{cases}$$

   In other words, at time t, firm i's belief of its competitors' prices for a past period $s < t$ is equal to the price set at s by the competitors, whereas its belief for a future period $s \geq t$, is the arithmetic average of the historical prices set at s by its competitors in previous selling horizons $h = -H, \ldots, -1$.

2. Let $(\omega^h)_{h=-H,\ldots,0}$ and $(\tau^h)_{h=-H,\ldots,-1}$ be series of nonnegative weights such that $\sum_{h=-H}^{0} \omega^h = 1$ and $\sum_{h=H}^{-1} \tau^h = 1$:

   $$\overline{p_{-i}}(s) = \begin{cases} \sum_{h=-H}^{0} \omega^h \widehat{p_{-i}}^h(s) & \text{if } s \leq t-1 \\ \sum_{h=-H}^{-1} \tau^h \widehat{p_{-i}}^h(s) & \text{if } s \geq t \end{cases}$$

   In other words, at time t, firm i's belief of its competitors' prices for a past period $s < t$ is equal to the weighted average of the prices set at s by the competitors in the current horizon and in past horizons, whereas its belief for a future period $s \geq t$, is the weighted average of the historical prices set at s by its competitors in previous selling horizons $h = -H, \ldots, -1$.

Step 2: Estimation of the Price Sensitivities

In this section, we address the estimation problem that each firm solves at each time period t of the selling horizon. At time t, each firm wants to update its estimate of the price sensitivities, based on the market information gathered up to $t-1$. As the sellers are assumed to be revenue maximizers, then the best-response prices that were found in Step 1 should be good estimates of the prices observed on the market. Therefore, the price sensitivities $\beta_{\mathbf{i}}$ ought to be such that the best-response prices $\mathcal{P}_i(s, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$ match the observed market prices. In other words, the parameters $\beta_{\mathbf{i}}$ should minimize the error between the observed market prices, and the price estimate given by the solution to each firm's best-response problem. By *error*, we mean the distance between the two prices, i.e either the absolute value error if the L_1 norm is used, or the squared error is the L_2 norm is used.
Note that since the best-response price in any period s, $\mathcal{P}_i(s, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$, depends in general on the parameters of all time periods, it follows that at time t, the firm must estimate the parameters of all time periods. Two cases arise, depending on whether the firm compares prices in past periods $s < t$, or in future periods $s \geq t$. Indeed, the information sets for $s < t$ and $s \geq t$ differ.

1. For $s < t$, the information set of firm i consists of the prices set in the current horizon at period s, and the prices set in historical horizons in period s. Hence, the estimation problem consists of minimizing the following errors:
    - For prices set in the current horizon the absolute value error $|\widehat{p_i}^0(s) - \mathcal{P}_i(s, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})|$ (resp. the squared error $\left(\widehat{p_i}^0(s) - \mathcal{P}_i(s, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})\right)^2$) where $\overline{\mathbf{p_{-i}}}$ is firm i's belief at time t regarding its competitors' strategy;
    - For prices set in the historical horizons h, the absolute value errors $|\widehat{p_i}^h(s) - \mathcal{P}_i(s, \mathbf{p_{-i}}(\mathbf{h}), \beta_{\mathbf{i}})|$ (resp. the squared error $\left(\widehat{p_i}^h(s) - \mathcal{P}_i(s, \mathbf{p_{-i}}(\mathbf{h}), \beta_{\mathbf{i}})\right)^2$).
2. For $s \geq t$, the information set of firm i consists only of the prices set in historical horizons in period s. Hence, the estimation problem consists of minimizing the error between historical prices and the best-response prices in absolute value $|\widehat{p_i}^h(s) - \mathcal{P}_i(s, \mathbf{p_{-i}}(\mathbf{h}), \beta_{\mathbf{i}})|$ (resp. the squared $\left(\widehat{p_i}^h(s) - \mathcal{P}_i(s, \mathbf{p_{-i}}(\mathbf{h}), \beta_{\mathbf{i}})\right)^2$).

The estimation problem at time t can therefore be written in absolute value terms as:

$$\min_{\beta_i \in \mathcal{B}_i} \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} |\widehat{p_i}^h(s) - \mathcal{P}_i(s, \mathbf{p_{-i}}(\mathbf{h}), \beta_{\mathbf{i}})| \} + \sum_{s=1}^{T} |\widehat{p_i}^0(s) - \mathcal{P}_i(s, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})| \tag{2.4}$$

Alternatively, the estimation problem at time t in squared error is:

$$\min_{\beta_i \in \mathcal{B}_i} \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} \left(\widehat{p_i}^h(s) - \mathcal{P}_i(s, \mathbf{p_{-i}}(\mathbf{h}), \beta_{\mathbf{i}})\right)^2 \} + \sum_{s=1}^{T} \left(\widehat{p_i}^0(s) - \mathcal{P}_i(s, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})\right)^2 \tag{2.5}$$

We denote by $\widehat{\beta_{\mathbf{i}}}^t$ the vector of optimal parameters computed given the information available at time t.

Step 3: Optimal Pricing Policy

In summary, at each period t of the selling horizon, firm i finds its best-response prices $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$, as a function of the unknown price sensitivities $\beta_{\mathbf{i}}$ and competitors' strategy in Step 1, and in Step 2, it finds the optimal estimates of the price sensitivities $\widehat{\beta_{\mathbf{i}}}^t$ given the information available at that time. Combining the results of both steps, we obtain that the vector $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \widehat{\beta_{\mathbf{i}}}^t)$ is the optimal pricing strategy at time t for the selling horizon.

2.2.3 The Dynamic Pricing Problem in Equilibrium

In this section, we investigate the dynamic pricing problem in the ideal case where the market has reached equilibrium. This means that for known price sensitivities β, learning of the optimal pricing strategy has been achieved. Indeed, recall that at each time period t and for known parameters β, each firm i selects its strategy as the best-response to its belief concerning its competitors' strategy. Firms' beliefs are updated at each period to incorporate the most current information. A similar iterative learning process was introduced for instance in Perakis and Sood [72]. They proved its convergence in a robust setting, i.e when the parameters $\beta_{\mathbf{i}}$ are unknown, but belong to an uncertainty set, and the firms are conservative and maximize their worst-case revenue. The intuition is that, if in the long run, the strategies converge to some $\mathcal{P}^*(\beta)$, then we have:

$$\mathcal{P}_i^*(\beta) \in \mathcal{BR}_i(\mathcal{P}_{-i}^*, \beta_{\mathbf{i}}) \quad \forall\ i$$

Hence, $\mathcal{P}^*(\beta)$ is a fixed point of the best-response mapping and is therefore a Nash equilibrium: assuming firm i's competitors set prices equal to the Nash equilibrium prices $\mathcal{P}_{-i}^*(\beta)$, then it is optimal for i to price at $\mathcal{P}_i^*(\beta)$, and it cannot gain from unilaterally deviating from the Nash policy.

As learning of the pricing strategies has already occurred in the equilibrium state, we can isolate the learning of the demand function. In equilibrium, the three-step problem which faces each firm at each time period is the following: in step 1, each firm aims at finding the Nash equilibrium prices in the market, assuming the price sensitivities are known. In step 2, each firm estimates its own as well as its competitors' price sensitivities. Finally, using the solutions of steps 1 and 2, each firm finds its optimal pricing policy.

As we will see shortly, all firms' problems at equilibrium are coupled, thus each firm has to solve simultaneously the problem for itself and its competitors. This makes the equilibrium problem more complex than its transient counterpart.

Step 1: Computation of the Nash Equilibrium Prices

When the market is in equilibrium, the firms' prices are fixed points of the best-response mapping. In other words, the prices simultaneously solve the N best-response problems. Let $\mathcal{GNE}(\alpha)$ denote the generalized Nash equilibrium in the market, when price sensitivities are equal to β:

A vector $\mathcal{P}(\beta)$ of prices solves the generalized Nash equilibrium iff $\mathcal{P}_i(\beta)$ solves $\mathcal{BR}_i(\mathcal{P}_{-i}(\beta))$ for all $i = 1, \ldots, N$.
Such an equilibrium is called a *generalized equilibrium* since each firm's strategy space depends on its competitors' strategy: indeed, the set of feasible prices for each firm depends on its competitors' prices through the capacity constraint:

$$\sum_{t=1}^{T} q_i(p_i(t), \overline{p_{-i}}(t), \beta_i(t)) \leq C_i$$

This is to be distinguished from the usual notion of Nash equilibrium, which corresponds to games for which each player's strategy space does not depend on its competitors' strategy.
In Chapter 4, we cover in more details the generalized Nash equilibrium. In particular, we investigate conditions for existence and uniqueness of the equilibrium.
<u>Remarks:</u>

1. Notice that since the Nash equilibrium prices simultaneously solve all firms' best-response problems, then the equilibrium prices of firm i depend not only on its own price sensitivities $\beta_{\mathbf{i}}$, but also on its competitors' sensitivities $\beta_{-\mathbf{i}}$. As a result, the problem of estimating firm i's prices is non-separable from firm j's estimation problem.

2. Notice also, that the Nash equilibrium prices for period s depend in general on the price sensitivities of all time periods, since the time periods are coupled by the capacity constraint.

Step 2: Estimation of the Price Sensitivities

Due to remarks 1 and 2, it appears that, contrary to the disequilibrium state, firm i cannot estimate its own price sensitivities independently of firm j. As a result, each firm has to jointly estimate the price sensitivities of all the firms in the market.
One possibility is to write the joint estimation as the simultaneous solution to the N individual estimation problems of each firm. The optimal vector of parameters for all firms and all time periods, denoted by $\widehat{\beta}^t$ is the simultaneous solution to:

$$\min_{\beta_i \in \mathcal{B}_i} \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} \left(\widehat{p_i}^h(s) - \mathcal{P}_i(s, \beta_{\mathbf{i}}, \widehat{\beta_{-\mathbf{i}}}) \right)^2 \} + \sum_{s=1}^{T} \left(\widehat{p_i}^0(s) - \mathcal{P}_i(s, \beta_{\mathbf{i}}, \widehat{\beta_{-\mathbf{i}}}) \right)^2 \quad (2.6)$$

However, another way to perform joint estimation, which renders the problem simpler from a mathematical perspective, is to choose as estimates the parameters which minimize the sum, for all firms and all periods, of all the errors between observed prices and equilibrium prices. Hence, the equilibrium estimation problem is formulated as:

$$\min_{\beta \in \mathcal{B}_1 \times \ldots \times \mathcal{B}_N} \sum_{i=1}^{N} \left\{ \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} (\widehat{p_i}^h(s) - \mathcal{P}_i(s, \beta_{\mathbf{i}}, \beta_{-\mathbf{i}}))^2 \} + \sum_{s=1}^{T} (\widehat{p_i}^0(s) - \mathcal{P}_i(s, \beta_{\mathbf{i}}, \beta_{-\mathbf{i}}))^2 \right\} \quad (2.7)$$

The above estimation problem differs from the transient case in two ways: first, the objective function corresponds to the sum of each firm's objective function in the transient case. Moreover, the prices as functions of the unknown sensitivities are Nash equilibrium prices, instead of best-response prices for a given competitors' strategy.
The solution to the estimation problem are the optimal time-t estimators $\widehat{\beta}^t$ (in the least square sense) of the price sensitivities for all firms and all time periods.

Step 3: Optimal Pricing Policy

In summary, at each period t in the selling horizon, each firm finds the Nash equilibrium prices, as a function of the unknown price sensitivities β in step 1, and finds the optimal estimates of the price sensitivities $\widehat{\beta}^t$ in step 2. Similarly to the disequilibrium state, the solutions to steps 1 and 2 yield each firm's optimal pricing strategy at time t.

2.2.4 Reformulations of Steps 1 and 2

In Chapter 4, we show that Steps 1 and 2 under Bertrand competition can be reformulated as one problem. Indeed, we show that the Bertrand best-response problem with learning can be reformulated as a bilevel problem, or equivalently as a mixed integer program. The Bertrand equilibrium problem can be reformulated as a Mathematical Program with Equilibrium Constraints (MPEC) or equivalently as a mixed integer program.
To achieve this, we will first focus on Step 1 in Chapter 3. We will then use the insights gained on Step 1 to consider Steps 1 and 2 together in Chapter 4.

2.3 Problem in Cournot Competition

In a Cournot competitive environment, firms compete via allocations, and not prices: each firm's price at each time period is determined as a function of its own, and its competitors' allocation decisions, as prescribed by the inverse demand function, also called price function.
The information observed by the firms at each time period are the quantities set by themselves and their competitors in past time periods of the current selling horizon, supplemented by historical quantities from past selling horizons.
In disequilibrium state, each firm wants to set its quantities optimally, given its belief concerning its competitors' allocation policy. This belief is based on the historical data collected by each firm concerning their competitors. In other words, each firm wants to find the vector of allocations which maximizes its total revenue over the selling horizon, and such that the total quantity sold over the horizon does not exceed capacity. Two forms of learning occur in this state:

- Learning of the price function through the allocation sensitivities. At each time period of the horizon, each firm observes the prices resulting from the allocation decisions in the previous period, and can incorporate this additional piece of information into its estimation of the price sensitivities.

- Learning of the market equilibrium: when the market is in disequilibrium, each firm's strategy consists of setting quantities that are the best-response to its belief regarding its competitors' strategy. The firms form a belief of their competitors' strategy by observing historical prices. Hence, at each additional price observed, firms update their belief, and compute their best-response to this new belief. This strategy is very much akin to learning strategies in Game Theory such as fictitious play or tatonnement. Eventually, through this dynamic strategic update process, the firms' strategies converge to equilibrium strategies.

In what follows, we first specify the price-demand relationship. Then we focus on the dynamic allocation with learning in disequilibrium state. Finally, we address the dynamic allocation with learning in equilibrium.

2.3.1 The Price-Demand Relationship

In this framework, the price-demand relationship is expressed though an inverse demand function, or price function: $p_i(q_i(t), q_{-i}(t), \alpha_i(t))$. The parameters of the price function are called allocation sensitivities, since they specify how the price varies when the allocation policy varies.
We discuss the inverse demand function in further details in Chapter 3, where we analyze the best-response problem. In particular, we define the feasible set of parameters, such that the inverse demand function satisfies certain properties.
For now, we denote by $\mathcal{A}_i$ the set of feasible parameters for firm i.

2.3.2 The Dynamic Allocation Problem in Disequilibrium

In this section, we focus on the dynamic allocation problem that each firm faces at each time period $t = 2, \dots, T$ of the selling horizon, when the market is in disequilibrium.
Firms seek to find their optimal allocation strategy, given their belief concerning their competitors' strategy. This implies that they have formed beliefs concerning the allocations of their competitors: using the historical allocations data from past selling horizons, as well as the market allocations observed in the current horizon so far, each firm infers its competitors' allocation policy. Each firm aims at achieving two types of learning: it seeks to learn the optimal allocation policies for itself and it competitors, as well as to learn the price-demand relationship by learning the allocation sensitivities of the price functions. At each period t, firm i's problem can be decomposed in three steps: in step 1, firm i determines its best-response quantities to its beliefs concerning its competitors' quantities, assuming its allocation sensitivities are known. In step 2, it estimates its allocation sensitivities, and finally, in step 3, it sets its optimal policy for future periods.

Step 1: Computations of the Best-Response Quantities

We assume that the allocation sensitivities α are known. Each firm i seeks to find its best-response quantities to its competitors' allocation policy $\overline{\mathbf{q_{-i}}}$: firm i's best-response allocations are those which maximize its total expected revenue $\pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$ over the selling horizon. Each firm therefore solves a best-response problem $\mathcal{BR}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$:

$$\begin{aligned} \max_{\mathbf{q_i}} \pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i}) \quad &= \textstyle\sum_{t=1}^{T} q_i(t) p_i(q_i(t), \overline{q_{-i}(t)}, \alpha_i(t)) && (2.8) \\ \text{s.t} \qquad & \textstyle\sum_{t=1}^{T} q_i(t) \leq C_i && (2.9) \\ & 0 \leq q_i(t) && (2.10) \end{aligned}$$

The objective function is the total expected revenue over the selling horizon, for fixed strategy of its competitors. Constraint (2.9) is the capacity constraint, which prescribes that the total quantity sold over the selling horizon should not exceed the capacity. Finally, constraint (2.10) restricts the allocations to be nonnegative. In Chapter 3, we will establish conditions under which this formulation is sufficient to ensure nonnegativity of the prices without introducing further constraints.
We denote by $\mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$ the best-response quantities of firm i, for fixed competitors' strategy $\overline{\mathbf{q_{-i}}}$.
Two cases may arise when solving the best-response problem:

- The capacity constraint is active or binding, i.e capacity is scarce. In this case, the best-response quantity in any period depends on the parameters in all time periods;

- The capacity constraint is inactive, i.e capacity is abundant. In this case, the problem is equivalent to a non-capacitated problem, and is separable in time. Hence, the best-response quantity in any period only depends on the parameters in that period.

In practice, each firm solves its best-response problem, for its belief regarding its competitors' allocation strategy, based on historical allocations from previous selling horizons, as well as the observed market allocations in the previous periods of the current horizon. Examples of such beliefs are as follows:

1. $$\overline{q_{-i}}(s) = \begin{cases} \widehat{q_{-i}}^0(s) & \text{if } s \leq t-1 \\ \sum_{h=-H}^{-1} \frac{1}{H} \widehat{q_{-i}}^h(s) & \text{if } s \geq t \end{cases}$$
   In other words, at time t, firm i's belief of its competitors' quantities for a past period $s < t$ is equal to the price set at s by the competitors, whereas its belief for a future period $s \geq t$, is the arithmetic average of the historical quantities set at s by its competitors in previous selling horizons $h = -H, \ldots, -1$.

2. Let $(\omega^h)_{h=-H,\ldots,0}$ and $(\tau^h)_{h=-H,\ldots,-1}$ be series of nonnegative weights such that $\sum_{h=-H}^{0} \omega^h = 1$ and $\sum_{h=H}^{-1} \tau^h = 1$:
   $$\overline{q_{-i}}(s) = \begin{cases} \sum_{h=-H}^{0} \omega^h \widehat{q_{-i}}^h(s) & \text{if } s \leq t-1 \\ \sum_{h=-H}^{-1} \tau^h \widehat{q_{-i}}^h(s) & \text{if } s \geq t \end{cases}$$

Step 2: Estimation of the Allocation Sensitivities

In this section, we address the estimation problem that each firm solves at each time period t of the selling horizon. At time t, each firm wants to update its estimate of the allocation sensitivities, based on the market information gathered up to $t-1$. As the sellers are assumed to be revenue maximizers, then the best-response allocations which were computed in Step 1 should be good estimates of the quantities observed on the market. Therefore, the allocation sensitivities $\alpha_{\mathbf{i}}$ ought to be such that the best-response allocations $\mathcal{Q}_i(s, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}})$ match the observed market quantities. In other words, the parameters $\alpha_{\mathbf{i}}$ should minimize the error between the observed market allocations, and the price estimate given by the solution to each firm's best-response problem, in terms of absolute value error or squared error.
Hence, the estimation problem under Cournot competition is similar to that under Bertrand competition, where quantities are substituted for prices. The estimation problem at time t can therefore be written in absolute value terms as:

$$\min_{\alpha_i \in \mathcal{A}_i} \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} |\widehat{q_i}^h(s) - \mathcal{Q}_i(s, \mathbf{q_{-i}(h)}, \alpha_{\mathbf{i}})| \} + \sum_{s=1}^{T} |\widehat{q_i}^0(s) - \mathcal{Q}_i(s, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}})| \tag{2.11}$$

Alternatively, the estimation problem at time t in squared error is:

$$\min_{\alpha_i \in \mathcal{A}_i} \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} \left(\widehat{q_i}^h(s) - \mathcal{Q}_i(s, \mathbf{q_{-i}(h)}, \alpha_{\mathbf{i}}) \right)^2 \} + \sum_{s=1}^{T} \left(\widehat{q_i}^0(s) - \mathcal{Q}_i(s, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}) \right)^2 \tag{2.12}$$

Alternatively, the square operator in the objective function can be replaced by the absolute value operator, to yield the estimation problem in absolute value error. We denote by $\widehat{\alpha_{\mathbf{i}}}^t$ the vector of optimal parameters computed given the information available at time t.

Step 3: Optimal Allocation Policy

In summary, at each period t of the selling horizon, firm i finds its best-response quantities $\mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}})$, as a function of the unknown allocation sensitivities $\alpha_{\mathbf{i}}$ and competitors' strategy in Step 1, and in Step 2, it finds the optimal estimates of the allocation sensitivities $\widehat{\alpha_{\mathbf{i}}}^t$ given the information available at that time. Combining the results of both steps, we obtain that the vector $\mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \widehat{\alpha_{\mathbf{i}}}^t)$ is the optimal allocation strategy at time t for the selling horizon.

2.3.3 The Dynamic Allocation Problem in Equilibrium

In this section, we focus on the problem, faced by each firm at the beginning of period $t \in \{2, \ldots, T\}$ when the market has already reached equilibrium. In other words, learning of the optimal allocation policies has been achieved. Hence, we can isolate learning of the price-demand relationship through the allocation sensitivities. The problem faced by each firm at each time period t of the selling horizon can be decomposed as follows: in step 1, each firm determines the Nash equilibrium quantities, assuming the allocation sensitivities are known; in step 2, it estimates the allocation sensitivities, and finally in step 3, it sets its optimal policy for future periods, and guesses its competitors's optimal strategy.

Step 1: Computation of the Nash Equilibrium Quantities

We assume that the allocation sensitivities α are known. Each firm i seeks to find the Nash equilibrium quantities that emerge in the market, i.e the quantities that simultaneously solve the N best-response problems:
$\mathcal{Q}_i(\alpha)$ solve $\mathcal{BR}_i(\mathcal{Q}_{-i}(\alpha), \alpha_{\mathbf{i}})$, for all $i = 1, \ldots, N$. We denote by $\mathcal{NE}(\alpha)$ the set of Nash equilibrium quantities.
<u>Remarks</u>:

1. The Nash equilibrium quantities are such that the optimal quantities for time t depend in general on the allocation sensitivities of all time periods. This is due to the capacity constraint which links the quantities in different periods with one another. Moreover, firm i's Nash quantities depend on all the firms' allocation sensitivities since the Nash quantities simultaneously solve the best-response problem of all the firms.

2. Recall that under Bertrand competition, the equilibrium is a *generalized* Nash equilibrium, since each firm' strategy space depends on its competitors' strategy. This is not the case under Cournot competition, since the capacity constraint involves the decision variables themselves, and not the parameters. We will elaborate more on the differences between generalized equilibrium and equilibrium in Chapter 3 of the thesis.

Step 2: Estimation of the Allocation Sensitivities

In this section, we address the estimation problem faced by each firm at each time t, when the market is in equilibrium. The firm aim at updating its own, and its competitors's estimate of the allocation sensitivities, based on the market data up to time $t-1$. Since the Nash equilibrium quantities depend on the allocation sensitivities of all the firms, then firm i's estimation problem is coupled with its competitors' estimation problems, and each firm needs to jointly estimate all the firms' sensitivities. For computational tractability, instead of solving N simultaneous estimation problems, we solve a single estimation problem, minimizing the sum of the estimation errors for all periods and all firms.

$$\min_{\alpha \in \mathcal{A}_1 \times \ldots \times \mathcal{A}_N} \sum_{i=1}^{N} \left\{ \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} (\widehat{q}_i^{\,h}(s) - \mathcal{Q}_i(s, \alpha_{\mathbf{i}}, \alpha_{-\mathbf{i}}))^2 \} + \sum_{s=1}^{T} (\widehat{q}_i^{\,0}(s) - \mathcal{Q}_i(s, \alpha_{\mathbf{i}}, \alpha_{-\mathbf{i}}))^2 \right\} \tag{2.13}$$

The objective function is the sum of the objective functions of each firm's estimation problem in the transient regime. The solution to the estimation problem is the set of optimal time-t estimators in the least squares sense of the allocation sensitivities $\widehat{\alpha}^t$, for all firms and all time periods.

Step 3: Optimal Allocation Policy

In Step 1, we computed the Nash equilibrium allocations for given parameters and in step 2, we computed the optimal parameters, for allocations equal to the Nash equilibrium allocations. Hence, both steps together yield the optimal vector allocation policy at time t, for all firms.

2.3.4 Reformulations of Steps 1 and 2

In Chapter 4, we show that Steps 1 and 2 under Cournot competition can be reformulated as one problem. Indeed, we show that the Cournot best-response problem with learning can be reformulated as a bilevel problem, or equivalently as a mixed integer program. The Cournot equilibrium problem can be reformulated as a Mathematical Program with Equilibrium Constraints (MPEC), which is also equivalent to a mixed integer program.
To achieve this, in Chapter 3, we will first focus on Step 1. We will then use the insights gained on Step 1 in Chapter 4 to consider Steps 1 and 2 jointly.

Chapter 3

Dynamic Policies When the Price-Demand Relationship Parameters Are Known

In this chapter, we assume that parameters of the price-demand relationship are known. That is, we assume that the price (resp. allocation) sensitivities of the parametric demand (resp. price) function are known under Bertrand (resp. Cournot) competition. Under each model of competition, we analyze the best-response and the equilibrium problems. We establish properties of the demand (resp. price) function under which there is existence and uniqueness of the best-response strategy and of the Nash equilibrium.

3.1 Study of the Best-Response Problem

In this section, we analyze the best-response problem when the parameters of the price-demand relationship are known. We focus on firm i, and fixed competitors' strategy. Since firms are assumed to be rational, then for known price-demand relationship, and given its competitors' strategy, firm i's optimal strategy is that which maximizes its total revenue over the selling horizon, while satisfying the feasibility and capacity constraints.
Under Bertrand (resp. Cournot) competition, we formulate the demand (resp. price) model. We then give alternative formulations of the best-response optimization probems as system of equalities and inequalities, using the Karush-Kuhn-Tucker conditions, and as variational inequality. We finally perform sensitivity analysis on the best-response solution. Indeed, sensitivity analysis is a key risk analysis to be carried out by companies, particularly in non robust settings. In disequilibrium, this analysis is twofold: firms seek to know how their best-response strategy changes for changes in the price-demand parameters. Furthermore, firms seek to know how their best-response strategy changes when their belief regarding their competitors' behavior changes. In equilibrium, firms want to keep track of the equilibrium strategy's sensitivity to changes in the price-demand parameters.

In equilibrium, we formulate the Nash equilibrium under Bertrand and Cournot competition as a system of equalities and inequalities, as a quasi variational inequality in the Bertrand case, and as a variational inequality in the Cournot case. Finally, we perform sensitivity analysis to quantify the effects of changes in parameters on the equilibrium.

3.1.1 The Bertrand Best-Response Problem

In Chapter 2, we defined as $q_i(\mathbf{p_i(t)}, \mathbf{p_{-i}(t)}, \beta_\mathbf{i}(\mathbf{t}))$ the demand function, for parameters $\beta_\mathbf{i}(\mathbf{t})$, as well as the best-response optimization problem:

$$\begin{aligned}\max_{\mathbf{p_i}} \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) &= \sum_{t=1}^{T} p_i(t) q_i(p_i(t), \overline{p_{-i}(t)}, \beta_i(t)) \\ \text{s.t} \quad & \sum_{t=1}^{T} q_i(p_i(t), \overline{p_{-i}(t)}, \beta_i(t)) \leq C_i \\ & 0 \leq p_i(t) \leq p_i^{max}\end{aligned}$$

For fixed parameter value $\beta_\mathbf{i}$ and fixed competitors' strategy $\overline{\mathbf{p_{-i}}}$, we denote by $\mathfrak{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ the feasible set of the best-response problem, in other words:

$$\mathfrak{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) = \{\mathbf{p_i} \ : \ \sum_{t=1}^{T} q_i(p_i(t), \overline{p_{-i}(t)}, \beta_i(t)) \leq C_i, \quad 0 \leq p_i(t) \leq p_i^{max}\}$$

The feasible set of parameters was denoted $\mathcal{B}_i$. We now state the Assumptions on the demand function that we consider in this Chapter. We then give examples of widely used demand functions in the revenue management and pricing literature for which these assumptions hold.

Assumptions on the Demand Function

Assumption 3.1. *We assume continuity and differentiability of the demand function q_i w.r.t $p_i(t)$.*

Assumption 3.2. *The market is for ordinary goods. This means that the demand function is increasing in the firm's own price, or equivalently: $\frac{\partial q_i}{\partial p_i(t)} < 0$.*

Assumption 3.3. *We focus on a market for substitutable products. This means that firm i's demand function is nondecreasing in firm j's price, for all $j \neq i$: in other words: $\frac{\partial q_i}{\partial p_j(t)} \geq 0$, $\forall j \neq i$.*

Assumption 3.4. *We assume strict concavity of the revenue rate $p_i(t)q_i(p_i(t), \overline{p_{-i}(t)})$ in $p_i(t)$. A weaker assumption is to assume strict pseudo-concavity of the revenue function $\pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}})$ in $\mathbf{p_i}$.*

Assumption 3.5. *We assume convexity of the demand function q_i in $p_i(t)$. A weaker assumption is to assume quasi-convexity of the intertemporal sum of the demand functions $\sum_{t=1}^{T} q_i(p_i(t), p_{-i}(t), \beta_i)$ in $\mathbf{p_i}$.*

Assumption 3.6. *We assume the existence of a maximum price p_i^{max} that each firm can set, and such that* $\inf_{p_{-i}} q_i(p_i^{max}, p_{-i}(t), \beta_i(t)) \geq 0$.

Assumption 3.7. *We assume that* $\inf_{p_{-i(t)}} q_i(0, p_{-i}(t), \beta_i(t)) > C_i$.

Assumption 3.1 guarantees that the best-response problem is a smooth optimization problem. Assumption 3.2 means that the demand for the good decreases when the price of the good increases, ceteris paribus. This is the case for most goods. A notable exception is luxury goods for which the demand may increase as the price increases. Assumption 3.3 implies that if firm j decreases its price, then some customers will likely substitute product j for product i, hence, firm i's demand may decrease. Strict pseudo-concavity in Assumption 3.4 means that the revenue function is differentiable and that the following holds: for all $\mathbf{p_i} \neq \tilde{\mathbf{p}}_\mathbf{i}$:

$$\nabla_i \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}})'(\tilde{\mathbf{p}}_\mathbf{i}, \mathbf{p_i}) < 0 \quad \Rightarrow \quad \pi_i(\tilde{\mathbf{p}}_\mathbf{i}, \overline{\mathbf{p_{-i}}}) < \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}})$$

Quasi-convexity in Assumption 3.5 means that the level sets

$$\{\mathbf{p_i} : \sum_{t=1}^{T} q_i(p_i(t), p_{-i}(t), \beta_i(t)) \leq M\}$$

are convex. Assumption 3.6 enables us essentially to reduce the set of feasible prices to a compact set. This is not restrictive in practice, since prices are always finite. The maximum allowable price is such that it generates a nonnegative demand, regardless of competitors' prices. Assumption 3.7 implies that if firm i prices at zero, then the demand generated, regardless of its competitors' pricing strategy, exceeds its capacity. It basically prevents a firm from posting a price of zero while still participating in the market.

Examples of Suitable Demand Functions

The following demand functions, which are widely used in the revenue management literature satisfy the above assumptions:

- Linear demand function:

$$q_i(p_i(t), p_{-i}(t), \beta_i(t)) = \beta_{i0}(t) - \beta_{ii}(t)p_i(t) + \sum_{j \neq i} \beta_{ij}(t)p_j(t)$$

  For $\beta_{ii}(t) > 0$ and $\beta_{ij}(t) \geq 0 \ \forall\ j \neq i$, then it is easy to verify that the linear demand function verifies assumptions 3.1 to 3.5.
  Assumption 3.7 prescribes that $\beta_{i0}(t) + \sum_{j \neq i} \beta_{ij}(t)p_j(t) > C_i$ for all i, t. It is sufficient that $\beta_{i0}(t) \geq C_i$ for all i, t.
  Hence, the set $\mathcal{B}_i$ of feasible parameters is:

$$\begin{aligned} \mathcal{B}_i = \{\beta : \qquad & \forall\ t, i, \beta_{ii}(t) \geq \varepsilon \\ & \forall\ i, t, \forall\ j \neq i, \beta_{ij}(t) \geq 0 \\ & \forall\ i, t, \beta_{i0}(t) \geq C_i\} \end{aligned}$$

where $\varepsilon > 0$.

- Loglinear demand function:

$$q_i(p_i(t), p_{-i}(t), \beta_i(t)) = \exp\left(\beta_{i0}(t) - \beta_{ii}(t)p_i(t) + \sum_{j \neq i} \beta_{ij}(t)p_j(t)\right)$$

The loglinear demand function verifies assumptions 3.1 to 3.5 for $0 < \beta_{ii}(t) < \frac{2}{p_i^{max}}$ and $\beta_{ij}(t) \geq 0$ for all $j \neq i$.
Assumption 3.7 states that $\beta_{i0}(t) + \sum_{j \neq i} \beta_{ij}(t)p_j(t) > \ln(C_i)$, hence, a sufficient condition for 3.7 to be satisfied for all $p_{-i}(t)$ is $\beta_{i0}(t) > \ln(C_i)$ for all i, t.
The corresponding feasible set of parameters is therefore:

$$\begin{aligned} \mathcal{B}_i = \{\beta : \quad & \forall\ t, i, \varepsilon \leq \beta_{ii}(t) \leq \tfrac{2}{p_i^{max}} - \varepsilon \\ & \forall\ i, t, \forall\ j \neq i, \beta_{ij}(t) \geq 0 \\ & \forall\ i, t, \beta_{i0}(t) \geq \ln(C_i)\} \end{aligned}$$

where $\varepsilon > 0$.

- Constant elasticity demand function:

$$q_i(p_i(t), p_{-i}(t), \beta_i(t)) = \beta_{i0}(t) \frac{\prod_{j \neq i} (p_j(t))^{\beta_{ij}(t)}}{(p_i(t))^{\beta_{ii}(t)}}$$

Assumptions 3.1 to 3.5 are satisfied when $0 < \beta_{ij}(t) < 1$ and $\beta_{ij}(t) \geq 0$ for all $j \neq i$.
Hence, the feasible set of parameters:

$$\begin{aligned} \mathcal{B}_i = \{\beta : \quad & \forall\ t, i, \varepsilon \leq \beta_{ii}(t) \leq 1 - \varepsilon \\ & \forall\ i, t, \forall\ j \neq i, \beta_{ij}(t) \geq 0\} \end{aligned}$$

where $\varepsilon > 0$.
Since the constant elasticity demand function is not defined for prices equal to zero, we have to choose a positive lower bound p_i^{min} for the prices. Assumption 3.7 then becomes:

$$\inf_{p_{-i}(t)} q_i(p_i^{min}, p_{-i}(t), \beta_i(t)) > C_i$$

Changing the lower bound from 0 to a positive amount does not change the rest of the analysis.

Existence and Uniqueness of the Best-Response Solution

Given the above assumptions we establish the following properties concerning the best-response problem:

Proposition 3.1. *The best-response problem under Bertrand competition has a unique solution, denoted* $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$*.*

Proof.

- Existence of the best-response solution:
  The revenue function is continuous in p_i and the feasible set is compact: indeed, it is a closed subset of the cube: $[0, p_i^{max}]^N$. Hence, by Weierstrass' s theorem, there exists a solution to the best-response problem.

- Uniqueness of the best-response solution:
  Assume that the best-response problem has two solutions $\hat{\mathbf{p}}_{\mathbf{i}}$ and $\check{\mathbf{p}}_{\mathbf{i}}$ such that $\hat{p}_i(t) \neq \check{p}_i(t)$ for some t.
  By convexity and continuity of q_i, the feasible set $\mathfrak{P}_i(\overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$ is closed convex, and π_i is continuously differentiable in $\mathbf{p_i}$. Hence, $\hat{\mathbf{p}}_{\mathbf{i}}$ and $\check{\mathbf{p}}_{\mathbf{i}}$ verify: for all $\mathbf{p_i} \in \mathfrak{P}_i(\overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$:

  $$-\nabla_i \pi_i(\hat{\mathbf{p}}_{\mathbf{i}}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})\prime(\mathbf{p_i} - \hat{\mathbf{p}}_{\mathbf{i}}) \geq 0$$

  We apply the inequality to $\mathbf{p_i} = \check{\mathbf{p}}_{\mathbf{i}}$ and get:

  $$-\nabla_i \pi_i(\hat{\mathbf{p}}_{\mathbf{i}}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})\prime(\check{\mathbf{p}}_{\mathbf{i}} - \hat{\mathbf{p}}_{\mathbf{i}}) \geq 0$$

  Hence, by strict pseudo-concavity of π_i w.r.t $\mathbf{p_i}$, the above inequality implies that

  $$\pi_i(\hat{\mathbf{p}}_{\mathbf{i}}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}}) < \pi_i(\check{\mathbf{p}}_{\mathbf{i}}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$$

  This contradicts the optimality of $\hat{\mathbf{p}}_{\mathbf{i}}$.
  Therefore, the solution to the best-response problem is unique.

□

Equivalence of the Best-Response Problem to a Variational Inequality

Here, we prove that the best-response optimization problem is equivalent to a *variational inequality*. A finite-dimensional variational inequality formulation is particularly convenient to study equilibrium problems, as shall be seen in the section corresponding to the analysis of the Nash equilibrium problem. It allows for a unified treatment of optimization problems and equilibrium problems. We now establish the equivalence of the best-response problem to a variational inequality.

Proposition 3.2. *The best response problem $\mathcal{BR}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ is equivalent to a variational inequality in the following sense:*
$\mathbf{p_i} = \mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ *solves the best-response problem iff it solves the following variational inequality denoted* $VI(-\nabla_i \pi_i(., \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}), \mathfrak{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}))$*:*

$$-\nabla_i \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})\prime(\tilde{\mathbf{p}}_\mathbf{i} - \mathbf{p_i}) \geq 0$$

for all $\tilde{\mathbf{p}}_\mathbf{i} \in \mathfrak{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$.

Proof.

- Let $\phi(t) = \pi_i(t.\tilde{\mathbf{p}}_\mathbf{i} + (1-t)\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}), \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$. By convexity of the feasible set, then for all $\tilde{\mathbf{p}}_\mathbf{i} \in \mathfrak{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ and for all $t \in [0,1]$, the point $t.\tilde{\mathbf{p}}_\mathbf{i} + (1-t)\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}), \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}$ is in the feasible set. Since $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ is the best-response solution, then $\phi(t)$ achieves its maximum at $t = 0$.
  Since π_i is continuously differentiable in $\mathbf{p_i}$, then ϕ is continuously differentiable in t. Thus $\frac{d\phi}{dt}(0) \geq 0$. This is equivalent to the variational inequality.
- By pseudo-concavity of $-\pi_i$, the variational inequality implies:

  $$\pi_i(\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}), \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \geq \pi_i(\tilde{\mathbf{p}}_\mathbf{i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$$

  for all feasible $\tilde{\mathbf{p}}_\mathbf{i}$. Hence, $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ is the best-response solution.

□

The variational inequality formulation states that the vector $-\nabla_i \pi_i(\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}), \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ makes an "acute" angle with any feasible vector emanating from $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$.

First Order Optimality Conditions

In what follows, we introduce a reformulation of the best-response problem based on the Karush-Kuhn-Tucker optimality conditions (hereafter called KKT conditions). We prove that under the assumptions made in Section 1, the KKT conditions are necessary and sufficient for optimality.
In what follows, we denote by $g_i(\mathbf{p_i}, \mathbf{p_{-i}}, \beta_\mathbf{i}) = \sum_{t=1}^{T} q_i(p_i(t), p_{-i}(t), \beta_i(t)) - C_i$ the capacity constraint.

Proposition 3.3. *Assume that if capacity is tight at $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$, then at least one of the best-response prices is interior to the compact $[0, p_i^{max}]$.*
Then $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ is the best-response solution iff there is a scalar $\lambda_i \geq 0$, and two vectors in $\mathbb{R}^T$ $\underline{\mu}_\mathbf{i}, \overline{\mu}_\mathbf{i} \geq \mathbf{0}$ such that the following system holds at $\mathbf{p_i} = \mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$:

$$-\nabla_i \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) + \lambda_i \nabla_i g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) + \sum_{t=1}^{T} (\overline{\mu}_i(t) - \underline{\mu}_i(t))\mathbf{e}(t) = 0 \tag{3.1}$$

$$\lambda_i g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) = 0 \tag{3.2}$$

$$\overline{\mu}_\mathbf{i}{}'(\mathbf{p_i} - \mathbf{p_i}^{\max}) = 0 \tag{3.3}$$

$$\underline{\mu}_\mathbf{i}{}'\mathbf{p_i} = 0 \tag{3.4}$$

$$g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \leq 0 \tag{3.5}$$

$$\mathbf{0} \leq \mathbf{p_i} \leq \mathbf{p_i^{max}} \tag{3.6}$$

$\lambda_i, \underline{\mu}_{\mathbf{i}}, \overline{\mu}_{\mathbf{i}}$ are called KKT multipliers corresponding to the constraints $g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}}) \leq 0$, $\mathbf{p_i} \geq \mathbf{0}$ and $\mathbf{p_i} \leq \mathbf{p_i^{max}}$ respectively.
In order to prove this, let us first introduce some notation. Let us fix $\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}}$. We denote by $\underline{\mathcal{I}}_i(\mathbf{p_i}) = \{t \ : \ p_i^t = 0\}$ the set of periods for which the price is zero, $\bar{\mathcal{I}}_i(\mathbf{p_i}) = \{t \ : \ p_i^t = p_i^{max}\}$ the set of periods for which the price is p_i^{max}, so that $\underline{\mathcal{I}}_i(\mathbf{p_i}) \bigcup \bar{\mathcal{I}}_i(\mathbf{p_i})$ is the set of periods of active price constraints.

Proof. Let $\mathbf{p_i} = \mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$.

- The KKT conditions are necessary to optimality if the set of gradients of the tight constraints is linearly independent.
  Let $\bar{t} \in \bar{\mathcal{I}}_i(\mathbf{p_i})$ and Let $\underline{t} \in \underline{\mathcal{I}}_i(\mathbf{p_i})$. The gradient of $p_i(\bar{t})$ is $\mathbf{e}(\bar{\mathbf{t}})$, the vector with all components zero except for component $\bar{t}$ equal to 1, and that of $p_i(\underline{t})$ is $\mathbf{e}(\underline{\mathbf{t}})$. Notice that $\underline{\mathcal{I}}_i(\mathbf{p_i}) \bigcap \bar{\mathcal{I}}_i(\mathbf{p_i}) = \emptyset$. Hence, if the capacity is not tight, the gradients of tight constraints are indeed linearly independent.
  Assume that capacity is tight. The gradient $\nabla_i g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$ is the vector of components $\frac{\partial g_i}{\partial p_i(t)}(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}}) = \frac{\partial q_i}{\partial p_i(t)}(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}})$.
  Assume that there exist a scalar γ_i and two vectors $\bar{\gamma}_{\mathbf{i}}$ and $\underline{\gamma_i}$ not all equal to zero such that:

$$\gamma_i \nabla_i g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}}) - \sum_{t \in \underline{\mathcal{I}}_i(\mathbf{p_i})} \underline{\gamma}_i(t) \mathbf{e}(\mathbf{t}) + \sum_{t \in \bar{\mathcal{I}}_i(\mathbf{p_i})} \bar{\gamma}_i(t) \mathbf{e}(\mathbf{t}) = \mathbf{0}$$

  Notice that since there exists t_0 such that $0 < p_i(t_0) < p_i^{max}$, then $\underline{\mathcal{I}}_i(\mathbf{p_i}) \bigcup \bar{\mathcal{I}}_i(\mathbf{p_i}) \subsetneq \{1, \ldots, T\}$.
  Component t_0 of the above linear combination is: $\gamma_i \frac{\partial g_i}{\partial p_i(t_0)} = 0$. Since $\frac{\partial g_i}{\partial p_i(t_0)} < 0$ by Assumption 3.2, then this yields $\gamma_i = 0$. Since $\underline{\mathcal{I}}_i(\mathbf{p_i}) \bigcap \bar{\mathcal{I}}_i(\mathbf{p_i}) = \emptyset$, then we must have: $\underline{\gamma}_i(t) = \bar{\gamma}_i(t') = 0$ for all $t \in \underline{\mathcal{I}}_i(\mathbf{p_i})$ and $t' \in \bar{\mathcal{I}}_i(\mathbf{p_i})$.
  This contradicts our initial assumption. Therefore, the gradient of active constraints are linearly independent.

- The KKT conditions are sufficient for optimality if the objective function π_i is pseudo concave and g_i is quasi-convex in $\mathbf{p_i}$, which is ensured by Assumptions 3.4 and 3.5.

$\square$

The complementary slackness conditions (3.2), (3.3) and (3.4) are special cases of disjunctive constraints. This implies that the system of KKT conditions has an inherent combinatorial nature. As a result, the best-response problem is NP-hard. In mathematical programming theory, a problem is said to be NP (nondeterministic polynomial) if it is solvable in polynomial time by a nondeterministic Turing machine. A problem is said to be NP-hard if an algorithm for solving it can be translated into an algorithm to solve any NP problem. In other words, it is "at least as hard" as an NP problem, although it might, in fact, be harder.
Furthermore, we prove that the KKT multipliers are unique.

Proposition 3.4. *Let $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ be the solution to the best-response problem $\mathcal{BR}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$. Assume that it is also solution to the KKT conditions. Then there is a unique scalar $\lambda_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$, and unique vectors $\bar{\mu}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ and $\underline{\mu}(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ solving the KKT conditions at $\mathbf{p_i} = \mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$*

Proof. The proof is in two parts:

1. Assume that capacity is not tight at $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$:
   Then $\lambda_i = 0$ and thus the KKT conditions become:

$$-\nabla_i \pi_i(p_i, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) + \sum_{t=1}^{T} (\bar{\mu}_i(t) - \underline{\mu}_i(t))\mathbf{e}(t) = 0$$

   For $t \in \underline{\mathcal{I}}_i(\mathbf{p_i})$, we have $\bar{\mu}_i(t) = 0$, and hence, $\underline{\mu}_i(t) = -\frac{\partial \pi_i}{\partial p_i(t)}(p_i, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$.
   For $t \in \bar{\mathcal{I}}_i(\mathbf{p_i})$, then $\underline{\mu}_i(t) = 0$, thus $\bar{\mu}_i(t) = -\frac{\partial \pi_i}{\partial p_i(t)}(p_i, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$.
   For $t \notin \underline{\mathcal{I}}_i(\mathbf{p_i}) \bigcup \bar{\mathcal{I}}_i(\mathbf{p_i})$, we have $\underline{\mu}_i(t) = \bar{\mu}_i(t) = 0$.
   Hence, the multipliers are uniquely defined.

2. Assume capacity is tight at $\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$. Then we have the following:

   - For $t \notin \underline{\mathcal{I}}_i(\mathbf{p_i}) \bigcup \bar{\mathcal{I}}_i(\mathbf{p_i})$, we have $\underline{\mu}_i(t) = \bar{\mu}_i(t) = 0$, hence, the KKT conditions are:

$$-\frac{\partial \pi_i}{\partial p_i(t)}(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) + \lambda_i \frac{\partial q_i}{\partial p_i(t)}(p_i(t), \overline{p_{-i}(t)}, \beta_i) = 0$$
$$\sum_{t=1}^{T} q_i(p_i, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) = C_i$$

     This means that for all $t \in \underline{\mathcal{I}}_i(\mathbf{p_i}) \bigcup \bar{\mathcal{I}}_i(\mathbf{p_i})$, the ratio of the derivative of the profit function w.r.t $p_i(t)$ and the derivative of the demand function w.r.t $p_i(t)$ is constant equal to λ_i.

   - We can now replace λ_i by its value, and solve for $\underline{\mu}_i(t)$ for all $t \in \underline{\mathcal{I}}_i(\mathbf{p_i})$, and $\bar{\mu}_i(t)$ for all $t \in \bar{\mathcal{I}}_i(\mathbf{p_i})$:

$$\underline{\mu}_i(t) = \lambda_i \frac{\partial q_i}{\partial p_i(t)}(p_i, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) - \frac{\partial \pi_i}{\partial p_i(t)}(p_i, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \quad \forall\, t \in \underline{\mathcal{I}}_i(\mathbf{p_i})$$
$$\bar{\mu}_i(t) = \lambda_i \frac{\partial q_i}{\partial p_i(t)}(p_i, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) - \frac{\partial \pi_i}{\partial p_i(t)}(p_i, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \quad \forall\, t \in \bar{\mathcal{I}}_i(\mathbf{p_i})$$

   Hence, the multipliers are uniquely defined.

□

Sensitivity Analysis with respect to Parameters

We aim at quantifying how changes in parameters affects the best-response solution. We prove that under some additional assumptions, the best response as a function of the parameters $\beta_{\mathbf{i}}$ is a well defined function which is piecewise continuously differentiable (PC1), whose directional derivative in any direction is given as the unique solution to a quadratic program.
Toward that end, we denote by $\mathcal{M}_i = (\lambda_i, \bar{\mu}_{\mathbf{i}}, \underline{\mu}_{\mathbf{i}})$ the set of KKT multipliers, and we define the *Lagrangian function* $\mathcal{L}_i(p_i, \overline{\mathbf{p}_{-\mathbf{i}}}, \beta_{\mathbf{i}}, \mathcal{M}_i)$ of the best-response problem:

$$\mathcal{L}_i(p_i, \overline{\mathbf{p}_{-\mathbf{i}}}, \beta_{\mathbf{i}}, \mathcal{M}_i) = -\pi_i(\mathbf{p_i}, \overline{\mathbf{p}_{-\mathbf{i}}}, \beta_{\mathbf{i}}) + \lambda_i g_i(\mathbf{p_i}, \overline{\mathbf{p}_{-\mathbf{i}}}, \beta_{\mathbf{i}}) + (\bar{\mu}_{\mathbf{i}} - \underline{\mu}_{\mathbf{i}})\mathbf{p_i}$$

Then notice that the first equation of the KKT conditions is simply:

$$\nabla_i \mathcal{L}_i(p_i, \overline{\mathbf{p}_{-\mathbf{i}}}, \beta_{\mathbf{i}}, \mathcal{M}_i) = 0$$

To establish the differentiability property of the best-response solution, we introduce the following additional assumption:

Assumption 3.8. *Let $\mathcal{M}_i = (\lambda_i, \bar{\mu}_{\mathbf{i}}, \underline{\mu}_{\mathbf{i}})$ be the KKT multipliers at $\mathcal{P}_i(\overline{\mathbf{p}_{-\mathbf{i}}}, \beta_{\mathbf{i}})$. For all $\mathbf{x_i} \neq \mathbf{0}$ such that $\nabla_i g_i(\mathbf{p_i}, \overline{\mathbf{p}_{-\mathbf{i}}}, \beta_{\mathbf{i}})'\mathbf{x_i} = 0$ if $\lambda_i > 0$, we have:*

$$\mathbf{x_i}'\nabla^2_{ii}\mathcal{L}_i(p_i, \overline{\mathbf{p}_{-\mathbf{i}}}, \beta_{\mathbf{i}}, \mathcal{M}_i)\mathbf{x_i} > 0$$

Assumption **??** says that the Hessian matrix of the Lagrangian function is positive definite on the set of feasible directions. For an unconstrained optimization problem, this condition would simplify to: the hessian matrix of the objective function is positive definite, which is a condition sufficient for optimality for an unconstrained optimization problem.
We can now establish the first result of this section, which can be found in [28]. Let us fix competitors' strategy at $\overline{\mathbf{p}_{-\mathbf{i}}}$ and thus drop it from the notation. Let us fix $\bar{\beta}_i \in \mathcal{B}_i$.

Proposition 3.5. *Let $\bar{\mathcal{P}}_i = \mathcal{P}_i(\bar{\beta}_i)$ be the best response solution for parameters $\bar{\beta}_i$. Assume that KKT conditions are necessary and sufficient at $\mathcal{P}_i(\bar{\beta}_i)$, and let $\lambda_i, \bar{\mu}_{\mathbf{i}}, \underline{\mu}_{\mathbf{i}}$ be the corresponding KKT multipliers. Under Assumptions 3.1 to 3.8, there exist open neighborhoods $\mathcal{U}$ of $\bar{\beta}_i$ and $\mathcal{V}$ of $\mathcal{P}_i(\bar{\beta}_i)$ such that:*

1. *$\mathcal{P}_i(.)$ is continuous in $\mathcal{U}$, such that $\mathcal{P}_i(\beta_i)$ is best-response solution for all $\beta_i \in \mathcal{U}$;*
2. *The function $\mathcal{P}_i(.)$ is directionally differentiable, and for all directions $\mathbf{d_i}$, the directional derivative $\mathbf{x_i} = \mathcal{P}'_i(\bar{\beta}_i; \mathbf{d_i})$ is the unique solution to the following convex quadratic problem, denoted $\mathcal{QP}(\bar{\beta}_i; \mathbf{d_i})$:*

$$\min_{\mathbf{x_i} \in \mathcal{C}(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i})} \frac{1}{2}\mathbf{x_i}'\nabla^2_{ii}\mathcal{L}_i(\bar{\mathcal{P}}_i, \bar{\beta}_{\mathbf{i}}, \mathcal{M}_i(\bar{\beta}_{\mathbf{i}}))\mathbf{x_i} + \mathbf{d_i}'\nabla^2_{i\beta_i}\mathcal{L}_i(\bar{\mathcal{P}}_i, \bar{\beta}_{\mathbf{i}}, \mathcal{M}_i(\bar{\beta}_{\mathbf{i}}))\mathbf{x_i} \tag{3.7}$$

where $\mathcal{C}(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i})$ is the critical cone at $\mathcal{P}_i(\bar{\beta}_i)$ in direction $\mathbf{d_i}$.

Let us denote by $\bar{\mathcal{I}}_i^0(\beta_{\mathbf{i}}) = \{t : \bar{\mu}_i(t) = 0\}$, $\underline{\mathcal{I}}_i^0(\beta_{\mathbf{i}}) = \{t : \underline{\mu}_i(t) = 0\}$ the sets of periods with zero multipliers, $\bar{\mathcal{I}}_i^+(\beta_{\mathbf{i}}) = \{t : \bar{\mu}_i(t) > 0\}$, $\underline{\mathcal{I}}_i^+(\beta_{\mathbf{i}}) = \{t : \underline{\mu}_i(t) > 0\}$, the sets of periods of positive multipliers. The critical cone at $\mathcal{P}_i(\bar{\beta}_i)$ in direction $\mathbf{d}$ is defined as follows:

- If $g_i(\bar{\mathcal{P}}_i, \overline{\mathbf{p_{-i}}}, \bar{\beta}_i) < 0$, then $\mathcal{C}(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i}) = \mathcal{C}^0(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i})$, where $\mathcal{C}^0(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i})$ is defined as:

$$\begin{aligned}\mathcal{C}^0(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i}) = \{\mathbf{x_i} \in \mathbb{R}^{(N+1)T} : \quad & \forall t \in \bar{\mathcal{I}}_i^0(\beta_{\mathbf{i}}),\ x_i(t) \leq 0 \\ & \forall t \in \underline{\mathcal{I}}_i^0(\beta_{\mathbf{i}}),\ x_i(t) \geq 0 \\ & \forall t \in \bar{\mathcal{I}}_i^+(\beta_{\mathbf{i}}) \bigcup \underline{\mathcal{I}}_i^+(\beta_{\mathbf{i}}),\ x_i(t) = 0\}\end{aligned}$$

- If $g_i(\bar{\mathcal{P}}_i, \overline{\mathbf{p_{-i}}}, \bar{\beta}_i) = 0$ and $\lambda_i > 0$, then:

$$\begin{aligned}&\mathcal{C}(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i}) \\ &= \mathcal{C}^0(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i}) \bigcup \{\mathbf{x_i} \in \mathbb{R}^{(N+1)T} : \nabla_i g_i(\bar{\mathcal{P}}_i, \overline{\mathbf{p_{-i}}}, \bar{\beta}_i)' \mathbf{x_i} + \nabla_{\beta_i} g_i(\bar{\mathcal{P}}_i, \overline{\mathbf{p_{-i}}}, \bar{\beta}_i)' \mathbf{d_i} = 0\}\end{aligned}$$

- If $g_i(\bar{\mathcal{P}}_i, \overline{\mathbf{p_{-i}}}, \bar{\beta}_i) = 0$ and $\lambda_i = 0$, then:

$$\begin{aligned}&\mathcal{C}(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i}) \\ &= \mathcal{C}^0(\bar{\beta}_{\mathbf{i}}; \mathbf{d_i}) \bigcup \{\mathbf{x_i} \in \mathbb{R}^{(N+1)T} : \nabla_i g_i(\bar{\mathcal{P}}_i, \overline{\mathbf{p_{-i}}}, \bar{\beta}_i)' \mathbf{x_i} + \nabla_{\beta_i} g_i(\bar{\mathcal{P}}_i, \overline{\mathbf{p_{-i}}}, \bar{\beta}_i)' \mathbf{d_i} \leq 0\}\end{aligned}$$

The critical cone has the following interpretation. Consider a constraint of the form $\mathbf{h_i}(\mathbf{p_i}) \leq 0$, and best-response solution $\bar{\mathcal{P}}_i$ at $\bar{\beta}_i$. Consider the following index sets: $\mathcal{I}_i^+(\bar{\beta}_i) = \{t : h_i^t(\bar{\mathcal{P}}_i) = 0 \text{ and } \mu_i^t > 0\}$ is the set of *strongly active* constraints, i.e active constraints for which the lagrangian multiplier is positive. $\mathcal{I}_i^0(\bar{\beta}_i) = \{t : h_i^t(\bar{\mathcal{P}}_i)) = 0 \text{ and } \mu_i^t = 0\}$ is the set of *weakly active* constraints, i.e active constraints for which the lagrangian multiplier is zero. Let $\mathcal{I}_i(\bar{\beta}_i) = \mathcal{I}_i^+(\bar{\beta}_i) \bigcup \mathcal{I}_i^0(\bar{\beta}_i)$ the set of indices of active constraints.

First of all, the set of directions $\{\mathbf{w} \in \mathbb{R}^T : \forall t \in \mathcal{I}_i(\bar{\beta}_i)\ \nabla h_i^t(\mathcal{P}_i(\bar{\beta}_i))' \mathbf{w} \leq 0\}$ contains the set of feasible directions at $\bar{\mathcal{P}}_i$, i.e the directions for which there is $\bar{\epsilon} > 0$ such that for all $0 < \epsilon < \bar{\epsilon}$, $\mathbf{h_i}(\bar{\mathcal{P}}_i + \epsilon \mathbf{w}) \leq 0$. In the Bertrand best-response problem, since the inequality constraints independent of β_i are linear, then the set $\{\mathbf{w} \in \mathbb{R}^T : \forall t \in \mathcal{I}_i(\bar{\beta}_i)\ \nabla h_i^t(\bar{\mathcal{P}}_i)' \mathbf{w} \leq 0\}$ is actually equal to the set of feasible directions at $\bar{\mathcal{P}}_i$.

Consider now a constraint of the form $\mathbf{h_i}(\mathbf{p_i}, \bar{\beta}_i) \leq 0$, best-response solution $\bar{\mathcal{P}}_i$ at $\bar{\beta}_i$, and a direction $\mathbf{d_i}$. Since the gradient of h_i^t is: $\nabla h_i^t(\mathbf{p_i}, \beta_i) = [\nabla_{p_i} h_i^t(\mathbf{p_i}, \beta_i), \nabla_{\beta_i} h_i^t(\mathbf{p_i}, \beta_i)]'$. The set of feasible directions now writes:

$$\{(\mathbf{w^1}, \mathbf{w^2}) \in \mathbb{R}^T \times \mathbb{R}^{(N+1)T} : \forall t \in \mathcal{I}_i(\bar{\beta}_i),\ \nabla h_i^t(\bar{\mathcal{P}}_i, \bar{\beta}_{\mathbf{i}})' [\mathbf{w^1}, \mathbf{w^2}]' \leq 0\}$$

If we fix direction $\mathbf{w^2} = \mathbf{d_i}$, we get the set of feasible directions $\{\mathbf{w} \in \mathbb{R}^T : \forall t \in \mathcal{I}_i(\bar{\beta}_i),\ \nabla h_i^t(\bar{\mathcal{P}}_i, \bar{\beta}_{\mathbf{i}})' [\mathbf{w}, \mathbf{d_i}]' \leq 0\}$.

Furthermore, the critical cone also differentiates between weakly and strongly active constraints. The critical cone also appears in KKT second order sufficient conditions for nonlinear problems: indeed, if the Hessian of the Lagrangian function is positive definite on the critical cone at a point $\bar{x}$, then $\bar{x}$ is an optimal solution to the problem. We now state the second result of this section:

Proposition 3.6. *Let $\bar{\beta}_i$ a vector of parameters, and $\bar{\mathcal{P}}_i$ the corresponding best-response solution. Assume Assumptions 3.1 to 3.8 hold. Then the best-response function $\mathcal{P}_i(.)$ whose existence was established in Proposition 3.5 is piecewise continuously differentiable, and the directional derivative $\mathcal{P}_i'(\bar{\beta}_\mathbf{i}; \mathbf{d_i})$ is a piecewise linear function which is unique solution to the convex quadratic program $\mathcal{QP}(\bar{\beta}_i; \mathbf{d_i})$.*

Proof. The proof is to be found in [28] under the Constant Rank Constraint Qualification: Any subset of the gradients of active constraints has the same rank for all points in a neighborhood of the optimal solution $\mathcal{P}_i(\bar{\beta}_\mathbf{i})$.
Since linear independence is a special case of the Constant Rank assumption and was proved above, then the proposition holds. □

Definition 3.1. *$\mathcal{P}_i(.)$ is said to be piecewise continuously differentiable function (PC1) near $\bar{\beta}_i$ if it is a continuous function, and there is a finite family of continuously differentiable functions $\{\phi_1(\beta_i), \dots, \phi_K(\beta_i)\}$ defined in a neighborhood of $\bar{\beta}_i$, such that for some k, $\mathcal{P}_i(\beta_i) = \phi_k(\beta_i)$ fir each β_i in that neighborhood.*

PC1 functions are not differentiable, but they have a directional derivative. They are also *Bouligand-differentiable*. That is, the directional derivative gives a first order approximation of the function:

$$\mathcal{P}_i(\bar{\beta}_i + \mathbf{d_i}) = \mathcal{P}_i(\bar{\beta}_i) + \mathcal{P}_i'(\bar{\beta}_i; \mathbf{d_i}) + o_{\|\mathbf{d_i}\| \to 0}(\|\mathbf{d_i}\|)$$

where $o(.)$ is such that $\lim_{\|\mathbf{x}\| \to 0} \frac{o(\|\mathbf{x}\|)}{\|\mathbf{x}\|} = 0$. We will the use B-differentiability property of the directional derivative in the design of an algorithm in Chapter 6.
Furthermore, since $\mathcal{P}_i'(\bar{\beta}_i; .)$ is a piecewise linear function, then if $\mathcal{P}_i(\beta_i) = \phi_k(\beta_i)$ for some k, we have:

$$\mathcal{P}_i'(\beta_i; \mathbf{d_i}) = \nabla \phi_k(\beta_i)' \mathbf{d_i}$$

We can therefore define the generalized Jacobian $\partial \mathcal{P}_i(.)$ of $\mathcal{P}_i(.)$:

$$\partial \mathcal{P}_i(\beta_i) = \{\nabla \phi_k(\beta_i) \ \forall k \text{ s.t } \mathcal{P}_i(\beta_i) = \phi_k(\beta_i)\}$$

Example:
In two dimensions, consider the following simple piecewise continuously differentiable function: $f(x, y) = |x - y|$. It is made of two pieces $f(x, y) = y - x = f_-(x, y)$, for all $x \leq y$, and $f(x, y) = x - y = f_+(x, y)$, for all $x \geq y$. It is not differentiable at $x = y$, but for $x < y$, it is differentiable, with gradient equal to $\nabla f_-(x, y) = (-1, 1)'$, and for $x > y$, it is differentiable, with gradient equal to $\nabla f_+(x, y) = (1, -1)'$.
On the plane $\{(x, y) \ : \ x = y\}$, the generalized jacobian is the set: $\partial f(x, y) = \{(-1, 1)'; (1, -1)'\}$.

Sensitivity Analysis w.r.t Competitors' Strategy

Another sensitivity feature of importance to firms is how sensitive their best response strategy is to changes in their competitors' behaviors. In particular, since the best-response strategies are computed using a firm's belief on its competitors, and that this belief might be inaccurate, firms should keep track of how their strategy would differ, should their competitors behave differently than they believed.
A similar theorem to that regarding sensitivity analysis w.r.t the parameters states that the best-response as a function of competitors' strategy is directionally differentiable and that the directional derivative is the unique solution to a convex quadratic program. In what follows, we consider the best-response for fixed parameters $\beta_{\mathbf{i}}$, and therefore drop the dependence on $\beta_{\mathbf{i}}$ from the notation.

Proposition 3.7. *Let $\mathcal{P}_i(\overline{\mathbf{p_{-i}}})$ be the best-response solution for competitors' strategy $\overline{\mathbf{p_{-i}}}$. Assume that KKT conditions are necessary and sufficient at $\mathcal{P}_i(\overline{\mathbf{p_{-i}}})$, and let $\lambda_i, \bar{\mu}_{\mathbf{i}}, \underline{\mu}_{\mathbf{i}}$ be the corresponding KKT multipliers. Under Assumptions 3.1 to 3.8, there exist open neighborhoods $\mathcal{U}$ of $\overline{\mathbf{p_{-i}}}$ and $\mathcal{V}$ of $\mathcal{P}_i(\overline{\mathbf{p_{-i}}})$ such that:*

1. *$\mathcal{P}_i(.)$ is continuous in $\mathcal{U}$, such that $\mathcal{P}_i(\mathbf{p_{-i}})$ is best-response solution for all $\mathbf{p_{-i}} \in \mathcal{U}$;*

2. *The function $\mathcal{P}_i(.)$ is directionally differentiable, and for all directions $\mathbf{d_{-i}} \in \mathbb{R}^{(N-1)T}$, the directional derivative $\mathbf{y_i} = \mathcal{P}_i'(\overline{\mathbf{p_{-i}}}; \mathbf{d_{-i}})$ is the unique solution to the following convex quadratic problem, denoted $\mathcal{QP}(\overline{\mathbf{p_{-i}}}; \mathbf{d_{-i}})$:*

$$\min_{\mathbf{y_i} \in \mathcal{C}(\overline{\mathbf{p_{-i}}}; \mathbf{d_{-i}})} \frac{1}{2} \mathbf{y_i}' \nabla^2_{ii} \mathcal{L}_i(\bar{\mathcal{P}}_i, \overline{\mathbf{p_{-i}}}, \mathcal{M}_i) \mathbf{y_i} + \mathbf{d_{-i}}' \nabla^2_{i,-i} \mathcal{L}_i(\bar{\mathcal{P}}_i, \overline{\mathbf{p_{-i}}}, \mathcal{M}_i) \mathbf{y_i} \tag{3.8}$$

   where $\mathcal{C}(\overline{\mathbf{p_{-i}}}; \mathbf{d_{-i}})$ is the critical cone at $\mathcal{P}_i(\overline{\mathbf{p_{-i}}})$ in direction $\mathbf{d_{-i}}$

3. *Furthermore, $\mathcal{P}_i(.)$ is piecewise continuously differentiable and $\mathcal{P}_i'(\overline{\mathbf{p_{-i}}}; .)$ is piecewise linear.*

3.1.2 The Cournot Best-Response Problem

In Chapter 2, we defined as $p_i(\mathbf{q_i}(\mathbf{t}), \mathbf{q_{-i}}(\mathbf{t}), \alpha_{\mathbf{i}}(\mathbf{t}))$ the inverse demand function, also called price function, for parameters $\alpha_{\mathbf{i}}(\mathbf{t})$, as well as the best-response optimization problem:

$$\begin{aligned} \max_{\mathbf{q_i}} \pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}) \quad &= \textstyle\sum_{t=1}^T q_i(t) p_i(q_i(t), \overline{q_{-i}(t)}, \alpha_i(t)) \\ \text{s.t} \qquad\qquad & \textstyle\sum_{t=1}^T q_i(t) \leq C_i \\ & 0 \leq q_i(t) \end{aligned}$$

For fixed parameter value $\alpha_{\mathbf{i}}$ and fixed competitors' strategy $\overline{\mathbf{q_{-i}}}$, we denote by $\mathfrak{Q}_i)$ the feasible set of the best-response problem, in other words:

$$\mathfrak{Q}_i = \{\mathbf{q_i} \;:\; \sum_{t=1}^T q_i(t) \leq C_i, \quad 0 \leq q_i(t) \leq q_i^{max}\}$$

Notice that unlike the Bertrand feasible set, the Cournot feasible set is independent of competitors' strategy and of the price-demand parameters.
The feasible set of parameters was denoted $\mathcal{A}_i$. We now state the assumptions on the price function that we consider in this Chapter. We then give examples of widely used inverse demand functions in the revenue management and pricing literature for which these assumptions hold.

Assumptions on the Price Function

Assumption 3.9. *First, we assume continuity and differentiability of the price function p_i w.r.t $q_i(t)$.*

Assumption 3.10. *The market is for ordinary goods, which means that the price function is increasing in the firm's own quantity, or equivalently: $\frac{\partial p_i}{\partial q_i(t)} < 0$.*

Assumption 3.11. *We focus on a market for substitutable products. In other words: $\frac{\partial p_i}{\partial q_j(t)} \leq 0$, $\forall j \neq i$.*

Assumption 3.12. *We assume strict concavity of the revenue rate $q_i(t)p_i(q_i(t), \overline{q_{-i}(t)})$ in $q_i(t)$. A weaker assumption is to assume strict pseudo-concavity of the revenue function $\pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}})$ in $\mathbf{q_i}$.*

Assumption 3.9 guarantees that the best-response problem is a smooth optimization problem. Assumption 3.10 implies that the demand for the good decreases when the price of the good increases, ceteris paribus. Assumption 3.11 means that firm i's price function is non increasing in firm j's quantity, for all $j \neq i$.Strict pseudo concavity in Assumption 3.12 has the same meaning as in Assumption 3.4.
Remarks:

- Recall that in the case of the Bertrand best-response problem, we assumed convexity of the demand function. This guarantees that the feasible set of the Bertrand best-response problem be a convex set. In the Cournot best-response problem, the feasible set is polyhedral, hence, convexity of the price function is not necessary. Furthermore, we assumed existence of a maximum price in order to guarantee compactness of the feasible set. In the Cournot best-response problem, $q_i(t) \geq 0$ and $\sum_{t=1}^{T} q_i(t) \leq C_i$ imply that each component $q_i(t)$ is at most C_i. Thus compactness holds without requiring existence of a maximum quantity. This difference is essentially technical and does not have an impact in real applications. Finally, we made the technical assumption that a firm cannot participate in the market while posting a zero price. This technical assumption finds its main applicability in the second bullet point below.

- We can obtain a formulation of the Bertrand best-response problem with a polyhedral feasible set, provided the following assumption holds:

  Assumption 3.13. *The demand function $q_i(., \overline{q_{-i}}(t), \beta_i(t))$ is invertible w.r.t $\mathbf{p_i}$, and we denote by $\breve{p}_i(., \overline{q_{-i}}(t), \beta_i(t))$ its inverse. In other words:*

  $$q_i(t) = q_i(p_i(t), \overline{q_{-i}}(t), \beta_i(t)) \Leftrightarrow p_i(t) = \breve{p}_i(q_i(t), \overline{q_{-i}}(t), \beta_i(t))$$

Let $\breve{\pi}_i(\mathbf{q_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) = \sum_{t=1}^{T} q_i(t)\breve{p}_i(q_i(t), \overline{q_{-i}}(t), \beta_i(t))$.

Proposition 3.8. *The Bertrand best-response problem can be reformulated as follows:*

$$\max_{\mathbf{q_i}} \quad \breve{\pi}_i(\mathbf{q_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \tag{3.9}$$

$$s.t \quad \sum_{t=1}^{T} q_i(t) \leq C_i \tag{3.10}$$

$$0 \leq q_i(t) \tag{3.11}$$

Proof. Let $\mathbf{p_i}$ be a feasible price vector for the Bertrand best response problem. Let us denote by $\breve{q}_i(t) = q_i(p_i(t), \overline{p_{-i}}(t), \beta_i(t))$. Constraint (3.9) follows by substitution of $q_i(p_i(t), \overline{p_{-i}}(t), \beta_i(t))$ by $q_i(t)$ in Constraint (2.2). Let us consider Constraint (2.3), i.e $0 \leq p_i(t) \leq p_i^{max}$ of the Bertrand best-response problem. Applying function q_i which is decreasing in $p_i(t)$ yields:

$$q_i(p_i^{max}, \overline{p_{-i}}(t), \beta_i(t)) \leq q_i(p_i(t), \overline{p_{-i}}(t), \beta_i(t)) \leq q_i(0, \overline{p_{-i}}(t), \beta_i(t))$$

By Assumption 3.6, the left hand side term is nonnegative, and by Assumption 3.7, the right hand side term is strictly greater than C_i. But since Constraint (3.9) implies that $q_i(t) \leq C_i$, then under our assumptions, Constraint (2.3) is equivalent to: $0 \leq q_i(t)$. □

Examples of Suitable Inverse Demand Functions

Among the demand functions which are widely used in the Revenue Management and Pricing literature, which ones have inverse demand functions that satisfy the above assumptions?

- Inverse linear demand function:

$$p_i(q_i(t), q_{-i}(t), \alpha_i(t)) = \alpha_{i0}(t) - \sum_{j=1}^{N} \alpha_{ij}(t) q_j(t)$$

Assumptions 3.9 to 3.12 are satisfied for $\alpha_{ii}(t) > 0$, $\alpha_{ij}(t) \geq 0$, $\forall\ j \neq i$. Hence, the feasible set $\mathcal{A}_i$ is:

$$\mathcal{A}_i = \{\alpha_\mathbf{i} : \alpha_{ii}(t) \geq \varepsilon\ \forall\ i, t \quad \alpha_{ij}(t) \geq 0\ \forall\ i, t\ \forall\ j \neq i\}$$

where $\varepsilon > 0$.

- Inverse loglinear demand function:

$$p_i(q_i(t), q_{-i}(t), \alpha_i(t)) = \alpha_{i0}(t) - \sum_{j=1}^{N} \alpha_{ij}(t) \ln(q_j(t))$$

Assumptions 3.9 to 3.12 are satisfied for $\alpha_{ii}^t > 0$ and $\alpha_{ij}^t \geq 0\ \forall\ j \neq i$.
Hence, the feasible set $\mathcal{A}_i$ is:

$$\mathcal{A}_i = \{\alpha_{\mathbf{i}} : \alpha_{ii}(t) \geq \varepsilon\ \forall\ i, t \qquad \alpha_{ij}(t) \geq 0\ \forall\ i, t\ \forall\ j \neq i\}$$

where $\varepsilon > 0$.

- Inverse constant elasticity demand function:

$$p_i(q_i(t), q_{-i}(t), \alpha_i(t)) = \exp\left(\alpha_{i0}(t) - \sum_{j=1}^{N} \alpha_{ij}(t) \ln(q_j(t))\right)$$

  Assumptions 3.9 to 3.12 are satisfied for $0 < \alpha_{ii}^t < 1$ and $\alpha_{ij}^t \geq 0$ for $i \neq j$.
  Hence, the feasible set $\mathcal{A}_i$ is:

$$\mathcal{A}_i = \{\alpha_{\mathbf{i}} : \varepsilon \leq \alpha_{ii}(t) \leq 1 - \varepsilon\ \forall\ i, t \qquad \alpha_{ij}(t) \geq 0\ \forall\ i, t\ \forall\ j \neq i\}$$

  where $\varepsilon > 0$.

Existence and Uniqueness of the Best-Response Solution

Given the above assumptions we establish the following properties concerning the best-response problem:

Proposition 3.9. *The best-response problem under Cournot competition has a unique solution, denoted $\mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}})$.*

Proof.
- Existence of the best-response solution:
  The revenue function is continuous in q_i and the feasible set is compact: it is a polyhedron included in the cube $[0, C_i]^N$. Hence, by Weierstrass' s theorem, there exists a solution to the best-response problem.

- Uniqueness of the best-response

  solution:
  The proof of uniqueness is similar to that of the Bertrand best-response problem.

□

Equivalence of the best-response problem to a variational inequality

We now establish the equivalence of the best-response problem to a variational inequality.

Proposition 3.10. *The best-response problem $\mathcal{BR}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$ is equivalent to a variational inequality in the following sense:*
$\mathbf{q_i} = \mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$ *solves the best-response problem iff it solves the following variational inequality, denoted $VI(-\nabla_i \pi_i(., \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i}), \mathfrak{Q}_i)$:*

$$-\nabla_i \pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})\prime(\tilde{\mathbf{q}}_\mathbf{i} - \mathbf{q_i}) \geq 0$$

for all $\tilde{\mathbf{q}}_\mathbf{i} \in \mathfrak{Q}_i$.

We omit the proof, since it is similar to that for the Bertrand best-response problem.

First Order Optimality Conditions

In what follows, we show that the Cournot best-response problem is equivalent to its KKT conditions.

Proposition 3.11. *Assume that there is a feasible quantity vector $\mathbf{q_i}$ for which $\pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i}) > 0$.*
Then $\mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$ is the best-response solution iff there is a scalar $\lambda_i \geq 0$, and a vector in $\mathbb{R}^T$ $\mu_\mathbf{i} \geq \mathbf{0}$ such that the following system holds at $\mathbf{q_i} = \mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$:

$$-\nabla_i \pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i}) + \lambda_i \nabla_i g_i(\mathbf{q_i}) - \sum_{t=1}^{T} \mu_i(t)\mathbf{e}(t) = 0 \tag{3.12}$$

$$\lambda_i g_i(\mathbf{q_i}) = 0 \tag{3.13}$$

$$\mu_\mathbf{i}{}'\mathbf{q_i} = 0 \tag{3.14}$$

$$g_i(\mathbf{q_i}) \leq 0 \tag{3.15}$$

$$\mathbf{0} \leq \mathbf{q_i} \tag{3.16}$$

$\lambda_i, \mu_\mathbf{i}$ are called KKT multipliers corresponding to the constraints $g_i(\mathbf{q_i}) \leq 0$, $\mathbf{q_i} \geq \mathbf{0}$.

Proof. Let us fix competitors' strategy $\overline{\mathbf{q_{-i}}}$ and parameters $\alpha_\mathbf{i}$.
Let $\mathbf{q_i} = \mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$. Let $\mathcal{I}_i(\mathbf{q_i}) = \{t \ : \ q_i(t) = 0\}$ be the set of active nonnegativity constraints.

- The KKT conditions are necessary to optimality if the set of gradients of the tight constraints is linearly independent. Since there exists a feasible quantity for which the revenue is strictly positive, then the optimal quantity cannot be zero. Hence, $\mathcal{I}_i(\mathbf{q_i}) \subsetneq \{1, \ldots, T\}$.
  Assume that capacity is not tight: then the set of gradients of active constraints is $\{\mathbf{e}(\mathbf{t}), t \in \mathcal{I}_i(\mathbf{q_i})\}$. And since $\mathcal{I}_i(\mathbf{q_i}) \subsetneq \{1, \ldots, T\}$, then the set is linearly independent.
  Assume that capacity is tight. The gradient $\nabla_i g_i(\mathbf{q_i}) = \mathbf{e} = \sum_{t=1}^{T} \mathbf{e}(\mathbf{t})$. But since $\mathcal{I}_i(\mathbf{q_i}) \subsetneq \{1, \ldots, T\}$, it follows that the gradients of active constraints are linearly independent.

- The KKT conditions are sufficient for optimality if the objective function π_i is pseudo-concave and g_i is quasi-convex in $\mathbf{p_i}$. Pseudo-concavity follows from Assumption 3.12, and quasi-convexity follows by linearity of g_i.

□

The fact that the best-response problem is equivalent to the KKT system of equations proves that the problem is NP-hard.
Furthermore, we prove that the KKT multipliers are unique.

Proposition 3.12. *Let $\mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$ be the solution to best-response problem $\mathcal{BR}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$. Assume that it is also solution to the KKT conditions. Then there is a unique scalar $\lambda_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$, and a unique vector $\mu_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$ solving the KKT conditions at $\mathbf{q_i} = \mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$*

Proof. The proof is in two parts:

1. Assume that capacity is not tight at $\mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$:
   Then $\lambda_i = 0$ and thus the KKT conditions become:

$$-\nabla_i \pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i}) - \sum_{t=1}^{T} \mu_i(t)\mathbf{e}(t) = 0$$

   For $t \in \mathcal{I}_i(\mathbf{q_i})$, we have $\mu_i(t) = 0$, else for $t \notin \mathcal{I}_i(\mathbf{q_i})$, we have $\mu_i(t) = -\frac{\partial \pi_i}{\partial q_i(t)}(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$.
   Hence, the multipliers are uniquely defined.

2. Assume capacity is tight at $\mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$. Then we have the following:

   - For $t \notin \mathcal{I}_i(\mathbf{q_i})$, we have $\mu_i(t) = 0$. Therefore the KKT conditions are:

$$-\frac{\partial \pi_i}{\partial q_i(t)}(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i}) + \lambda_i = 0$$

$$\sum_{t=1}^{T} q_i(t) = C_i$$

     Hence,, $\lambda_i = \frac{\partial \pi_i}{\partial q_i(t)}(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$, for all $t \notin \mathcal{I}_i(\mathbf{q_i})$.
   - We can now replace λ_i by its value, and solve for $\mu_i(t)$ for all $t \in \mathcal{I}_i(\mathbf{q_i})$:

$$\mu_i(t) = \lambda_i - \frac{\partial \pi_i}{\partial q_i(t)}(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})$$

   Therefore the multipliers are uniquely defined.

□

Sensitivity Analysis w.r.t Parameters

In this section, we aim at quantifying how much the best-response

quantity varies when the value of the parameters changes. We prove that under some additional assumptions, the best-response as a function of the parameters α_i is a well defined function which is piecewise continuously differentiable (PC1), whose directional derivative in any direction is given as the unique solution to a quadratic program.

We are able to leverage a key difference between the Bertrand and Cournot competitive settings. Indeed, under Cournot competition, the feasible set of the best-response problem, i.e $\mathfrak{Q}_i = \{\mathbf{q} \geq \mathbf{0} \ : \ \sum_{t=1}^{T} q_i(t) \leq C_i\}$ is a polyhedral set, independent on the value of the parameter α. Sensitivity analysis results in the polyhedral case take a somewhat simpler form.

Let us fix a vector of parameters $\bar{\alpha}_{\mathbf{i}}$. $\bar{\mathcal{Q}}_i = \mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \bar{\alpha}_{\mathbf{i}})$ is the corresponding best-response solution for competitors' strategy equal to $\overline{\mathbf{p_{-i}}}$. We define the following polyhedral cone:

- If the capacity constraint is tight, i.e $\sum_{s=1}^{T} \bar{\mathcal{Q}}_i(s) = C_i$:

$$\begin{aligned}\mathfrak{Q}_i\perp(\bar{\alpha}_{\mathbf{i}}) = \{\mathbf{x_i} \in \mathbb{R}^T \ : \ & \nabla_i \pi_i(\bar{\mathcal{Q}}_i, \overline{\mathbf{q_{-i}}}, \bar{\alpha}_{\mathbf{i}})'\mathbf{x_i} = 0 \\ & x_i(t) \geq 0 \ \forall \ t \in \mathcal{I}_i(\bar{\mathcal{Q}}_i) \\ & \sum_{s=1}^{T} x_i(s) \leq 0\}\end{aligned}$$

- If the capacity constraint is not tight, i.e $\sum_{s=1}^{T} \bar{\mathcal{Q}}_i(s) < C_i$:

$$\begin{aligned}\mathfrak{Q}_i\perp(\bar{\alpha}_{\mathbf{i}}) = \{\mathbf{x_i} \in \mathbb{R}^T \ : \ & \nabla_i \pi_i(\bar{\mathcal{Q}}_i, \overline{\mathbf{q_{-i}}}, \bar{\alpha}_{\mathbf{i}})'\mathbf{x_i} = 0 \\ & x_i(t) \geq 0 \ \forall \ t \in \mathcal{I}_i(\bar{\mathcal{Q}}_i)\}\end{aligned}$$

To gain intuition on the meaning of the cone, assume that $\mathcal{Q}_i$ is differentiable. Notice that if $\mathcal{Q}_i(\alpha_{\mathbf{i}}) \in \mathfrak{Q}_i$ is the best-response quantity for parameter value $\alpha_{\mathbf{i}}$, then direction $\mathcal{Q}_i(\alpha_{\mathbf{i}}) - \bar{\mathcal{Q}}_i$ belongs to $\{\mathbf{x_i} : x_i(t) \geq 0 \ \forall \ t \in \mathcal{I}_i(\bar{\mathcal{Q}}_i)\}$ if capacity is not tight at $\bar{\mathcal{Q}}_i$, and $\{\mathbf{x_i} : x_i(t) \geq 0 \ \forall \ t \in \mathcal{I}_i(\bar{\mathcal{Q}}_i), \ \sum_{s=1}^{T} x_i(s) \leq 0\}$ otherwise.

Furthermore, we have the following fist order approximation for π_i:

$$\begin{aligned}&\pi_i(\mathcal{Q}_i(\alpha_{\mathbf{i}}), \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}) - \pi_i(\bar{\mathcal{Q}}_i, \overline{\mathbf{q_{-i}}}, \bar{\alpha}_{\mathbf{i}}) \\ &\quad \approx \ \nabla_i \pi_i(\bar{\mathcal{Q}}_i, \overline{\mathbf{q_{-i}}}, \bar{\alpha}_{\mathbf{i}})'\mathcal{Q}_i'(\overline{\mathbf{q_{-i}}}; \mathbf{d_i}) + \nabla_i \pi_i(\bar{\mathcal{Q}}_i, \overline{\mathbf{q_{-i}}}, \bar{\alpha}_{\mathbf{i}})'\mathbf{d_i}\end{aligned}$$

Since $\mathcal{Q}_i(\alpha_{\mathbf{i}}) - \bar{\mathcal{Q}}_i \approx \mathcal{Q}_i'(\bar{\alpha}_{\mathbf{i}}; \mathbf{d_i})$ where $\mathbf{d_i} = \alpha_{\mathbf{i}} - \bar{\alpha}_{\mathbf{i}}$, then the first term of the right hand side of the equation vanishes.

We introduce the following assumption, under which the differentiability result holds:

Assumption 3.14. *The matrix* $\nabla^2_{i,i}\pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}})$ *is negative definite on the linear subspace* $\mathfrak{Q}_i - \mathfrak{Q}_i\perp(\bar{\alpha}_{\mathbf{i}}) = \{\mathbf{q_i} \ : \ \exists \mathbf{d_i} \in \mathfrak{Q}_i\perp(\bar{\alpha}_{\mathbf{i}}) \mid \mathbf{q_i} + \mathbf{d_i} \in \mathfrak{Q}_i\}$.

Note that, without the restriction on the linear subspace, the above assumption would be the negative definiteness of the Hessian matrix of the revenue function, which is equivalent to the concavity of the revenue function.

Proposition 3.13. *Under differentiability of $\nabla_i \pi_i$ w.r.t $\mathbf{q_i}$ and $\alpha_\mathbf{i}$ in a neighborhood of $(\bar{\mathcal{Q}}_i, \bar{\alpha}_\mathbf{i})$, and under Assumption 3.14, there exist neighborhoods $\mathcal{U}$ of $\bar{\alpha}_\mathbf{i}$ and $\mathcal{V}$ of $\bar{\mathcal{Q}}_i$, and a Lipschitz continuous function $\mathcal{Q}_i(.) : \mathcal{U} \mapsto \mathcal{V}$ such that:*

1. *for all $\alpha_\mathbf{i} \in \mathcal{U}$, $\mathcal{Q}_i(\alpha_\mathbf{i})$ is the solution to the best-response problem at parameter $\alpha_\mathbf{i}$;*
2. *for all $\alpha_\mathbf{i} \in \mathcal{U}$, $\mathcal{Q}_i(\alpha_\mathbf{i}) - \bar{\mathcal{Q}}_i \in \mathfrak{Q}_i\perp(\bar{\alpha}_\mathbf{i})$;*
3. *$\mathcal{Q}_i(.)$ is piecewise continuously differentiable at $\bar{\alpha}_\mathbf{i}$ and the directional derivative in any direction $\mathbf{d_i}$, denoted $\mathcal{Q}_i'(\bar{\alpha}_\mathbf{i}; \mathbf{d}_i)$ is the unique solution to the following convex quadratic optimization problem $\mathcal{QP}(\bar{\alpha}_i; \mathbf{d})$:*

$$\begin{array}{ll} \min\limits_{\mathbf{x_i}} & -\frac{1}{2}\mathbf{x_i}'\nabla^2_{i,i}\pi_i(\bar{\mathcal{Q}}_i, \overline{\mathbf{p_{-i}}}, \bar{\alpha}_\mathbf{i})\mathbf{x_i} + \mathbf{x_i}'\nabla^2_{i,\alpha_i}\pi_i(\bar{\mathcal{Q}}_i, \overline{\mathbf{p_{-i}}}, \bar{\alpha}_\mathbf{i})\mathbf{d_i} \\ s.t & \mathbf{x_i} \in \mathfrak{Q}_i\perp(\bar{\alpha}_\mathbf{i}) \end{array}$$

Proof. The proof is to be found in [71]. □

Since $\mathcal{Q}_i(.)$ is PC1, there is a finite family of continuously differentiable functions $\{\psi_1(\alpha_\mathbf{i}), \ldots, \psi_K(\alpha_\mathbf{i})\}$ such that $\mathcal{Q}_i(\alpha_\mathbf{i}) = \psi_k(\alpha_\mathbf{i})$ for some k. As a result, the directional derivative is such that:

$$\mathcal{Q}_i'(\alpha_\mathbf{i}; \mathbf{d_i}) = \nabla\psi_k(\alpha_\mathbf{i})'\mathbf{d_i}$$

so that the generalized jacobian of $\mathcal{Q}_i(.)$ is:

$$\partial\mathcal{Q}_i(\alpha_\mathbf{i}) = \{\nabla\psi_k(\alpha_\mathbf{i})\ \forall\ k \text{ s.t } \mathcal{Q}_i(\alpha_\mathbf{i}) = \psi_k(\alpha_\mathbf{i})\}$$

Furthermore, $\mathcal{Q}_i(.)$ is Bouligand differentiable, which means that the directional derivative is a first order approximation of the function:

$$\mathcal{Q}_i(\alpha_\mathbf{i} + \mathbf{d_i}) = \mathcal{Q}_i(\alpha_\mathbf{i}) + \mathcal{Q}_i'(\alpha_\mathbf{i}; \mathbf{d_i}) + o_{\mathbf{d_i}\to 0}(\|\mathbf{d_i}\|)$$

We will use this property in Chapter 6.

Sensitivity Analysis w.r.t Competitors' Strategy

It is also important for companies to track how sensitive their best-response quantity is to changes in their competitors' strategy. We can establish a similar result to the sensibility to parameters.
We fix parameters to $\alpha_\mathbf{i}$ and drop the dependence on parameters from the notation.

Proposition 3.14. *Let $\overline{\mathbf{q_{-i}}}$ and $\bar{\mathcal{Q}}_i = \mathcal{Q}_i(\overline{\mathbf{q_{-i}}})$ be the corresponding best-response solution. Under differentiability of $\nabla_i \pi_i$ w.r.t $\mathbf{q_i}$ and $\alpha_\mathbf{i}$ in a neighborhood of $(\bar{\mathcal{Q}}_i, \bar{\alpha}_\mathbf{i})$, and under Assumption 3.14, there exist a neighborhood $\mathcal{U}$ of $\overline{\mathbf{q_{-i}}}$ and $\mathcal{V}$ of $\bar{\mathcal{Q}}_i$, and a Lipschitz continuous function $\mathcal{Q}_i(.) : \mathcal{U} \mapsto \mathcal{V}$ such that:*

1. *for all* $\overline{\mathbf{q_{-i}}} \in \mathcal{U}$, $\mathcal{Q}_i(\mathbf{p_{-i}})$ *is the solution to the best-response problem at parameter* $\mathbf{p_{-i}}$*;*

2. *for all* $\mathbf{p_{-i}} \in \mathcal{U}$, $\mathcal{Q}_i(\mathbf{p_{-i}}) - \mathcal{Q}_i(\overline{\mathbf{q_{-i}}}) \in \mathfrak{Q}_i \perp (\alpha_\mathbf{i})$*;*

3. $\mathcal{Q}_i(.)$ *is piecewise continuously differentiable at* $\overline{\mathbf{q_{-i}}}$ *and the directional derivative in any direction* $\mathbf{d_{-i}}$*, denoted* $\mathcal{Q}_i'(\overline{\mathbf{q_{-i}}};\mathbf{d_{-i}})$ *is the unique solution to the following convex quadratic optimization problem* $\mathcal{QP}(\overline{\mathbf{q_{-i}}};\mathbf{d_{-i}})$*:*

$$\begin{array}{ll} \min\limits_{\mathbf{y_i}} & -\frac{1}{2}\mathbf{y_i}'\nabla^2_{i,i}\pi_i(\bar{\mathcal{Q}}_i,\overline{\mathbf{p_{-i}}})\mathbf{y_i} + \mathbf{y_i}'\nabla^2_{i,-i}\pi_i(\bar{\mathcal{Q}}_i,\overline{\mathbf{p_{-i}}})\mathbf{d_{-i}} \\ s.t & \mathbf{y_i} \in \mathfrak{Q}_i \perp (\alpha_\mathbf{i}) \end{array}$$

3.2 Study of the Equilibrium Problem

In this section, we study the equilibrium problem for known price-demand relationship. The problem was introduced in Section 2.2.3 under Bertrand competition, and in Section 2.3.3 under Cournot competition. We reformulate the equilibrium problem under Bertrand competition as a quasi variational inequality, and under Cournot competition as a variational inequality. Using tools from variational inequalities and quasi variational inequalities, we establish existence and uniqueness of the Nash equilibrium. We then perform sensitivity analysis on the Nash equilibrium, in order to quantify how changes in parameters affects the equilibrium.

3.2.1 The Bertrand Equilibrium Problem

In Section 2.2.3, we formulated the equilibrium problem as a fixed point of the best-response mappings of the firms:

$$\mathcal{P}_i(\beta) = \mathcal{P}_i(\mathcal{P}_{-i}(\beta), \beta_\mathbf{i}) \quad \forall\ i = 1, \ldots, N$$

In other words, the set of Nash equilibrium prices $(\mathcal{P}_i(\beta))_{i=1,\ldots,N}$ is the simultaneous solution to the N best-response problems of the firms.
Recall that the Bertrand equilibrium problem is a *generalized* equilibrium, in the sense that each firm's set of feasible prices depends on its competitors' strategy. As a result of this special feature of the Bertrand equilibrium problem, we need specific tools to analyze the problem. In what follows, we establish that the generalized Nash equilibrium is equivalent to a quasi variational inequality. We then prove existence of the equilibrium solution, and its uniqueness under additional assumptions.

Equivalence of the Generalized Nash Equilibrium to a Quasi Variational Inequality

A *quasi variational inequality* formulation is more general than a variational inequality formulation in the sense that instead of being solved over a set of feasible solutions, it is solved over a point-to-set mapping of feasible solutions:

Definition 3.2. *Let $X \subseteq \mathbb{R}^n$, and $f : X \to \mathbb{R}^m$ a function. Let $K : X \to 2^{\mathbb{R}^n}$ be a point-to-set mapping, i.e if $x \in X$, then $K(x)$ is a subset of $\mathbb{R}^n$. We define the quasi variational inequality, denoted $QVI(f, K)$ as follows: $x^* \in K(x^*)$ solves $QVI(f, K)$ if for all $x \in K(x^*)$, we have:*

$$f(x^*)'(x - x^*) \geq 0$$

In what follows we define as $\mathfrak{P}(\mathbf{p}, \beta) = \{\tilde{\mathbf{p}} : \tilde{\mathbf{p}}_\mathbf{i} \in \mathcal{P}_i(\mathbf{p}_{-\mathbf{i}}, \beta_\mathbf{i})\}$ the point-to-set mapping of the Bertrand generalized equilibrium. We now establish the equivalence result.

Proposition 3.15. *The generalized Nash equilibrium $\mathcal{GNE}(\beta)$ is equivalent to a quasi variational inequality in the following sense:*
$\mathcal{P}(\beta)$ is the generalized Nash equilibrium iff it solves the following QVI: for all $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_N})$ such that for all $\mathbf{p} \in \mathfrak{P}(\mathbf{p}, \beta)$:

$$\sum_{i=1}^{N} -\nabla_i \pi_i(\mathcal{P}_i(\beta), \mathcal{P}_{-i}(\beta), \beta_\mathbf{i})'(p_i - \mathcal{P}_i(\beta)) \geq 0$$

Proof. The result was established by Perakis and Sood [72]. □

Existence of a Generalized Nash Equilibrium

To establish a solution to the generalized Nash equilibrium problem, we cannot use traditional results from the game theoretic literature. Indeed, the traditional game theoretic framework assumes that the strategy set of each participant in the game is independent of the strategy used by its competitors. As mentioned in Chapter 2, this assumption is violated under Bertrand competition, due to the capacity constraint which involves competitors' strategy.
The literature on generalized Nash games is much more scarce than that on proper Nash games, and existence and uniqueness results are harder to obtain.
Existence of a generalized Nash equilibrium was established by Rosen [78] for concave games.

Proposition 3.16. *There exists a generalized Nash equilibrium solution.*

Proof. See Rosen [78]. □

Here we outline the idea of the proof. The proof relies on the application of Kakutani's fixed point theorem to a particular function. Rosen proves that a generalized Nash equilibrium solution is a fixed point of the mapping Γ such that $\Gamma(\mathbf{p}) = \arg\max_{\tilde{\mathbf{p}}} \varphi(\tilde{p}, \mathbf{p})$, where the function φ is defined on the set $\mathfrak{P} = \{\mathbf{p} : \mathbf{p_i} = \mathfrak{P}_i(\mathbf{p}_{-\mathbf{i}})\}$ as:

$$\varphi(\hat{\mathbf{p}}, \check{\mathbf{p}}) = \sum_{i=1}^{N} \pi_i(\hat{\mathbf{p}}_\mathbf{i}, \check{\mathbf{p}}_{-\mathbf{i}})$$

Uniqueness of the Generalized Nash Equilibrium

In general, as documented by Sood [81], the generalized Nash equilibrium is not unique. In this case, the Nash equilibrium has little predictive power as far as the behavior of the firms in the market is concerned. Indeed, suppose that there exist two distinct Nash equilibria $\mathbf{p}^*$ and $\mathbf{p}^{**}$. Then firm i may use $\mathbf{p_i^*}$ as pricing strategy, whereas firm j uses $\mathbf{p_j^{**}}$. But $\mathbf{p_i^*}$ is a best-response to $\mathbf{p_{-i}^*}$, not to $\mathbf{p_{-i}^{**}}$, and $\mathbf{p_j^{**}}$ is best-response to $\mathbf{p_{-j}^{**}}$, not to $\mathbf{p_{-j}^*}$.
Rosen [78] establishes a condition on the pseudo gradient of the revenue function, under which the generalized Nash equilibrium is guaranteed to have a unique solution. We first introduce the notion of *pseudo-gradient* and of *diagonal strict concavity* of the revenue function, and then state the uniqueness result.

Definition 3.3. *Let* $\sigma(\mathbf{p},\mathbf{r}) = \sum_{i=1}^N r_i\pi_i(\mathbf{p_i},\mathbf{p_{-i}})$.
The pseudo-gradient of σ *denoted* $\mathbf{h}(\mathbf{p},\mathbf{r})$ *is the* N *dimensional vector with components* $r_i\nabla_i\pi_i(\mathbf{p_i},\mathbf{p_{-i}})$.

Definition 3.4. *Let* $\sigma(\mathbf{p},\mathbf{r}) = \sum_{i=1}^N r_i\pi_i(\mathbf{p_i},\mathbf{p_{-i}})$ *with pseudo-gradient* $\mathbf{h}(\mathbf{p},\mathbf{r})$.
Let $\mathfrak{P} = \{\mathbf{p}\ :\ \forall i,\ \mathbf{p_i} \in \mathfrak{P}_i(\mathbf{p_{-i}})\}$.
σ *is said to be diagonally strictly concave if* $\forall\ \mathbf{p^1},\mathbf{p^2} \in \mathfrak{P}$ *with* $\mathbf{p_1} \neq \mathbf{p_2}$, *and for all* $\mathbf{r} \geq 0$, *we have:*

$$(\mathbf{p^1}-\mathbf{p^2})'\mathbf{h}(\mathbf{p^2},\mathbf{r}) + (\mathbf{p^2}-\mathbf{p^1})'\mathbf{h}(\mathbf{p^1},\mathbf{r}) > 0$$

In other words, σ is diagonally strictly concave if:

$$(-\mathbf{h}(\mathbf{p^1},\mathbf{r}) + \mathbf{h}(\mathbf{p^2},\mathbf{r}))'(\mathbf{p^1}-\mathbf{p^2}) > 0$$

Hence, strict pseudo-concavity of σ is equivalent to strict monotonicity of $-\mathbf{h}$.
Note that diagonally strict concavity implies for $r_i = 1$ for all i that:

$$\sum_{i=1}^N \left(-\nabla_i\pi_i(\mathbf{p_i^1},\mathbf{p_{-i}^1}) + \nabla_i\pi_i(\mathbf{p_i^2},\mathbf{p_{-i}^2})\right)'(\mathbf{p_i^1}-\mathbf{p_i^2})$$

Hence, diagonally strict concavity is more stringent than strict monotonicity of $-\nabla_i\pi_i(\mathbf{p_i},\mathbf{p_{-i}})$.
For instance, for the linear demand function, diagonal strict concavity is equivalent to: for all $\mathbf{r} \geq 0$,

$$2r_i\beta_{ii}(t) > r_i\sum_{j\neq i}\beta_{ij}(t)$$
$$2r_i\beta_{ii}(t) > \sum_{j\neq i} r_j\beta_{ji}(t)$$

We now establish the uniqueness result due to Rosen [78]:

Proposition 3.17. *Assume that the interior of* $\mathfrak{P} = \{\mathbf{p}\ :\ \mathbf{p_i} \in \mathfrak{P}_i(\mathbf{p_{-i}})\}$ *is nonempty.*
Assume furthermore that $\sigma(\mathbf{p},\mathbf{r}) = \sum_{i=1}^N r_i\pi_i(\mathbf{p_i},\mathbf{p_{-i}})$ *is strictly pseudo concave.*
Then the generalized Nash equilibrium is unique.

First Order Optimality Conditions

We prove that the generalized equilibrium is equivalent to a system of equalities and inequalities obtained from the KKT conditions for each best-response problem.

Proposition 3.18. *Assume that if capacity is tight at $\mathcal{P}_i(\beta)$, then at least one of the best-response prices is interior to the compact $[0, p_i^{max}]$.*
Then $\mathcal{P}(\beta)$ is the generalized Nash equilibrium iff there are scalars $\lambda_1, \ldots, \lambda_N \geq 0$, and $2N$ vectors in $\mathbb{R}^T$ $\underline{\mu}_\mathbf{1}, \ldots, \underline{\mu}_\mathbf{N}$ and $\overline{\mu}_\mathbf{1}, \ldots, \overline{\mu}_\mathbf{N} \geq \mathbf{0}$ such that the following system holds at $\mathbf{p} = \mathcal{P}(\beta)$:
for all $i = 1, \ldots, N$:

$$-\nabla_i \pi_i(\mathbf{p_i}, \mathbf{p_{-i}}, \beta_\mathbf{i}) + \lambda_i \nabla_i g_i(\mathbf{p_i}, \mathbf{p_{-i}}, \beta_\mathbf{i}) + \sum_{t=1}^{T} (\overline{\mu}_i(t) - \underline{\mu}_i(t))\mathbf{e}(t) = 0 \tag{3.17}$$

$$\lambda_i g_i(\mathbf{p_i}, \mathbf{p_{-i}}, \beta_\mathbf{i}) = 0 \tag{3.18}$$

$$\overline{\mu}_\mathbf{i}{}' \mathbf{p_i} = 0 \tag{3.19}$$

$$\underline{\mu}_\mathbf{i}{}' \mathbf{p_i} = 0 \tag{3.20}$$

$$g_i(\mathbf{p_i}, \mathbf{p_{-i}}, \beta_\mathbf{i}) \leq 0 \tag{3.21}$$

$$\mathbf{0} \leq \mathbf{p_i} \leq \mathbf{p_i^{max}} \tag{3.22}$$

Proof.

- Assume that $\mathcal{P}(\beta)$ is the generalized Nash equilibrium. Then for all i, $\mathcal{P}_i(\beta)$ solves the best-response problem $\mathcal{BR}_i(\mathcal{P}_{-i}(\beta), \beta_\mathbf{i})$. Hence, by Proposition 3.3, $\mathcal{P}_i(\beta)$ solves the KKT conditions with $\overline{\mathbf{p_{-i}}} = \mathcal{P}_{-i}(\beta)$. Since this is true for all i, then the above system holds at $\mathbf{p} = \mathcal{P}(\beta)$.
- Assume that the above system holds at $\mathbf{p} = \mathcal{P}(\beta)$. Then in particular the KKT conditions hold for each i for $\overline{\mathbf{p_{-i}}} = \mathcal{P}_{-i}(\beta)$. Hence, by virtue of Proposition 3.3, $\mathcal{P}_i(\beta)$ solves the best-response problem $\mathcal{BR}_i(\mathcal{P}_{-i}(\beta), \beta_\mathbf{i})$ for each i. By definition, $\mathcal{P}(\beta)$ is therefore the generalized Nash equilibrium.

□

Sensitivity Analysis

We focus on quantifying how much the generalized Nash equilibrium changes, as the parameters change. We state that the generalized Nash equilibrium solution is directionally differentiable. The result is due to Outrata [70] and relies on the quasi variational inequality formulation of the generalized Nash equilibrium. Toward this end, we introduce some notation.
Let $\mathbf{p}, \beta$ be a set of prices and parameters. For ease of notation, we denote the constraints of the Nash equilibrium by $h_k(\mathbf{p}, \mathbf{p}', \beta)$, for $k = 1, \ldots, (2T+1)N$:

$$h_k(\mathbf{p}, \mathbf{p}', \beta) = \sum_{t=1}^{T} q_i(p_i(t), p'_{-i}(t), \beta_i(t)) - C_i, \ k = 1, \ldots, N$$
$$h_k(\mathbf{p}, \mathbf{p}', \beta) = p_i(t) - p_i^{max}, \ k = N+1, \ldots, N(T+1)$$
$$h_k(\mathbf{p}, \mathbf{p}', \beta) = -p_i(t), \ k = N(T+1)+1, \ldots, N(2T+1)$$

Hence, the feasible mapping of the QVI writes:

$$\mathfrak{P}(\mathbf{p},\beta) = \{\mathbf{p}' \ : \ h_k(\mathbf{p},\mathbf{p}',\beta) \leq 0 \ \forall \ i = 1,\ldots,N\}$$

We denote by $\mathbf{F}(\mathbf{p},\beta)$ its objective function, i.e:

$$\mathbf{F_i}(\mathbf{p},\beta) = -\nabla_i \pi_i(\mathbf{p}_i, , \mathbf{p_{-i}}, \beta_\mathbf{i})$$

The QVI writes: find $\mathbf{p} \in \mathfrak{P}(\mathbf{p},\beta)$ such that the following holds for all $\mathbf{p}' \in \mathfrak{P}(\mathbf{p},\beta)$:

$$\mathbf{F}(\mathbf{p},\beta)'(\mathbf{p}' - \mathbf{p}) \geq 0$$

It is well-known that the QVI is equivalent to finding the fixed point of the projection:

$$\mathbf{p} = \mathrm{Proj}_{\mathfrak{P}(\mathbf{p},\beta)}(\mathbf{p} - \mathbf{F}(\mathbf{p},\beta))$$

We denote by $Z(\mathbf{p},\beta)$ the projection operator, and by $W(\mathbf{p},\beta) = \mathbf{p} - Z(\mathbf{p},\beta)$.
For ease of notation, we drop the dependence in β and simply denote by $\mathcal{P}$ the generalized Nash equilibrium when the parameters are equal to β.
Let $\mathcal{I}(\mathcal{P},\beta) = \{k \ : \ h_k(\mathcal{P},\mathcal{P},\beta) = 0\}$, $\mathcal{I}^+(\mathcal{P},\beta) = \{k \in \mathcal{I}(\mathcal{P},\beta) : \lambda_k > 0\}$ where λ_k is the multiplier corresponding to constraint k, and $\mathcal{I}^0(\mathcal{P},\beta) = \{k \in \mathcal{I}(\mathcal{P},\beta) : \lambda_k = 0\}$.
We denote by $\mathbf{H}(\mathbf{p},\mathbf{p}',\beta)$ the vector with components $h_k(\mathbf{p},\mathbf{p}',\beta)$, $k = 1,\ldots,N(2T+1)$.
We define the Lagrangian function of the projection operator as:

$$\mathcal{L}_Z(\mathbf{p},\mathbf{p}',\beta,\lambda) = \mathbf{p}' - \mathbf{p} + \mathbf{F}(\mathbf{p},\beta) + \nabla_{\mathbf{p}'}\mathbf{H}(\mathbf{p},\mathbf{p}',\beta)'\lambda$$

The Lagrangian function associated with the QVI is:

$$\mathcal{L}(\mathbf{p},\beta,\lambda) = \mathbf{F}(\mathbf{p},\beta) + \nabla_{\mathbf{p}'}\mathbf{H}(\mathbf{p},\mathbf{p},\beta)'\lambda$$

In particular, we have: $\mathcal{L}(\mathbf{p},\beta,\lambda) = \mathcal{L}_Z(\mathbf{p},\mathbf{p},\beta,\lambda)$.
If I is an index set, then $\mathbf{M}_I$ denotes the submatrix of row vectors $M_i, \ i \in I$ and $\mathbf{v}_I$ denotes the subvector of components $v_i \ i \in I$.
We introduce the following additional assumption, in terms of the Lagrangian of the projection operator Z, under which directional differentiability holds:

Assumption 3.15. *For all $i \in \mathcal{I}^0(\mathcal{P},\beta)$, there exist matrices $\mathbf{A_i},\mathbf{B_i},\mathbf{C_i},\mathbf{D_i}$ solving the following equations:*

$$\begin{aligned}
-\nabla_{\mathbf{p}'}\mathcal{L}_Z(\mathbf{p},\mathbf{p}',\beta,\lambda)\mathbf{A_i} + \nabla_{\mathbf{p}'}\mathbf{H}_{\mathcal{I}^+\bigcup\{i\}}(\mathbf{p},\mathbf{p}',\beta)'\mathbf{B_i} &= -\nabla_\beta\mathcal{L}_Z(\mathbf{p},\mathbf{p}',\beta,\lambda)\\
\nabla_{\mathbf{p}'}\mathbf{H}_{\mathcal{I}^+\bigcup\{i\}}(\mathbf{p},\mathbf{p}',\beta)\mathbf{A_i} &= \nabla_\beta\mathbf{H}_{\mathcal{I}^+\bigcup\{i\}}(\mathbf{p},\mathbf{p}',\beta)\\
\nabla_{\mathbf{p}'}\mathcal{L}_Z(\mathbf{p},\mathbf{p}',\beta,\lambda)(\mathbf{E}-\mathbf{C_i}) + \nabla_{\mathbf{p}'}\mathbf{H}_{\mathcal{I}^+\bigcup\{i\}}(\mathbf{p},\mathbf{p}',\beta)'\mathbf{D_i} &= -\nabla_{\mathbf{p}}\mathcal{L}_Z(\mathbf{p},\mathbf{p}',\beta,\lambda)\\
-\nabla_{\mathbf{p}'}\mathbf{H}_{\mathcal{I}^+\bigcup\{i\}}(\mathbf{p},\mathbf{p}',\beta)(\mathbf{E}-\mathbf{C_i}) &= \nabla_{\mathbf{p}}\mathbf{H}_{\mathcal{I}^+\bigcup\{i\}}(\mathbf{p},\mathbf{p}',\beta)
\end{aligned}$$

where $\mathbf{E}$ is the matrix of coefficients equal to 1.

Proposition 3.19. *Under Assumptions 3.1 through 3.15, the generalized Nash equilibrium solution $\mathcal{P}(\beta)$ is directionally differentiable, and its directional derivative in any direction $\mathbf{d} \in \mathbb{R}^{NT}$, denoted $\mathcal{P}'(\beta;\mathbf{d})$ is solution to the following linear quasi variational inequality: $\mathbf{x} = \mathcal{P}'(\beta;\mathbf{d})$ solves for all $\mathbf{y} \in \mathcal{C}'(\mathbf{x})$:*

$$(\nabla_{\mathbf{p}}\mathcal{L}(\mathbf{p},\beta,\lambda)\mathbf{x} + \nabla_{\beta}\mathcal{L}(\mathbf{p},\beta,\lambda)\mathbf{d})'\,(\mathbf{y}-\mathbf{x})) \geq 0$$

where $\mathcal{C}'(.)$ is the point-to-set mapping:

$$\begin{aligned}\mathcal{C}'(\mathbf{x}) = \{\mathbf{y} : \nabla_{\beta}h_k(\mathbf{p},\mathbf{p},\beta)'\mathbf{d} + \nabla_{\mathbf{p}}h_k(\mathbf{p},\mathbf{p},\beta)'\mathbf{x} + \nabla_{\mathbf{p}'}h_k(\mathbf{p},\mathbf{p},\beta)'\mathbf{y} = 0 \;\forall\; k \in \mathcal{I}^+(\mathbf{p},\beta)\\ \nabla_{\beta}h_k(\mathbf{p},\mathbf{p},\beta)'\mathbf{d} + \nabla_{\mathbf{p}}h_k(\mathbf{p},\mathbf{p},\beta)'\mathbf{p} + \nabla_{\mathbf{p}'}h_k(\mathbf{p},\mathbf{p},\beta)'\mathbf{y} = 0 \;\forall\; k \in \mathcal{I}^0(\mathbf{p},\beta)\}\end{aligned}$$

3.2.2 The Cournot Equilibrium Problem

In section 2.3.3, we formulated the market equilibrium problem under Cournot competition as fixed point of the best-response mappings of all the firms:

$$\mathcal{Q}_i(\alpha) = \mathcal{Q}_i(\mathcal{Q}_{-i}(\alpha),\alpha_{\mathbf{i}}) \quad \forall i = 1,\dots,N$$

In other words, the set of Nash equilibrium quantities $(\mathcal{Q}_i(\alpha))_{i=1,\dots,N}$ is the simultaneous solution to the N best-response problems of the firms.
Unlike its Bertrand counterpart, the Cournot equilibrium problem is a *proper* Nash equilibrium. In what follows, we prove equivalence of the Cournot Nash equilibrium to a variational inequality, existence of its solution, and uniqueness under additional assumption. We then perform sensitivity analysis on the Nash equilibrium, quantifying how much the equilibrium is affected by changes in the parameters.

Equivalence of the Nash Equilibrium to a Variational Inequality

The feasible set of the Cournot Nash equilibrium is denoted $\mathfrak{Q} = \mathfrak{Q}_i \times \mathfrak{Q}_N$, where $\mathfrak{Q}_i$ are the feasible set of the best-response problems.

Proposition 3.20. *The Cournot Nash equilibrium $\mathcal{Q}(\alpha)$ is equivalent to a variational inequality in the following sense:*
$\mathcal{Q}(\alpha)$ is the Nash equilibrium iff it solves the following variational inequality: for all $\mathbf{q} \in \mathfrak{Q}$, we have:

$$\sum_{i=1}^{N} -\nabla_i\pi_i(\mathcal{Q}_i(\alpha),\mathcal{Q}_{-i}(\alpha),\alpha_{\mathbf{i}})'(\mathbf{q_i} - \mathcal{Q}_i(\alpha)) \geq 0$$

Proof.
- Assume $\mathcal{Q}(\alpha)$ is solution to the Nash equilibrium, i.e for all i, $\mathcal{Q}_i(\alpha)$ is solution to the best-response problem $\mathcal{BR}_i(\mathcal{Q}_{-i}(\alpha),\alpha_{\mathbf{i}})$. Thus for all i, $\mathcal{Q}_i(\alpha)$ solves the variational inequality of Proposition 3.10.
  Summing up the VIs yields: For all $\mathbf{q} = (\mathbf{q_1},\dots,\mathbf{q_N})$, such that for all i, $\mathbf{q_i} \in \mathfrak{Q}_i$:

$$\sum_{i=1}^{N} -\nabla_i\pi_i(\mathcal{Q}_i(\alpha),\mathcal{Q}_{-i}(\alpha),\alpha_{\mathbf{i}})'(\mathbf{q_i} - \mathcal{Q}_i(\alpha)) \geq 0$$

- Assume that $\mathcal{Q}(\alpha)$ solves the above variational inequality. Pick $i \in \{1,\ldots,N\}$. Let $\mathbf{q_i} \in \mathfrak{Q}_i$. Let $\mathbf{q} = (\mathcal{Q}_1(\alpha),\ldots,\mathcal{Q}_{i-1}(\alpha),\mathbf{q_i},\mathcal{Q}_{i+1}(\alpha),\ldots,\mathcal{Q}_N(\alpha))$ and apply the variational inequality to $\mathbf{q}$: all the terms $j \neq i$ cancel and there remains:

$$-\nabla_i \pi_i(\mathcal{Q}_i(\alpha),\mathcal{Q}_{-i}(\alpha),\alpha_\mathbf{i})'(\mathbf{q_i} - \mathcal{Q}_i(\alpha)) \geq 0$$

Hence, by Proposition 3.10, $\mathcal{Q}_i(\alpha)$ solves $\mathcal{BR}_i(\mathcal{Q}_{-i}(\alpha),\alpha_\mathbf{i})$. Since this holds for all i, then $\mathcal{Q}(\alpha)$ is the Nash equilibrium.

□

Uniqueness of the Nash Equilibrium

Without any additional assumption, the Cournot Nash equilibrium might not be unique. To see this, consider the linear allocation model, with the following set of parameters:

$$\begin{aligned} \alpha_{ii}(t) &= \alpha & \forall\, i,t \\ \alpha_{ij}(t) &= 2\alpha & \forall\, j \neq i,t \\ \alpha_{i0}(t) &= \alpha_0 & \forall\, i \end{aligned}$$

Assume that $C_i > \frac{\alpha_0}{4\alpha}T$ for all i, and that $C_i = C_j = C$ for all $i, j \neq i$.
Then the first order optimality conditions yield: for all i:

$$\alpha_0 - 2\alpha(q_i(t) - \frac{1}{N-1}\sum_{j\neq i} q_j(t)) = 0$$

Hence, in matrix notation, the first order optimality conditions are:

$$\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \mathbf{q(t)} = \frac{\alpha_0}{2\alpha}\mathbf{e}$$

Since the matrix is not invertible, several solutions exist.
We therefore introduce an additional assumptions sufficient for uniqueness of the equilibrium.
Let $\mathbf{F}(\mathbf{q},\alpha)$ the function such that for all i, $\mathbf{F_i}(\mathbf{q},\alpha) = -\nabla_i \pi_i(\mathbf{q_i},\mathbf{q_{-i}},\alpha_\mathbf{i})$.
We also denote by $\mathfrak{Q} = \{\mathbf{q} \ : \ \mathbf{Aq} \leq \mathbf{b}\}$ the polyhedral feasible set of the Cournot Nash equilibrium.

Assumption 3.16. $\mathbf{F}(\mathbf{q},\alpha)$ *is strictly monotone on* $\mathfrak{Q} = \mathfrak{Q}_1 \times \ldots \times \mathfrak{Q}_N$.
In other words, $\forall\ \hat{\mathbf{q}},\check{\mathbf{q}} \in \mathfrak{Q}$ *such that* $\hat{\mathbf{q}} \neq \check{\mathbf{q}}$*, we have:*

$$(\mathbf{F}(\hat{\mathbf{q}},\alpha) - \mathbf{F}(\check{\mathbf{q}},\alpha))'(\hat{\mathbf{q}} - \check{\mathbf{q}}) > 0$$

Proposition 3.21. *Under Assumption 3.16, The Nash equilibrium is unique.*

Proof. The Nash equilibrium is equivalent to the following variational inequality: $\forall\ \mathbf{q} \in \mathfrak{Q}$:

$$\mathbf{F}(\mathcal{Q}(\alpha), \alpha)'(\mathbf{q} - \mathcal{Q}(\alpha)) \geq 0$$

It is a well-known result that a variational inequality with strictly monotone function over a convex and compact set has a unique solution (see for instance Nagurney [68]). □

First Order Optimality Conditions

We prove that the Nash equilibrium is equivalent to a system of equalities and inequalities obtained from the KKT conditions of the firms' best-response problems.

Proposition 3.22. *Assume that there is a feasible vector* $\mathbf{q}$ *for which* $\pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i}) > 0$ *for all* $i = 1, \ldots, N$.
Then $\mathcal{Q}(\alpha)$ *is the Nash equilibrium solution iff there are scalars* $\lambda_1, \ldots, \lambda_N \geq 0$, *and* N *vectors in* $\mathbb{R}^T$ $\mu_\mathbf{1}, \ldots, \mu_\mathbf{N} \geq \mathbf{0}$ *such that the following system holds at* $\mathbf{q} = \mathcal{Q}(\alpha)$*:*
for all $i = 1, \ldots, N$*:*

$$-\nabla_i \pi_i(\mathbf{q_i}, \mathbf{q_{-i}}, \alpha_\mathbf{i}) + \lambda_i \nabla_i g_i(\mathbf{q_i}) - \sum_{t=1}^{T} \mu_i(t)\mathbf{e}(t) = 0 \tag{3.23}$$

$$\lambda_i g_i(\mathbf{q_i}) = 0 \tag{3.24}$$

$$\mu_\mathbf{i}{}'\mathbf{q_i} = 0 \tag{3.25}$$

$$g_i(\mathbf{q_i}) \leq 0 \tag{3.26}$$

$$\mathbf{0} \leq \mathbf{q_i} \tag{3.27}$$

The proof is very similar to the Bertrand case and is therefore omitted.

Sensitivity Analysis

We focus on quantifying changes in the Nash equilibrium as the parameters change. We now prove that the Nash equilibrium is directionally differentiable, and that the directional derivative is obtained as the solution to an affine variational inequality. We utilize results from sensitivity analysis for variational inequalities, making explicit use of the fact that the feasible set of the variational inequality corresponding to the Cournot Nash equilibrium is polyhedral.
Let us fix $\bar{\alpha} \in \mathcal{A}$. We define the following polyhedral cone at $\mathcal{Q}(\alpha)$ denoted $\mathfrak{Q}\perp$:

$$\mathfrak{Q}\perp = \{\mathbf{d} \ : \ \mathbf{F}(\mathcal{Q}(\bar{\alpha}), \alpha)'\mathbf{d} = 0 \\ \mathbf{A}_\mathcal{I}\mathbf{q} \leq \mathbf{0}\}$$

where $\mathbf{A}_\mathcal{I}$ denotes the rows of the matrix $\mathbf{A}$ corresponding to the active constraints of $\mathbf{Aq} \leq \mathbf{b}$, and $\mathbf{b}_\mathcal{I}$ is the corresponding subvector, so that $\mathbf{A}_\mathcal{I}\mathbf{q} = \mathbf{b}_\mathcal{I}$.
With this notation at hand, we introduce the following assumption sufficient for the differentiability result to hold.

Assumption 3.17. *The matrix* $\nabla_{\mathbf{q}}\mathbf{F}(\mathcal{P}(\bar{\alpha}), \bar{\alpha})$ *is positive definite on the span of* $\mathfrak{Q}\perp$.

Proposition 3.23. *Under differentiability assumption of* $\mathbf{F}$ *w.r.t* $\mathbf{q}$ *and* α*, and under Assumption 3.17, then there exist neighborhoods* $\mathcal{U}$ *of* $\bar{\alpha}$ *and* $\mathcal{V}$ *of* $\bar{\mathcal{Q}}$ *and a Lipschitz continuous function* $\mathcal{Q}(.) : \mathcal{U} \mapsto \mathcal{V}$ *such that:*

1. *for all* $\alpha \in \mathcal{U}$, $\mathcal{Q}(\alpha)$ *is the solution to the best-response problem at parameter* α*;*

2. *for all* $\alpha \in \mathcal{U}$, $\mathcal{Q}(\alpha) - \mathcal{Q}(\bar{\alpha}) \in \mathfrak{Q}\perp$*;*

3. $\mathcal{Q}(.)$ *is piecewise continuously differentiable at* $\bar{\alpha}$ *and the directional derivative in any direction* $\mathbf{d}$*, denoted* $\mathcal{Q}'(\bar{\alpha}; \mathbf{d})$ *is the unique solution to the following affine variational inequality:* $\mathbf{x} = \mathcal{Q}'(\bar{\alpha}; \mathbf{d})$ *solves for all* $\mathbf{x}' \in \mathfrak{Q}\perp$*:*

$$\left(\nabla_{\mathbf{q}}\mathbf{F}(\bar{\mathcal{Q}}, \alpha)\mathbf{x} + \nabla_{\alpha}\mathbf{F}(\bar{\mathcal{Q}}, \alpha)\mathbf{d}\right)' (\mathbf{x}' - \mathbf{x}) \geq 0$$

Chapter 4

Dynamic Policies with Learning

In this chapter, we consider the joint dynamic control policies with learning. That is, unlike Chapter 3, where the parameters of the price-demand relationship were known, here we consider that the parametric form of the price-demand relationship is known, but the value of its parameters are unknown and thus need to be learned. The goal of the firms in this setting is therefore twofold. First, the firms seek to learn the price-demand relationship. For this purpose, they use a data-driven approach, dynamically updating their estimates of the parameters of the price-demand relationship with the most up-to-date market information. The firms also seek to find the optimal pricing or allocation policies.

The content of this chapter is as follows. We first study the market in disequilibrium, i.e when learning of the equilibrium strategies and of the price-demand relationship occur concomitantly. Under Bertrand and Cournot competition, we give two alternative formulations of the joint dynamic control with learning: as a mixed integer program and as a bilevel problem. We then study the market in equilibrium, i.e when learning of the equilibrium strategies has occurred and the firms seek to learn the price-demand relationship. Under Bertrand and Cournot competition, we give two alternative formulations of the joint dynamic control with learning: as a mixed integer program, and as a Mathematical Program with Equilibrium Constraints (MPEC).

4.1 Dynamic Policies with Learning in Disequilibrium

In this section, we consider the problem faced by each firm at each period in the selling horizon when the market is in disequilibrium. In Chapter 2, we have seen that each firm wants to achieve two goals: learn the price-demand relationship, and find its optimal strategy, given its belief regarding its competitors' behavior. We formulated the problem as a two-step problem: first, each firm determines its best-response strategy, assuming the parameters governing the price-demand relationship are known. Then each firm updates its estimate of those parameters, integrating the most up-to-date market information.

We show that these two steps can be performed jointly and formulate them as a single problem. They are indeed equivalent to a bilevel problem, and also to a mixed integer program.

4.1.1 Dynamic Pricing with Learning in Disequilibrium

We focus on the problem under Bertrand competition. At period t, if firm i believes its competitors use pricing strategy $\overline{\mathbf{p_{-i}}}$, then the first two steps of the dynamic pricing with learning problem are as follows:

1. Computation of the best-response strategy, assuming the price sensitivities $\beta_{\mathbf{i}}$ are known:

$$\begin{aligned} \max_{\mathbf{p_i}} \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}}) &= \textstyle\sum_{t=1}^{T} p_i(t) q_i(p_i(t), \overline{p_{-i}(t)}, \beta_i(t)) \\ \text{s.t} \quad & \textstyle\sum_{t=1}^{T} q_i(p_i(t), \overline{p_{-i}(t)}, \beta_i(t)) \leq C_i \\ & 0 \leq p_i(t) \leq p_i^{max} \end{aligned}$$

2. Estimation of the price sensitivities, assuming that the best-response of firm i is known:

$$\min_{\beta_i \in \mathcal{B}_i} \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} \left(\widehat{p_i}^h(s) - \mathcal{P}_i(s, \mathbf{p_{-i}}(\mathbf{h}), \beta_{\mathbf{i}}) \right)^2 \} + \sum_{s=1}^{T} \left(\widehat{p_i}^0(s) - \mathcal{P}_i(s, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}}) \right)^2$$

We show using the reformulations of the best-response problem given in Chapter 3 that the two steps can be formulated as a single problem in two ways: as a bilevel problem, or as a mixed integer program.

Formulation as a Bilevel Problem

Step 2 takes as input the solution of Step 1. Thus Steps 1 and 2 can be seen as the leader and follower of a Stackelberg game:

- Step 1 is the follower: its strategic variable is the price $\mathbf{p_i}$, which is chosen once the parameter $\beta_{\mathbf{i}}$ is known.
- Step 2 represents the leader: its strategic variable is $\beta_{\mathbf{i}}$, and is chosen knowing the strategy of the follower.

As a result Steps 1 and 2 together can be formulated as a bilevel problem: the objective of the leader forms the upper level optimization problem, whereas the objective of the follower is the lower level:

$$\min_{\beta_i \in \mathcal{B}_i, \mathbf{p_i}} \quad \textstyle\sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} \left(\widehat{p_i}^h(s) - p_i(s) \right)^2 \} + \sum_{s=1}^{T} \left(\widehat{p_i}^0(s) - p_i(s) \right)^2 \tag{4.1}$$

$$\text{s.t} \ \max_{\mathbf{p_i}} \quad \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_{\mathbf{i}}) = \textstyle\sum_{t=1}^{T} p_i(t) q_i(p_i(t), \overline{p_{-i}(t)}, \beta_i(t)) \tag{4.2}$$

$$\sum_{t=1}^{T} q_i(p_i(t), \overline{p_{-i}(t)}, \beta_i(t)) \leq C_i \quad (4.3)$$
$$0 \leq p_i(t) \leq p_i^{max} \quad (4.4)$$

Formulation as a Mixed Integer Program

In 3.1.1, we proved that the best-response problem for known price sensitivities is equivalent to the system of equalities and inequalities formed by its KKT conditions. Hence, in the above bilevel problem, the lower level best-response problem can be replaced by the system of KKT conditions. This yields the following single-level optimization problem:

$$\begin{aligned}
\min_{\beta_i, \mathbf{p_i}, \overline{\mu}_\mathbf{i}, \underline{\mu}_\mathbf{i}, \lambda_i} \quad & \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} \left(\widehat{p_i}^h(s) - p_i(s) \right)^2 \} + \sum_{s=1}^{T} \left(\widehat{p_i}^0(s) - p_i(s) \right)^2 \\
\text{s.t} \quad & -\nabla_i \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) + \lambda_i \nabla_i g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) + \sum_{t=1}^{T} (\overline{\mu}_i(t) - \underline{\mu}_i(t)) \mathbf{e}(t) = 0 \\
& \lambda_i g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) = 0 \\
& \overline{\mu}_\mathbf{i}'(\mathbf{p_i} - \mathbf{p_i}^{\max}) = 0 \\
& \underline{\mu}_\mathbf{i}' \mathbf{p_i} = 0 \\
& \beta_i \in \mathcal{B}_i \\
& g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \leq 0 \\
& \mathbf{0} \leq \mathbf{p_i} \leq \mathbf{p_i^{max}} \\
& \lambda_i \geq 0, \overline{\mu}_\mathbf{i}, \underline{\mu}_\mathbf{i} \geq \mathbf{0}
\end{aligned}$$

As noted in 3.1.1, the KKT conditions contain the disjunctive constraints:

$$\begin{aligned}
\lambda_i g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) &= 0 \\
\overline{\mu}_\mathbf{i}'(\mathbf{p_i} - \mathbf{p_i}^{\max}) &= 0 \\
\underline{\mu}_\mathbf{i}' \mathbf{p_i} &= 0
\end{aligned}$$

A disjunctive constraint of the form $x.y = 0$ can be reformulated as a mixed integer constraint by introducing two auxiliary binary variables $x', y' \in \{0, 1\}$ and replacing constraint $x.y = 0$ by:

$$\begin{aligned}
x &\leq Mx' \\
y &\leq My' \\
x' + y' &\leq 1
\end{aligned}$$

where M is taken large enough so that $x, y \leq M$.
Consider now a disjunctive constraint of the form $\mathbf{x'y} = 0$. We introduce two vectors of binary variables $\mathbf{x'}, \mathbf{y'}$ and replace the former constraint by:

$$\begin{aligned}
\mathbf{x} &\leq M\mathbf{x'} \\
\mathbf{y} &\leq M\mathbf{y'}
\end{aligned}$$

$$\mathbf{x}' + \mathbf{y}' \leq \mathbf{e}$$

As a result, we can rewrite the bilevel problem as a single level mixed integer program.

$$\begin{array}{rc}
\min & \sum_{h=-H}^{-1}\{\sum_{s=1}^{T}\left(\widehat{p_i}^h(s) - p_i(s)\right)^2\} + \sum_{s=1}^{T}\left(\widehat{p_i}^0(s) - p_i(s)\right)^2 \\
\text{s.t} & -\nabla_i \pi_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) + \lambda_i \nabla_i g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) + \sum_{t=1}^{T}(\overline{\mu}_i(t) - \underline{\mu}_i(t))\mathbf{e}(t) = 0 \\
& \lambda_i \leq M x_i \\
& g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \leq M x_i' \\
& x_i + x_i' \geq 1 \\
& \overline{\mu}_\mathbf{i} \leq M \bar{\mathbf{y}}_\mathbf{i} \\
& \mathbf{p_i} - \mathbf{p_i}^{\max} \leq M \bar{\mathbf{y}}_\mathbf{i}' \\
& \bar{\mathbf{y}}_\mathbf{i} + \bar{\mathbf{y}}_\mathbf{i}' \geq M\mathbf{e} \\
& \underline{\mu}_\mathbf{i} \leq M \underline{\mathbf{y}}_\mathbf{i} \\
& \mathbf{p_i} \leq M \underline{\mathbf{y}}_\mathbf{i}' \\
& \underline{\mathbf{y}}_\mathbf{i} + \underline{\mathbf{y}}_\mathbf{i}' \geq \mathbf{e} \\
& g_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \leq 0 \\
& \mathbf{0} \leq \mathbf{p_i} \leq \mathbf{p_i^{max}} \\
& \lambda_i \geq 0, \overline{\mu}_\mathbf{i}, \underline{\mu}_\mathbf{i} \geq \mathbf{0} \\
& x_i, x_i' \in \{0,1\}, \overline{\mu}_\mathbf{i}, \overline{\mu}_\mathbf{i}', \underline{\mu}_\mathbf{i}, \underline{\mu}_\mathbf{i}' \in \{0,1\}^T \\
& \beta_\mathbf{i} \in \mathcal{B}_i
\end{array}$$

4.1.2 Dynamic Allocation with Learning in Disequilibrium

We focus on the problem under Cournot competition. At period t, if firm i believes its competitors use allocation strategy $\overline{\mathbf{p_{-i}}}$, then the first two steps of the dynamic allocation with learning problem are as follows:

1. Computation of the best-response allocation strategy, assuming the allocation sensitivities $\alpha_\mathbf{i}$ are known:

$$\begin{array}{rl}
\max\limits_{\mathbf{q_i}} \pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i}) & = \sum_{t=1}^{T} q_i(t) p_i(q_i(t), \overline{q_{-i}(t)}, \alpha_i(t)) \\
\text{s.t} & \sum_{t=1}^{T} q_i(t) \leq C_i \\
& 0 \leq q_i(t)
\end{array}$$

2. Estimation of the allocation sensitivities, assuming that the best-response of firm i is known:

$$\min_{\alpha_\mathbf{i} \in \mathcal{A}_i} \sum_{h=-H}^{-1}\{\sum_{s=1}^{T}\left(\widehat{q_i}^h(s) - \mathcal{Q}_i(s, \mathbf{p_{-i}}(\mathbf{h}), \alpha_\mathbf{i})\right)^2\} + \sum_{s=1}^{T}\left(\widehat{q_i}^0(s) - \mathcal{Q}_i(s, \overline{\mathbf{q_{-i}}}, \alpha_\mathbf{i})\right)^2$$

We show using the reformulations of the best-response problem given in Chapter 3 that the two steps can be formulated as a bilevel problem, or as a mixed integer program.

Formulation as a Bilevel Problem

Similarly to the Bertrand case, Step 2 of the Cournot case can be considered as the leader, and Step 1 of the Cournot case as the follower of a Stackelberg game. As a result, Step 2 is the upper level objective of a bilevel problem with Step 1 as the lower-level objective:

$$
\begin{array}{ll}
\min\limits_{\alpha_{\mathbf{i}} \in \mathcal{A}_i, \mathbf{q_i}} & \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} \left(\widehat{q}_i^{\,h}(s) - \mathcal{Q}_i(s, \mathbf{q_{-i}}(\mathbf{h}), \alpha_{\mathbf{i}}) \right)^2 \} + \sum_{s=1}^{T} \left(\widehat{q}_i^{\,0}(s) - \mathcal{Q}_i(s, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}) \right)^2 \\
\text{s.t} & \max_{\mathbf{q_i}} \pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}) = \sum_{t=1}^{T} q_i(t) p_i(q_i(t), \overline{q_{-i}(t)}, \alpha_i(t)) \\
& \sum_{t=1}^{T} q_i(t) \leq C_i \\
& 0 \leq q_i(t)
\end{array}
$$

Formulation as a Mixed Integer Program

In 3.2.1, we proved that the Cournot best-response problem for known allocation sensitivities is equivalent to the system of equalities and inequalities formed by its KKT conditions. Hence, in the above bilevel problem, the lower level best-response problem can be replaced by the system of KKT conditions. This yields the following single-level optimization problem:

$$
\begin{array}{ll}
\min\limits_{\alpha_i \in \mathcal{A}_i, \mathbf{q_i} \lambda_i, \mu_{\mathbf{i}}} & \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} \left(\widehat{q}_i^{\,h}(s) - \mathcal{Q}_i(s, \mathbf{q_{-i}}(\mathbf{h}), \alpha_i) \right)^2 \} + \sum_{s=1}^{T} \left(\widehat{q}_i^{\,0}(s) - \mathcal{Q}_i(s, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}) \right)^2 \\
\text{s.t} & -\nabla_i \pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}) + \lambda_i \nabla_i g_i(\mathbf{q_i}) - \sum_{t=1}^{T} \mu_i(t) \mathbf{e}(t) = 0 \\
& \lambda_i g_i(\mathbf{q_i}) = 0 \\
& \mu_{\mathbf{i}}' \mathbf{q_i} = 0 \\
& g_i(\mathbf{q_i}) \leq 0 \\
& \mathbf{0} \leq \mathbf{q_i}
\end{array}
$$

The disjunctive complementary slackness constraints can be reformulated by introducing binary variables. As a result, the above optimization problem can be reformulated as a mixed integer program as follows:

$$
\begin{array}{ll}
\min & \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} \left(\widehat{q}_i^{\,h}(s) - \mathcal{Q}_i(s, \mathbf{q_{-i}}(\mathbf{h}), \alpha_{\mathbf{i}}) \right)^2 \} + \sum_{s=1}^{T} \left(\widehat{q}_i^{\,0}(s) - \mathcal{Q}_i(s, \overline{\mathbf{q_{-i}}}, \alpha_i) \right)^2 \\
\text{s.t} & -\nabla_i \pi_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}) + \lambda_i \nabla_i g_i(\mathbf{q_i}) - \sum_{t=1}^{T} \mu_i(t) \mathbf{e}(t) = 0 \\
& \lambda_i \leq M x_i \\
& g_i(\mathbf{q_i}) \leq M x_i' \\
& x_i + x_i' \leq 1
\end{array}
$$

$$
\begin{aligned}
\mu_{\mathbf{i}} &\leq M\mathbf{y_i} \\
\mathbf{q_i} &\leq M\mathbf{y'_i} \\
\mathbf{y_i} + \mathbf{y'_i} &\leq \mathbf{e} \\
g_i(\mathbf{q_i}) &\leq 0 \\
\mathbf{0} &\leq \mathbf{q_i} \\
x_i, x'_i \in \{0,1\}, \mathbf{y_i}, \mathbf{y'_i} &\in \{0,1\}^T \\
\alpha_{\mathbf{i}} &\in \mathcal{A}_i
\end{aligned}
$$

4.2 Dynamic Policies with Learning in Equilibrium

In this section, we consider the problem faced by each firm at each period in the selling horizon when the market has achieved equilibrium. In Chapter 2, we have seen that each firm wants to achieve two goals: learn the price-demand relationship for itself and its competitors, and find the Nash equilibrium strategies. We formulated the problem as a two-step problem: first, each firm determines the Nash equilibrium strategies, assuming the parameters governing the price-demand relationship are known. Then each firm updates its estimate of the parameters for itself and its competitors, integrating the most up-to-date market information.
We show that these two steps can be performed jointly and formulated as a single problem. They are indeed equivalent to a Mathematical Program with Equilibrium Constraints (MPEC), and also to a mixed integer program.

4.2.1 Dynamic Pricing with Learning in Equilibrium

We focus on the problem under Bertrand competition. At time t, the two steps of the dynamic pricing with learning are as follows:

1. Computation of the Nash equilibrium strategies, assuming the price sensitivities β are known:

$$\mathcal{P}_i(\beta) = \mathcal{P}_i(\mathcal{P}_{-i}(\beta), \beta_{\mathbf{i}}) \ \forall \ i = 1, \ldots, N$$

2. Computation of the price sensitivities, assuming the Nash equilibrium strategies are known:

$$\min_{\beta \in \mathcal{B}} \sum_{i=1}^{N} \left\{ \sum_{h=-H}^{-1} \{\sum_{s=1}^{T} (\widehat{p_i}^h(s) - \mathcal{P}_i(s, \beta_{\mathbf{i}}, \beta_{-\mathbf{i}}))^2\} + \sum_{s=1}^{T} (\widehat{p_i}^0(s) - \mathcal{P}_i(s, \beta_{\mathbf{i}}, \beta_{-\mathbf{i}}))^2 \right\}$$

Using the reformulations of the generalized Nash equilibrium given in Chapter 3, we show that the two steps can be formulated as a single problem in two ways: as an MPEC, or as a mixed integer program.

Formulation as an MPEC

First, we showed in Proposition 3.15, that the generalized Nash equilibrium is equivalent to a quasi variational inequality. Since Step 2 takes the optimal solution of Step 1 as input, then Step 1 is the lower level equilibrium constraint, of an MPEC whose upper level objective is Step 1. Hence, Step 1 and 2 taken together form the following MPEC:

$$\min_{\beta\in\mathcal{B},\mathbf{p}\in\mathfrak{P}(\mathbf{p})} \quad \sum_{i=1}^{N}\left\{\sum_{h=-H}^{-1}\{\sum_{s=1}^{T}(\widehat{p}_i^{\,h}(s)-p_i(s))^2\}+\sum_{s=1}^{T}(\widehat{p}_i^{\,0}(s)-p_i(s))^2\right\} \tag{4.5}$$

$$\text{s.t } \ \forall\ \tilde{\mathbf{p}}\in\mathfrak{P}(\mathbf{p}) \quad \sum_{i=1}^{N}-\nabla_i\pi_i(\mathbf{p_i},\mathbf{p_{-i}},\beta_\mathbf{i})'(\tilde{p}_i-\mathbf{p_i})\geq 0 \tag{4.6}$$

Formulation as a Mixed Integer Program

In Proposition 3.17, we showed that the generalized Nash equilibrium is equivalent to a system of equalities and inequalities. Hence, if we replace the quasi variational inequality of the lower level in the above MPEC, we can reformulate the dynamic pricing with learning in equilibrium as a single level optimization problem.

$$\begin{aligned}
\min_{\beta,\mathbf{p},\lambda,\bar{\mu},\underline{\mu}} \quad & \sum_{i=1}^{N}\left\{\sum_{h=-H}^{-1}\{\sum_{s=1}^{T}(\widehat{p}_i^{\,h}(s)-p_i(s))^2\}+\sum_{s=1}^{T}(\widehat{p}_i^{\,0}(s)-p_i(s))^2\right\}\\
\text{s.t} \quad & -\nabla_i\pi_i(\mathbf{p_i},\mathbf{p_{-i}},\beta_\mathbf{i})+\lambda_i\nabla_i g_i(\mathbf{p_i},\mathbf{p_{-i}},\beta_\mathbf{i})+\sum_{t=1}^{T}(\overline{\mu}_i(t)-\underline{\mu}_i(t))\mathbf{e}(t)=0 \ \ \forall\ i\\
& \lambda_i g_i(\mathbf{p_i},\mathbf{p_{-i}},\beta_\mathbf{i})=0 \ \ \forall\ i\\
& \overline{\mu}_\mathbf{i}{}'(\mathbf{p_i}-\mathbf{p_i}^{max})=0 \ \ \forall\ i\\
& \underline{\mu}_\mathbf{i}{}'\mathbf{p_i}=0 \ \ \forall\ i\\
& \beta_i\in\mathcal{B}_i \ \ \forall\ i\\
& g_i(\mathbf{p_i},\mathbf{p_{-i}},\beta_\mathbf{i})\leq 0 \ \ \forall\ i\\
& \mathbf{0}\leq\mathbf{p_i}\leq\mathbf{p_i^{max}} \ \ \forall\ i
\end{aligned}$$

We can reformulate the complementary slackness constraints by introducing binary variables. This enables us to rewrite the above problem as a mixed integer program:

$$\begin{aligned}
\min \quad & \sum_{i=1}^{N}\left\{\sum_{h=-H}^{-1}\{\sum_{s=1}^{T}(\widehat{p}_i^{\,h}(s)-p_i(s))^2\}+\sum_{s=1}^{T}(\widehat{p}_i^{\,0}(s)-p_i(s))^2\right\}\\
\text{s.t} \quad & -\nabla_i\pi_i(\mathbf{p_i},\mathbf{p_{-i}},\beta_\mathbf{i})+\lambda_i\nabla_i g_i(\mathbf{p_i},\mathbf{p_{-i}},\beta_\mathbf{i})+\sum_{t=1}^{T}(\overline{\mu}_i(t)-\underline{\mu}_i(t))\mathbf{e}(t)=0 \ \ \forall\ i\\
& \lambda_i\leq Mx_i\\
& g_i(\mathbf{p_i},\mathbf{p_{-i}},\beta_\mathbf{i})\leq Mx_i' \ \ \forall\ i\\
& x_i+x_i'\leq 1 \ \ \forall\ i\\
& \bar{\mu}_\mathbf{i}\leq M\bar{\mathbf{y}}_\mathbf{i} \ \ \forall\ i\\
& \mathbf{p_i}\leq M\bar{\mathbf{y}}'_\mathbf{i} \ \ \forall\ i\\
& \bar{\mathbf{y}}_\mathbf{i}+\bar{\mathbf{y}}'_\mathbf{i}\leq\mathbf{e} \ \ \forall\ i\\
& \underline{\mu}_\mathbf{i}\leq M\underline{\mathbf{y}}_\mathbf{i} \ \ \forall\ i\\
& \mathbf{p_i}\leq M\underline{\mathbf{y}'}_\mathbf{i} \ \ \forall\ i
\end{aligned}$$

$$\underline{\mathbf{y}}_{\mathbf{i}} + \underline{\mathbf{y}}'_{\mathbf{i}} \leq \mathbf{e} \ \ \forall\, i$$
$$\beta_i \in \mathcal{B}_i \ \ \forall\, i$$
$$g_i(\mathbf{p_i}, \mathbf{p_{-i}}, \beta_{\mathbf{i}}) \leq 0 \ \ \forall\, i$$
$$\mathbf{0} \leq \mathbf{p_i} \leq \mathbf{p_i^{max}} \ \ \forall\, i$$
$$\lambda_i \geq 0, \underline{\mu}_{\mathbf{i}}, \bar{\mu}_{\mathbf{i}} \geq \mathbf{0} \ \ \forall\, i$$
$$x_i, x'_i \in \{0,1\}, \bar{\mathbf{y}}_{\mathbf{i}}, \bar{\mathbf{y}}'_{\mathbf{i}}, \underline{\mathbf{y}}_{\mathbf{i}}, \underline{\mathbf{y}}'_{\mathbf{i}} \in \{0,1\}^T \ \ \forall\, i$$

4.2.2 Dynamic Allocation with Learning in Equilibrium

For the Cournot market in Equilibrium, the problem at time t has the following two steps:

1. Computation of the Nash equilibrium strategies, assuming the allocation sensitivities α are known:

$$\mathcal{Q}_i(\alpha)\mathcal{Q}_i(\mathcal{Q}_{-i}(\alpha)) \ \forall\ i = 1, \ldots, N$$

2. Computation of the price sensitivities, assuming the Nash equilibrium strategies are known:

$$\min_{\alpha \in \mathcal{A}} \sum_{i=1}^{N} \left\{ \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} (\widehat{q}_i^{\,h}(s) - \mathcal{Q}_i(s, \alpha_{\mathbf{i}}, \alpha_{-\mathbf{i}}))^2 \} + \sum_{s=1}^{T} (\widehat{q}_i^{\,0}(s) - \mathcal{Q}_i(s, \alpha_{\mathbf{i}}, \alpha_{-\mathbf{i}}))^2 \right\}$$

We show using the reformulations of the Nash equilibrium given in Chapter 3 that the two steps can be formulated as a single problem in two ways: as an MPEC, or as a mixed integer program.

Formulation as an MPEC

First, we showed in Proposition 3.20, that the Cournot Nash equilibrium is equivalent to a variational inequality. Since Step 2 takes the optimal solution of Step 1 as input, then Step 1 is the lower level equilibrium constraint of an MPEC whose upper level objective is Step 1. Hence, Step 1 and 2 taken together form the following MPEC:

$$\begin{array}{ll} \min\limits_{\alpha \in \mathcal{A}, \mathbf{q} \in \mathfrak{Q}} & \left\{ \sum_{h=-H}^{-1} \{ \sum_{s=1}^{T} (\widehat{q}_i^{\,h}(s) - \mathcal{Q}_i(s, \alpha_{\mathbf{i}}, \alpha_{-\mathbf{i}}))^2 \} + \sum_{s=1}^{T} (\widehat{q}_i^{\,0}(s) - \mathcal{Q}_i(s, \alpha_{\mathbf{i}}, \alpha_{-\mathbf{i}}))^2 \right\} \\ \text{s.t } \forall\ \tilde{\mathbf{q}} \in \mathfrak{Q} & \sum_{i=1}^{N} -\nabla_i \pi_i(\mathcal{Q}_i(\alpha), \mathcal{Q}_{-i}(\alpha), \alpha_{\mathbf{i}})'(\mathbf{q_i} - \mathcal{Q}_i(\alpha)) \geq 0 \end{array}$$

Formulation as a Mixed Integer Program

By replacing the lower level Nash equilibrium by its equivalent system of equalities and inequalities derived in Proposition 3.21, we can transform the above MPEC into a single-level optimization problem:

$$\begin{array}{cc}
\min\limits_{\alpha,\mathbf{q},\lambda,\mu} & \left\{\sum_{h=-H}^{-1}\{\sum_{s=1}^{T}(\widehat{q_i}^h(s)-\mathcal{Q}_i(s,\alpha_{\mathbf{i}},\alpha_{-\mathbf{i}}))^2\}+\sum_{s=1}^{T}(\widehat{q_i}^0(s)-\mathcal{Q}_i(s,\alpha_{\mathbf{i}},\alpha_{-\mathbf{i}}))^2\right\} \\
\text{s.t} & -\nabla_i\pi_i(\mathbf{q_i},\mathbf{q_{-i}},\alpha_{\mathbf{i}})+\lambda_i\nabla_i g_i(\mathbf{q_i})-\sum_{t=1}^{T}\mu_i(t)\mathbf{e}(t)=0 \ \ \forall\ i \\
& \lambda_i g_i(\mathbf{q_i})=0 \ \ \forall\ i \\
& \mu_{\mathbf{i}}{}'\mathbf{q_i}=0 \ \ \forall\ i \\
& g_i(\mathbf{q_i})\le 0 \ \ \forall\ i \\
& \mathbf{0}\le\mathbf{q_i} \ \ \forall\ i \\
& \lambda_i\ge 0,\mu_{\mathbf{i}}\ge\mathbf{0} \ \ \forall\ i
\end{array}$$

We can reformulate the disjunctive complementary slackness constraints using binary variables. Thus the MPEC is equivalent to a mixed integer program:

$$\begin{array}{cc}
\min\limits_{\alpha,\mathbf{q},\lambda,\mu} & \left\{\sum_{h=-H}^{-1}\{\sum_{s=1}^{T}(\widehat{q_i}^h(s)-\mathcal{Q}_i(s,\alpha_{\mathbf{i}},\alpha_{-\mathbf{i}}))^2\}+\sum_{s=1}^{T}(\widehat{q_i}^0(s)-\mathcal{Q}_i(s,\alpha_{\mathbf{i}},\alpha_{-\mathbf{i}}))^2\right\} \\
\text{s.t} & -\nabla_i\pi_i(\mathbf{q_i},\mathbf{q_{-i}},\alpha_{\mathbf{i}})+\lambda_i\nabla_i g_i(\mathbf{q_i})-\sum_{t=1}^{T}\mu_i(t)\mathbf{e}(t)=0 \ \ \forall\ i \\
& \lambda_i\le Mx_i \ \ \forall\ i \\
& g_i(\mathbf{q_i})\le Mx_i' \ \ \forall\ i \\
& x_i+x_i'\le 1 \ \ \forall\ i \\
& \mu_{\mathbf{i}}\le M\mathbf{y_i} \ \ \forall\ i \\
& \mathbf{q_i}\le M\mathbf{y_i'} \ \ \forall\ i \\
& \mathbf{y_i}+\mathbf{y_i'}\le\mathbf{e} \ \ \forall\ i \\
& g_i(\mathbf{q_i})\le 0 \ \ \forall\ i \\
& \mathbf{0}\le\mathbf{q_i} \ \ \forall\ i \\
& \lambda_i\ge 0,\mu_{\mathbf{i}}\ge\mathbf{0} \ \ \forall\ i \\
& x_i,x_i'\in\{0,1\},\mathbf{y_i},\mathbf{y_i'}\in\{0,1\}^T \ \ \forall\ i
\end{array}$$

Chapter 5

Key Learning Result

In this chapter, we prove that the dynamic pricing (resp. allocation) problem with learning in disequilibrium converges. By that, we mean that the two types of learning are achieved in the long run: learning of the price-demand relationship and learning of the equilibrium strategies. The result holds when all firms in the market apply the dynamic pricing (resp. allocation) with learning procedure. Furthermore, the result holds for a variety of estimation schemes such that the optimal vector of price-demand parameters does not vary too much from one period to the next.
In Section 1, we address the dynamic pricing with learning problem. We introduce the assumptions under which the learning result holds. We then state the result, and proceed to prove it. Section 2 focuses on the dynamic allocation with learning, and contains the assumptions, the result and its proof.

5.1 Learning under Bertrand Competition

In this section, we review the learning approach under Bertrand competition. We then study the convergence of the approach First of all, we introduce the assumptions under which the dynamic pricing with learning approach converges. The result relies on the reformulation of the Bertrand best-response problem of Proposition 3.8. Indeed, the Bertrand best-response problem is hard to analyse due to the fact that the strategy set of the best-response problem of one firm depends both on its competitors' strategy and on its price sensitivities. We then state and prove the learning result.

5.1.1 The Learning Approach

The learning approach is an iterative process whereby firms update their best-response strategy and the estimate regarding their price sensitivities at each period. All the firms in the market are assumed to use the learning approach.
The approach can be described by the following algorithm: at each period m, given competitors' vector of prices for the previous period $\mathbf{p_{-i}}^{(m-1)}$, each firm computes its best-response price as a function of price sensitivities, i.e $\mathcal{P}_i(\mathbf{p_{-i}}^{(m-1)}, \beta_\mathbf{i})$. Each firm then uses the best-response price function in the estimation problem, in order

to update its estimate of the price sensitivities. The new price sensitivities at m are denoted $\beta_{\mathbf{i}}^{(m)}$. Hence, firm i's vector of best-response prices at period m is $\mathbf{p_i}^{(m)} = \mathcal{P}_i(\mathbf{p_{-i}}^{(m-1)}, \beta_{\mathbf{i}}^{(m)})$.
Learning of the equilibrium strategies is said to be achieved when the difference between the vectors of best-response prices at period m and at period $m-1$ do not differ by more than a small constant $\eta > 0$, in other words:

$$\|\mathbf{p}^{(m-1)} - \mathbf{p}^{(m)}\| \leq \eta$$

Constant η is the precision of learning. Similarly, learning of the price sensitivities is achieved in precision η if:

$$\|\beta^{(m-1)} - \beta^{(m)}\| \leq \eta$$

Thus the learning approach under Bertrand competition can be described by the below algorithm.

Algorithm 1: Learning approach under Bertrand competition
Input: A set of initial values for the prices and price sensitivities
Output: The equilibrium prices and optimal price sensitivities

for $i = 1$ **to** N
  Initialize $\mathbf{p_i} \leftarrow \mathbf{p_i}^{(0)}$
  Initialize $\beta_{\mathbf{i}} \leftarrow \beta_{\mathbf{i}}^{(0)}$
  Initialize $m = 1$
  repeat
    for $i = 1$ **to** N
      Compute $\mathcal{P}_i(\mathbf{p_{-i}}^{(m-1)}, \beta_{\mathbf{i}})$
      Compute $\beta_{\mathbf{i}}^{(m)}$
      Set $\mathbf{p_i}^{(m)} \leftarrow \mathcal{P}_i(\mathbf{p_{-i}}^{(m-1)}, \beta_{\mathbf{i}}^{(m)})$
      $m \leftarrow m + 1$
    until $\|\mathbf{p}^{(m-1)} - \mathbf{p}^{(m)}\| \leq \eta$ **and** $\|\beta^{(m-1)} - \beta^{(m)}\| \leq \eta$

Note that in this chapter, we do not explain how the best-response prices and estimated parameters are computed practically. Their computation will be discussed in depth in Chapter 6.

5.1.2 Assumptions

We first restate the assumption needed to transform the Bertrand best-response problem into a problem with a polyhedral feasible set that does not depend on parameters and competitors' strategy.

Assumption 3.13. *The demand function $q_i(., \overline{q_{-i}}(t), \beta_i(t))$ is invertible w.r.t $\mathbf{p_i}$, and we denote by $\breve{p}_i(., \overline{q_{-i}}(t), \beta_i(t))$ its inverse. In other words:*

$$q_i(t) = q_i(p_i(t), \overline{q_{-i}}(t), \beta_i(t)) \Leftrightarrow p_i(t) = \breve{p}_i(q_i(t), \overline{q_{-i}}(t), \beta_i(t))$$

Let $\breve{\pi}_i(\mathbf{q_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) = \sum_{t=1}^{T} q_i(t)\breve{p}_i(q_i(t), \overline{p_{-i}}(t), \beta_i(t))$.

Proposition 3.8. *The Bertrand best-response problem can be reformulated as follows:*

$$\max_{\mathbf{q_i}} \quad \breve{\pi}_i(\mathbf{q_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}) \tag{5.1}$$

$$s.t \quad \sum_{t=1}^{T} q_i(t) \leq C_i \tag{5.2}$$

$$0 \leq q_i(t) \tag{5.3}$$

The following assumptions are made on the modified price and revenue function:

Assumptions 5.1. *1. $\nabla\breve{\pi}_i$ is Lipschitz continuous w.r.t $\mathbf{p_{-i}}$ with constant $\mathcal{L}_{\pi_i,p_{-i}}(\beta_\mathbf{i})$*

2. $\nabla\breve{\pi}_i$ is Lipschitz continuous w.r.t $\beta_\mathbf{i}$ with constant $\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q_i}, \mathbf{p_{-i}})$

3. $-\nabla\breve{\pi}_i$ is strongly monotone w.r.t $\mathbf{q_i}$ with constant $\mathcal{M}_{\pi_i,q_i}(\beta_\mathbf{i})$

4. $\breve{p}_i$ is Lipschitz continuous w.r.t $\mathbf{q_i}$ with constant $\mathcal{L}_{\breve{p}_i,q_i}(\beta_\mathbf{i})$;

5. $\breve{p}_i$ is Lipschitz continuous w.r.t $\mathbf{p_{-i}}$ with constant $\mathcal{L}_{\breve{p}_i,p_{-i}}(\beta_\mathbf{i})$

6. $\breve{p}_i$ is Lipschitz continuous w.r.t $\beta_\mathbf{i}$ with constant $\mathcal{L}_{\breve{p}_i,\beta_i}(\mathbf{q_i}, \mathbf{p_{-i}})$

In other words, we assume the following behavior concerning the modified price and revenue functions:

1. For all $\mathbf{q_{-i}} \in \mathfrak{Q}_i$, for all $\hat{\mathbf{p}}_{-\mathbf{i}}, \check{\mathbf{p}}_{-\mathbf{i}} \in [0, \mathbf{p_{-i}}^{max}]$, for all $\beta_\mathbf{i} \in \mathcal{B}_i$:

$$\|\nabla\breve{\pi}_i(\mathbf{q_i}, \hat{\mathbf{p}}_{-\mathbf{i}}, \beta_\mathbf{i}) - \nabla\breve{\pi}_i(\mathbf{q_i}, \check{\mathbf{p}}_{-\mathbf{i}}, \beta_\mathbf{i})\| \leq \mathcal{L}_{\pi_i,p_{-i}}(\beta_\mathbf{i})\|\hat{\mathbf{p}}_{-\mathbf{i}} - \check{\mathbf{p}}_{-\mathbf{i}}\|$$

2. For all $\mathbf{q_{-i}} \in \mathfrak{Q}_i$, for all $\mathbf{p_{-i}} \in [0, \mathbf{p_{-i}}^{max}]$, for all $\hat{\beta}_\mathbf{i}, \check{\beta}_\mathbf{i} \in \mathcal{B}_i$:

$$\|\nabla\breve{\pi}_i(\mathbf{q_i}, \mathbf{p_{-i}}, \hat{\beta}_\mathbf{i}) - \nabla\breve{\pi}_i(\mathbf{q_i}, \mathbf{p_{-i}}, \check{\beta}_\mathbf{i})\| \leq \mathcal{L}_{\pi_i,\beta_i}(\mathbf{q_i}, \mathbf{p_{-i}})\|\hat{\beta}_\mathbf{i} - \check{\beta}_\mathbf{i}\|$$

3. For all $\hat{\mathbf{q}}_{-\mathbf{i}}, \check{\mathbf{q}}_{-\mathbf{i}} \in \mathfrak{Q}_i$, for all $\mathbf{p_{-i}} \in [0, \mathbf{p_{-i}}^{max}]$, for all $\beta_\mathbf{i} \in \mathcal{B}_i$:

$$(-\nabla\breve{\pi}_i(\hat{\mathbf{q}}_\mathbf{i}, \mathbf{p_{-i}}, \beta_\mathbf{i}) + \nabla\breve{\pi}_i(\check{\mathbf{q}}_\mathbf{i}, \mathbf{p_{-i}}, \beta_\mathbf{i}))'(\hat{\mathbf{q}}_\mathbf{i} - \check{\mathbf{q}}_\mathbf{i}) \geq \mathcal{M}_{\pi_i,q_i}(\beta_\mathbf{i})\|\hat{\mathbf{q}}_\mathbf{i} - \check{\mathbf{q}}_\mathbf{i}\|^2$$

4. For all $\hat{\mathbf{q}}_{-\mathbf{i}}, \check{\mathbf{q}}_{-\mathbf{i}} \in \mathfrak{Q}_i$, for all $\mathbf{p_{-i}} \in [0, \mathbf{p_{-i}}^{max}]$, for all $\beta_\mathbf{i} \in \mathcal{B}_i$:

$$\|\breve{p}_i(\hat{\mathbf{q}}_\mathbf{i}, \mathbf{p_{-i}}, \beta_\mathbf{i} - \breve{p}_i(\check{\mathbf{q}}_\mathbf{i}, \mathbf{p_{-i}}, \beta_\mathbf{i}\| \leq \mathcal{L}_{\breve{p}_i,q_i}(\beta_\mathbf{i})\|\hat{\mathbf{q}}_\mathbf{i} - \check{\mathbf{q}}_\mathbf{i}\|$$

5. For all $\mathbf{q_{-i}} \in \mathfrak{Q}_i$, for all $\hat{\mathbf{p}}_{-\mathbf{i}}, \check{\mathbf{p}}_{-\mathbf{i}} \in [0, \mathbf{p_{-i}}^{max}]$, for all $\beta_\mathbf{i} \in \mathcal{B}_i$:

$$\|\breve{p}_i(\mathbf{q_i}, \hat{\mathbf{p}}_{-\mathbf{i}}, \beta_\mathbf{i}) - \breve{p}_i(\mathbf{q_i}, \check{\mathbf{p}}_{-\mathbf{i}}, \beta_\mathbf{i})\| \leq \mathcal{L}_{\breve{p}_i,p_{-i}}(\beta_\mathbf{i})\|\hat{\mathbf{p}}_{-\mathbf{i}} - \check{\mathbf{p}}_{-\mathbf{i}}\|$$

6. For all $\mathbf{q_{-i}} \in \mathfrak{Q}_i$, for all $\mathbf{p_{-i}} \in [0, \mathbf{p_{-i}}^{max}]$, for all $\hat{\beta}_\mathbf{i}, \check{\beta}_\mathbf{i} \in \mathcal{B}_i$:

$$\|\breve{p}_i(\mathbf{q_i}, \mathbf{p_{-i}}, \hat{\beta}_\mathbf{i} - \breve{p}_i(\mathbf{q_i}, \mathbf{p_{-i}}, \check{\beta}_\mathbf{i}\| \leq \mathcal{L}_{\breve{p}_i,\beta_i}(\mathbf{q_i}, \mathbf{p_{-i}})\|\hat{\beta}_\mathbf{i} - \check{\beta}_\mathbf{i}\|$$

Lipschitz continuity intuitively means that the function behaves smoothly. In particular, Lipschitz continuity implies differentiability. Furthermore, it implies that the derivative is bounded by a constant. Hence, $\nabla_i \breve{\pi}_i$ has bounded variations in $\mathbf{p_{-i}}$, and $\breve{p}_i$ has bounded variations in its three variables.
Strong monotonicity is the generalization of the strictly increasing property to a function in a topological vector space.
Before we introduce the second set of assumptions, let us define the following constants:

$$\begin{aligned}
\mathcal{K}_i(\beta_{\mathbf{i}}) &= \left(\frac{\mathcal{L}_{\pi_i,p_{-i}}(\beta_{\mathbf{i}})\mathcal{L}_{\breve{p}_i,q_i}(\beta_{\mathbf{i}})}{\mathcal{M}_{\pi_i,q_i}(\beta_{\mathbf{i}})} + \mathcal{L}_{\breve{p}_i,p_{-i}}(\beta_{\mathbf{i}})\right) \\
\mathcal{K}'_i(\mathbf{q_i}, \mathbf{p_{-i}}, \beta_{\mathbf{i}}, \beta_{\mathbf{i}}') &= \left(\mathcal{L}_{\breve{p}_i,\beta_i}(\mathbf{q_i}, \mathbf{p_{-i}}) + \frac{\mathcal{L}_{\breve{p}_i,q_i}(\beta_{\mathbf{i}}')\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q_i}, \mathbf{p_{-i}})}{\mathcal{M}_{\pi_i,q_i}(\beta_{\mathbf{i}})}\right)
\end{aligned}$$

Assumptions 5.2.

- *At each period m of the approach, we define $\mathcal{B}_i^{(m)} = \{\beta_{\mathbf{i}} \in \mathcal{B}_i \; : \; \|\beta_{\mathbf{i}} - \beta_{\mathbf{i}}{}^{(m-1)}\| \leq \mathcal{K}''_i \|\mathbf{p_{-i}}{}^{(m-1)} - \mathbf{p_{-i}}{}^{(m-2)}\|, \;\; \|\beta_{\mathbf{i}}\| \leq M\}$. The estimation step is performed on the reduced feasible set $\mathcal{B}_i^{(m)}$.*

- *We define the following constants:*

$$\begin{aligned}
\mathcal{K}_i = \max_{\beta_{\mathbf{i}}} \quad & \mathcal{K}(\beta_{\mathbf{i}}) \\
s.t \quad & \|\beta_{\mathbf{i}}\| \leq M \\
\mathcal{K}'_i = \max_{\mathbf{q_i}, \mathbf{p_{-i}}, \beta_{\mathbf{i}}, \beta_{\mathbf{i}}'} \quad & \mathcal{K}'(\mathbf{q_i}, \mathbf{p_{-i}}, \beta_{\mathbf{i}}, \beta_{\mathbf{i}}') \\
s.t \quad & \mathbf{0} \leq \mathbf{q_i} \leq C_i \mathbf{e} \\
& \mathbf{0} \leq \mathbf{p_{-i}} \leq \mathbf{p_{-i}}^{max} \\
& \|\beta_{\mathbf{i}}\|, \|\beta_{\mathbf{i}}'\| \leq M
\end{aligned}$$

  There exists a constant $\eta > 0$ such that:

$$\mathcal{K}_i + \mathcal{K}'_i \mathcal{K}''_i \leq \frac{1-\eta}{N-1}$$

The first bullet of Assumption 5.2 basically reduces the choice of bounded parameters to a set of parameters which shrinks at least as fast as the best-response prices shrink. The second bullet of Assumption 5.2 ensures that the mapping corresponding to the algorithm is a contraction.

5.1.3 Statement and Proof

Theorem 5.1. *Let $(\beta^{(0)}, \mathbf{p}^{(0)})$ be some starting values. under Assumptions 5.1 and 5.2, the sequence of iterates $\{\beta^{(m)}, \mathbf{p}^{(m)}\}$ generated by the approach converges as m goes to infinity to $(\beta^*, \mathbf{p}^*)$ such that $\mathbf{p}^*$ is the set of Nash equilibrium policies corresponding to parameters β^*.*

In what follows, we denote by $\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})$ the best-response quantity, when competitors' prices are $\mathbf{p_{-i}}^{(m-1)}$, and $\mathbf{q_i}^{(m)} = \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}^{(m)})$. The corresponding prices are $\mathcal{P}_i^{(m)}(\beta_{\mathbf{i}})$ and $\mathbf{p_i}^{(m)}$ respectively.
The proof relies on a contraction argument. We indeed prove that under the Lipschitz continuity, strong monotonicity assumptions, and when the estimation is performed on the reduced set of parameters, the mapping corresponding to the approach is a contraction. Hence, the sequence of prices and parameters converge.

Proof. We use the variational inequality formulation of the best-response problem $\mathcal{BR}_i(\mathbf{p_{-i}}^{(m-1)}, \beta_{\mathbf{i}})$ applied at $\mathbf{q_i} = \mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}})$ and that of $\mathcal{BR}_i(\mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}})$ applied at $\mathbf{q}_i = \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})$:

$$-\nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m-1)}, \beta_{\mathbf{i}})'(\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})) \geq 0 \tag{5.4}$$

$$-\nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}})'(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}})) \geq 0 \tag{5.5}$$

We sum both variational inequalities and get:

$$\begin{aligned}\Big(-\nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m-1)}, \beta_{\mathbf{i}}) + \nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}})\Big)'& \\ (\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})) \geq 0&\end{aligned} \tag{5.6}$$

We add and subtract $\breve{\pi}_i(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}})$ in the left hand side of the inequality, and rearrange terms to obtain:

$$\begin{aligned}\Big(-\nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m-1)}, \beta_{\mathbf{i}}) + \nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}})\Big)' (\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})) \geq \\ \Big(-\nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}}) + \nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}})\Big)' (\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}))\end{aligned} \tag{5.7}$$

By strong monotonicity of $-\nabla_i \breve{\pi}_i$ w.r.t $\mathbf{q_i}$, the right hand side of the inequality is bounded below by:

$$\mathcal{M}_{\pi_i, q_i}(\beta_{\mathbf{i}}) \|\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})\|^2$$

Hence, it is nonnegative, and the left hand side of the inequality is equal to its absolute value. Using Cauchy-Schwartz inequality, it is bounded above by:

$$\| - \nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m-1)}, \beta_{\mathbf{i}}) + \nabla_i \breve{\pi}_i(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}})\|.\|\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})\|$$

By Lipschitz continuity of $-\nabla \breve{\pi}_i$ w.r.t $\mathbf{p_{-i}}$, the above term is itself bounded above by:

$$\mathcal{L}_{\pi_i, p_{-i}}(\beta_{\mathbf{i}}) \|\mathbf{p_{-i}}^{(m)} - \mathbf{p_{-i}}^{(m-1)}\|.\|\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})\|$$

Putting things together yields the following inequality:

$$\mathcal{M}_{\pi_i, q_i}(\beta_{\mathbf{i}}) \|\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})\| \leq \mathcal{L}_{\pi_i, p_{-i}}(\beta_{\mathbf{i}}) \|\mathbf{p_{-i}}^{(m)} - \mathbf{p_{-i}}^{(m-1)}\| \tag{5.8}$$

The next step is to establish an inequality between $\|\mathcal{P}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{P}_i^{(m)}(\beta_{\mathbf{i}})\|$ and $\|\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})\|$.

$$\begin{aligned}&\|\mathcal{P}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{P}_i^{(m)}(\beta_{\mathbf{i}})\| \\ &\quad \leq \|\breve{\mathbf{p}}_{\mathbf{i}}(\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}}) - \breve{\mathbf{p}}_{\mathbf{i}}(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m-1)}, \beta_{\mathbf{i}})\| \\ &\quad \leq \|\breve{\mathbf{p}}_{\mathbf{i}}(\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}}) - \breve{\mathbf{p}}_{\mathbf{i}}(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p_{-i}}^{(m)}, \beta_{\mathbf{i}})\|\end{aligned}$$

$$\begin{aligned}
&+ \quad \|\breve{\mathbf{p}}_{\mathbf{i}}(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p}_{-\mathbf{i}}{}^{(m)}, \beta_{\mathbf{i}}) - \breve{\mathbf{p}}_{\mathbf{i}}(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}})\| \\
&\leq \quad \mathcal{L}_{\breve{p}_i,q_i}(\beta_{\mathbf{i}})\|\mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})\| + \mathcal{L}_{\breve{p}_i,p_{-i}}(\beta_{\mathbf{i}})\|\mathbf{p}_{-\mathbf{i}}{}^{(m)} - \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}\|
\end{aligned}$$

The last inequality follows from Lipschitz continuity of $\breve{\mathbf{p}}_{\mathbf{i}}$ w.r.t $\mathbf{q}_{\mathbf{i}}$ and w.r.t $\mathbf{p}_{-\mathbf{i}}$. Using the inequality we established 5.8, we obtain:

$$\|\mathcal{P}_i^{(m+1)}(\beta_{\mathbf{i}}) - \mathcal{P}_i^{(m)}(\beta_{\mathbf{i}})\| \leq \left(\frac{\mathcal{L}_{\pi_i,p_{-i}}(\beta_{\mathbf{i}})\mathcal{L}_{\breve{p}_i,q_i}(\beta_{\mathbf{i}})}{\mathcal{M}_{\pi_i,q_i}(\beta_{\mathbf{i}})} + \mathcal{L}_{\breve{p}_i,p_{-i}}(\beta_{\mathbf{i}})\right)\|\mathbf{p}_{-\mathbf{i}}{}^{(m)} - \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}\| \quad (5.9)$$

Since $\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})$ is the solution to best-response problem $\mathcal{BR}_i(\mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}})$, and $\mathbf{q}_{\mathbf{i}}{}^{(m)}$ to best-response problem $\mathcal{BR}_i(\mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}}{}^{(m)})$, we have:

$$-\nabla_i\breve{\pi}_i(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}})'(\mathbf{q}_{\mathbf{i}}{}^{(m)} - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})) \geq 0$$

$$-\nabla_i\breve{\pi}_i(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}}{}^{(m)})'(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}) - \mathbf{q}_{\mathbf{i}}{}^{(m)}) \geq 0$$

Summing both inequalities, adding and subtracting $\nabla_i\breve{\pi}_i(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}})$, and rearranging terms, we get:

$$\begin{aligned}
&\left(-\nabla_i\breve{\pi}_i(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}}{}^{(m)}) + \nabla_i\breve{\pi}_i(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}})\right)'(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}) - \mathbf{q}_{\mathbf{i}}{}^{(m)}) \geq \\
&\quad \left(-\nabla_i\breve{\pi}_i(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}}) + \nabla_i\breve{\pi}_i(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}}{}^{(m)})\right)'(d\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}) - \mathbf{q}_{\mathbf{i}}{}^{(m)})
\end{aligned}$$

Using Lipschitz continuity of $-\nabla_i\breve{\pi}_i$ w.r.t $\beta_{\mathbf{i}}$, and strong monotonicity w.r.t $\mathbf{q}_{\mathbf{i}}$, as was done for Inequality 5.8, we get:

$$\mathcal{M}_{\pi_i,q_i}(\beta_{\mathbf{i}}{}^{(m)})\|\mathbf{q}_{\mathbf{i}}{}^{(m)} - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})\| \leq \mathcal{L}_{\pi_i,\beta_{\mathbf{i}}}(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)})\|\beta_{\mathbf{i}} - \beta_{\mathbf{i}}{}^{(m)}\|$$

As a result, we have:

$$\|\mathbf{q}_{\mathbf{i}}{}^{(m)} - \mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}})\| \leq \frac{\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)})}{\mathcal{M}_{\pi_i,q_i}(\beta_{\mathbf{i}}{}^{(m)})}\|\beta_{\mathbf{i}} - \beta_{\mathbf{i}}{}^{(m)}\| \quad (5.10)$$

Similarly, we have:

$$\|\mathbf{q}_{\mathbf{i}}{}^{(m+1)} - \mathcal{Q}_i^{(m+1)}(\beta_{\mathbf{i}})\| \leq \frac{\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q}_{\mathbf{i}}{}^{(m+1)}, \mathbf{p}_{-\mathbf{i}}{}^{(m)})}{\mathcal{M}_{\pi_i,q_i}(\beta_{\mathbf{i}}{}^{(m+1)})}\|\beta_{\mathbf{i}} - \beta_{\mathbf{i}}{}^{(m+1)}\| \quad (5.11)$$

We use the above two inequalities towards deriving two inequalities, between $\|\mathbf{p}_{\mathbf{i}}{}^{(m)} - \mathcal{P}_i^{(m)}(\beta_{\mathbf{i}})\|$ and $\|\beta_{\mathbf{i}} - \beta_{\mathbf{i}}{}^{(m)}\|$, and between $\|\mathbf{p}_{\mathbf{i}}{}^{(m+1)} - \mathcal{P}_i^{(m+1)}(\beta_{\mathbf{i}})\|$ and $\|\beta_{\mathbf{i}} - \beta_{\mathbf{i}}{}^{(m+1)}\|$:

$$\begin{aligned}
&\|\mathbf{p}_{\mathbf{i}}{}^{(m)} - \mathcal{P}_i^{(m)}(\beta_{\mathbf{i}})\| \\
&\quad = \quad \|\breve{\mathbf{p}}_{\mathbf{i}}(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}}{}^{(m)}) - \breve{\mathbf{p}}_{\mathbf{i}}(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}})\| \\
&\quad \leq \quad \|\breve{\mathbf{p}}_{\mathbf{i}}(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}}{}^{(m)}) - \breve{\mathbf{p}}_{\mathbf{i}}(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}})\| \\
&\quad + \quad \|\breve{\mathbf{p}}_{\mathbf{i}}(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}}) - \breve{\mathbf{p}}_{\mathbf{i}}(\mathcal{Q}_i^{(m)}(\beta_{\mathbf{i}}), \mathbf{p}_{-\mathbf{i}}{}^{(m-1)}, \beta_{\mathbf{i}})\| \\
&\quad \leq \quad \mathcal{L}_{\breve{p}_i,\beta_i}(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)})\|\beta_{\mathbf{i}}{}^{(m)} - \beta_{\mathbf{i}}\| + \mathcal{L}_{\breve{p}_i,q_i}(\beta_{\mathbf{i}})\|\mathcal{Q}_i(\beta_{\mathbf{i}}) - \mathbf{q}_{\mathbf{i}}{}^{(m)}\|
\end{aligned}$$

The third inequality comes from Lipschitz continuity of $\breve{\mathbf{p}}_{\mathbf{i}}$ w.r.t β_i and $\mathbf{q}_{\mathbf{i}}$. Now, using Inequality 5.10 above, the second term in the last inequality is bounded above by :

$$\frac{\mathcal{L}_{\breve{p}_i,q_i}(\beta_{\mathbf{i}})\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q}_{\mathbf{i}}{}^{(m)}, \mathbf{p}_{-\mathbf{i}}{}^{(m-1)})}{\mathcal{M}_{\pi_i,q_i}(\beta_{\mathbf{i}}{}^{(m)})}\|\beta_{\mathbf{i}}{}^{(m)} - \beta_{\mathbf{i}}\|$$

Putting things together, we get the following inequality:

$$
\begin{aligned}
&\|\mathbf{p_i}^{(m)} - \mathcal{P}_i^{(m)}(\beta_\mathbf{i})\| && (5.12)\\
\leq\ & \left(\mathcal{L}_{\breve{p}_i,\beta_i}(\mathbf{q_i}^{(m)}, \mathbf{p_{-i}}^{(m-1)}) + \frac{\mathcal{L}_{\breve{p}_i,q_i}(\beta_\mathbf{i})\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q_i}^{(m)}, \mathbf{p_{-i}}^{(m-1)})}{\mathcal{M}_{\pi_i,q_i}(\beta_\mathbf{i}^{(m)})}\right) && (5.13)\\
.\ & \|\beta_\mathbf{i} - \beta_\mathbf{i}^{(m)}\| && (5.14)
\end{aligned}
$$

Similarly, using Inequality 5.11, we get the following inequality:

$$
\begin{aligned}
&\|\mathbf{p_i}^{(m+1)} - \mathcal{P}_i^{(m+1)}(\beta_\mathbf{i})\| && (5.15)\\
\leq\ & \left(\mathcal{L}_{\breve{p}_i,\beta_i}(\mathbf{q_i}^{(m+1)}, \mathbf{p_{-i}}^{(m)}) + \frac{\mathcal{L}_{\breve{p}_i,q_i}(\beta_\mathbf{i})\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q_i}^{(m+1)}, \mathbf{p_{-i}}^{(m)})}{\mathcal{M}_{\pi_i,q_i}(\beta_\mathbf{i}^{(m+1)})}\right) && (5.16)\\
.\ & \|\beta_\mathbf{i} - \beta_\mathbf{i}^{(m+1)}\| && (5.17)
\end{aligned}
$$

Using the above inequalities, we establish:

$$
\begin{aligned}
&\|\mathbf{p_i}^{(m+1)} - \mathbf{p_i}^{(m)}\|\\
\leq\ & \|\mathbf{p_i}^{(m+1)} - \mathcal{P}_i^{(m+1)}(\beta_\mathbf{i})\| + \|\mathcal{P}_i^{(m+1)}(\beta_\mathbf{i}) - \mathcal{P}_i^{(m)}(\beta_\mathbf{i})\| + \|\mathbf{p_i}^{(m)} - \mathcal{P}_i^{(m)}(\beta_\mathbf{i})\|
\end{aligned}
$$

The first term in the right hand side of the inequality can be bounded above using Inequality 5.13; the third term can be bounded above using Inequality 5.12; the second term can be bounded above using Inequality 5.9.

$$
\begin{aligned}
&\|\mathbf{p_i}^{(m+1)} - \mathbf{p_i}^{(m)}\|\\
\leq\ & \left(\mathcal{L}_{\breve{p}_i,\beta_i}(\mathbf{q_i}^{(m+1)}, \mathbf{p_{-i}}^{(m)}) + \frac{\mathcal{L}_{\breve{p}_i,q_i}(\beta_\mathbf{i})\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q_i}^{(m+1)}, \mathbf{p_{-i}}^{(m)})}{\mathcal{M}_{\pi_i,q_i}(\beta_\mathbf{i}^{(m+1)})}\right) \|\beta_\mathbf{i} - \beta_\mathbf{i}^{(m+1)}\|\\
+\ & \left(\frac{\mathcal{L}_{\pi_i,p_{-i}}(\beta_\mathbf{i})\mathcal{L}_{\breve{p}_i,q_i}(\beta_\mathbf{i})}{\mathcal{M}_{\pi_i,q_i}(\beta_\mathbf{i})} + \mathcal{L}_{\breve{p}_i,p_{-i}}(\beta_\mathbf{i})\right) \|\mathbf{p_{-i}}^{(m)} - \mathbf{p_{-i}}^{(m-1)}\|\\
+\ & \left(\mathcal{L}_{\breve{p}_i,\beta_i}(\mathbf{q_i}^{(m)}, \mathbf{p_{-i}}^{(m-1)}) + \frac{\mathcal{L}_{\breve{p}_i,q_i}(\beta_\mathbf{i})\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q_i}^{(m)}, \mathbf{p_{-i}}^{(m-1)})}{\mathcal{M}_{\pi_i,q_i}(\beta_\mathbf{i}^{(m)})}\right) \|\beta_\mathbf{i} - \beta_\mathbf{i}^{(m)}\|
\end{aligned}
$$

Rewriting the inequality for $\beta_\mathbf{i} = \beta_\mathbf{i}^{(m+1)}$, we get:

$$
\begin{aligned}
&\|\mathbf{p_i}^{(m+1)} - \mathbf{p_i}^{(m)}\|\\
\leq\ & \left(\frac{\mathcal{L}_{\pi_i,p_{-i}}(\beta_\mathbf{i}^{(m+1)})\mathcal{L}_{\breve{p}_i,q_i}(\beta_\mathbf{i}^{(m+1)})}{\mathcal{M}_{\pi_i,q_i}(\beta_\mathbf{i}^{(m+1)})} + \mathcal{L}_{\breve{p}_i,p_{-i}}(\beta_\mathbf{i}^{(m+1)})\right) \|\mathbf{p_{-i}}^{(m)} - \mathbf{p_{-i}}^{(m-1)}\|\\
+\ & \left(\mathcal{L}_{\breve{p}_i,\beta_i}(\mathbf{q_i}^{(m)}, \mathbf{p_{-i}}^{(m-1)}) + \frac{\mathcal{L}_{\breve{p}_i,q_i}(\beta_\mathbf{i}^{(m+1)})\mathcal{L}_{\pi_i,\beta_i}(\mathbf{q_i}^{(m)}, \mathbf{p_{-i}}^{(m-1)})}{\mathcal{M}_{\pi_i,q_i}(\beta_\mathbf{i}^{(m)})}\right) \|\beta_\mathbf{i}^{(m+1)} - \beta_\mathbf{i}^{(m)}\|
\end{aligned}
$$

As a result, we have the following:

$$\|\mathbf{p_i}^{(m+1)} - \mathbf{p_i}^{(m)}\| \leq \mathcal{K}_i\|\mathbf{p_{-i}}^{(m)} - \mathbf{p_{-i}}^{(m-1)}\| + \mathcal{K}'_i\|\beta_\mathbf{i}^{(m+1)} - \beta_\mathbf{i}^{(m)}\|$$

Now, using the fact that $\|\beta_\mathbf{i}^{(m+1)} - \beta_\mathbf{i}^{(m)}\| \leq \mathcal{K}''_i\|\mathbf{p_{-i}}^{(m)} - \mathbf{p_{-i}}^{(m-1)}\|$, we get:

$$\|\mathbf{p_i}^{(m+1)} - \mathbf{p_i}^{(m)}\| \leq (\mathcal{K}_i + \mathcal{K}'_i\mathcal{K}''_i)\,\|\mathbf{p_{-i}}^{(m)} - \mathbf{p_{-i}}^{(m-1)}\|$$

From Assumption 5.2, we have $\mathcal{K}_i + \mathcal{K}'_i\mathcal{K}''_i \leq \frac{1-\eta}{N-1}$. Hence:

$$\|\mathbf{p_i}^{(m+1)} - \mathbf{p_i}^{(m)}\| \leq \frac{1-\eta}{N-1}\|\mathbf{p_{-i}}^{(m)} - \mathbf{p_{-i}}^{(m-1)}\|$$

Summing over all i, we get:

$$\|\mathbf{p}^{(m+1)} - \mathbf{p}^{(m)}\| \leq (1-\eta)\|\mathbf{p}^{(m)} - \mathbf{p}^{(m-1)}\|$$

Hence, the mapping of the approach is a contraction. As a result, the prices converge to a vector of prices denoted $\mathbf{p}^*$.
Since $\|\beta_{\mathbf{i}}{}^{(m+1)} - \beta_{\mathbf{i}}{}^{(m)}\| \leq \mathcal{K}''_i\|\mathbf{p_{-i}}^{(m)} - \mathbf{p_{-i}}^{(m-1)}\|$, the coefficients also converge to a value denoted β^*.
For all i, $\mathbf{p_i^*}$ solves $\mathcal{BR}_i(\mathbf{p^*_{-i}}, \beta^*_{\mathbf{i}})$. Hence, $\mathbf{p}^*$ is the Nash equilibrium corresponding to price sensitivities β^*. □

5.2 Learning under Cournot Competition

In this section, we review the learning approach under Cournot competition. We then study the convergence of the approach. First, we introduce the assumptions under which the dynamic allocation with learning approach converges. Then we state and prove the learning result.

5.2.1 The Learning Approach

The learning approach is an iterative process whereby firms update their best-response strategy and the estimate regarding their allocation sensitivities at each period. All the firms in the market are assumed to use the learning approach.
The approach can be described by the following algorithm: at each period m, given competitors' vector of quantities for the previous period $\mathbf{q_{-i}}^{(m-1)}$, each firm computes its best-response quantity as a function of allocation sensitivities, i.e $\mathcal{Q}_i(\mathbf{q_{-i}}^{(m-1)}, \alpha_{\mathbf{i}})$. Each firm then uses the best-response price function in the estimation problem, in order to update its estimate of the price sensitivities. The new sensitivities at m are denoted $\alpha_{\mathbf{i}}{}^{(m)}$. Hence, firm i's vector of best-response quantities at period m is $\mathbf{q_i}^{(m)} = \mathcal{Q}_i(\mathbf{q_{-i}}^{(m-1)}, \alpha_{\mathbf{i}}{}^{(m)})$.
Learning of the equilibrium strategies is said to be achieved when the difference between the vectors of best-response quantities at period m and at period $m-1$ do not differ by more than a small constant $\eta > 0$, in other words:

$$\|\mathbf{q}^{(m-1)} - \mathbf{q}^{(m)}\| \leq \eta$$

Constant η is the precision of learning. Similarly, learning of the price sensitivities is achieved in precision η if:

$$\|\alpha^{(m-1)} - \alpha^{(m)}\| \leq \eta$$

Algorithm 2: Learning approach under Cournot competition
Input: A set of initial values for the quantities and sensitivities
Output: The equilibrium quantities and optimal sensitivities

```
for i = 1 to N
    Initialize q_i ← q_i^(0)
    Initialize α_i ← α_i^(0)
    Initialize m = 1
    repeat
        for i = 1 to N
            Compute Q_i(q_{-i}^(m-1), α_i)
            Compute α_i^(m)
            Set q_i^(m) ← Q_i(q_{-i}^(m-1), α_i^(m))
            m ← m + 1
        until ‖q^(m-1) − q^(m)‖ ≤ η and ‖α^(m-1) − α^(m)‖ ≤ η
```

Note that in this chapter, we do not explain how the best-response quantities and estimated parameters are computed in practice. Their computation will be discussed in depth in Chapter 6.

5.2.2 Assumptions

We make the following assumptions on the gradient of the revenue function:

Assumptions 5.3.

- *$-\nabla_i \pi_i$ is Lipschitz continuous w.r.t $\mathbf{q_{-i}}$ with constant $\mathcal{L}_{\pi_i,q_{-i}}(\alpha_\mathbf{i})$*
- *$-\nabla_i \pi_i$ is strongly monotone w.r.t $\mathbf{q_i}$ with constant $\mathcal{M}(\pi_i, q_i)(\alpha_\mathbf{i})$*
- *$-\nabla_i \pi_i$ is Lipschitz continuous w.r.t $\alpha_\mathbf{i}$ with constant $\mathcal{L}_{\pi,\alpha_i}(\mathbf{q_i}, \mathbf{q_{-i}})$*

We introduce the following constants:

$$\mathcal{K}_i(\alpha_\mathbf{i}) = \frac{\mathcal{L}_{\pi_i,q_{-i}}(\alpha_\mathbf{i})}{\mathcal{M}_{\pi_i,q_i}(\alpha_\mathbf{i})}$$

$$\mathcal{K}'_i(\mathbf{q_i}, \mathbf{q_{-i}}, \alpha_\mathbf{i}) = \frac{\mathcal{L}_{\pi_i,\alpha_i}(\mathbf{q_i}, \mathbf{q_{-i}})}{\mathcal{M}_{\pi_i,q_i}(\alpha_\mathbf{i})}$$

We make the following assumptions regarding the above constants:

Assumptions 5.4.

- *At each period (m) of the approach, we define $\mathcal{A}_i^{(m)} = \{\alpha_\mathbf{i} \in \mathcal{A}_i \ : \ \|\alpha_\mathbf{i} - {\alpha_\mathbf{i}}^{(m)}\| \leq \mathcal{K}''_i \|{\mathbf{q_i}}^{(m)} - {\mathbf{q_i}}^{(m-1)}, \ \|\mathbf{q_i}\| \leq M\}$.*
The estimation step is performed on the reduced feasible set $\mathcal{A}_i^{(m)}$

- *We define the following constants:*

$$
\begin{array}{rc}
\mathcal{K}_i = & \max_{\alpha_\mathbf{i}} \mathcal{K}_i(\alpha_\mathbf{i}) \\
s.t & \|\alpha_\mathbf{i}\| \leq M \\
\mathcal{K}'_i = & \max_{\mathbf{q_i}, \mathbf{q_{-i}}, \alpha_\mathbf{i}} \mathcal{K}'_i(\mathbf{q_i}, \mathbf{q_{-i}}, \alpha_\mathbf{i}) \\
s.t & \mathbf{0} \leq \mathbf{q_i} \leq C_i \mathbf{e} \\
 & \mathbf{0} \leq \mathbf{q_{-i}} \leq \mathbf{C_{-i}} \\
 & \|\alpha_\mathbf{i}\| \leq M
\end{array}
$$

Let $\mathcal{A} = \max_{i,\alpha_i} \mathcal{A}_i(\alpha_i)$ *and* $\mathcal{B} = \max_{i,\alpha_i} \mathcal{B}_i(\alpha_i)$.
There exists $\eta > 0$ *such that* $\mathcal{A}\mathcal{K} + \mathcal{B}(N-1) \leq 1 - \eta$

5.2.3 Statement and Proof

Theorem 5.2. *Under Assumptions 5.3 and 5.4, the sequence of iterates* $\{\alpha^{(m)}, \mathbf{q}^{(m)}\}$ *generated by the approach converges as* m *goes to infinity to* $(\alpha^*, \mathbf{q}^*)$ *such that* $\mathbf{q}^*$ *is the set of Nash equilibrium policies corresponding to parameters* α^*.

In what follows, we denote by $\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i})$ the best-response quantity, when competitors' prices are $\mathbf{q_{-i}}^{(m-1)}$, and $\mathbf{q_i}^{(m)} = \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}{}^{(m)})$. The corresponding prices are $\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i})$ and $\mathbf{q_i}^{(m)}$ respectively.
The proof relies on a contraction argument. We indeed prove that under the above assumptions, the mapping corresponding to the approach is a contraction. Hence, the sequence of quantities and parameters converge.

Proof. The variational inequality corresponding to $\mathcal{BR}_i(q_{-i}^{(m-1)}, \alpha_\mathbf{i})$ is:

$$-\nabla_i \pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), q_{-i}^{(m-1)}, \alpha_\mathbf{i})'(\mathbf{q_i} - \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i})) \geq 0 \qquad \forall\ \alpha_\mathbf{i} \in \mathcal{A}_i,\ \forall\ \mathbf{q_i} \in \mathfrak{Q}_i$$

Similarly, the variational inequality corresponding to $\mathcal{BR}_i(q_{-i}^{(m)}, \alpha_\mathbf{i})$ is:

$$
\begin{gathered}
-\nabla_i \pi_i(\mathbf{q_i}^{(m+1)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m)}, \alpha_\mathbf{i})' \\
(\mathbf{q_i} - \mathbf{q_i}^{(m+1)}(\alpha_\mathbf{i})) \geq 0 \qquad \forall\ \alpha_\mathbf{i} \in \mathcal{A}_i,\ \forall\ \mathbf{q_i} \in \mathfrak{Q}_i
\end{gathered}
$$

Applying $\mathbf{q_i} = \mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i})$ to the former, and $\mathbf{q_i} = \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i})$ to the latter and summing them up yields:

$$
\begin{aligned}
&\Big(-\nabla_i \pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i}) + \nabla_i \pi_i(\mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m)}, \alpha_\mathbf{i})\Big)' \\
&\quad . \quad (\mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i}) - \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i})) \geq 0
\end{aligned}
$$

Adding and subtracting $\nabla_i \pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m)}, \alpha_\mathbf{i})$ and rearranging terms, the above inequality becomes:

$$
\begin{aligned}
&\Big(-\nabla_i \pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i}) + \nabla_i \pi_i(q_i^{(m)}(\alpha_i), q_{-i}^{(m)}, \alpha_i)\Big)' (q_i^{(m+1)}(\alpha_i) - q_i^{(m)}(\alpha_i)) \geq \\
&\Big(-\nabla_i \pi_i(\mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m)}, \alpha_\mathbf{i}) + \nabla_i \pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m)}, \alpha_\mathbf{i})\Big)' (\mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i}) - \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}))
\end{aligned}
$$

By strong monotonicity of $-\nabla_i\pi_i$ w.r.t $\mathbf{q_i}$, the right hand side is nonnegative, hence the left hand side is equal to its absolute value, to which we can apply Cauchy-Schwartz inequality:

$$\begin{aligned} &|\left(-\nabla_i\pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i}) + \nabla_i\pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m)}, \alpha_\mathbf{i})\right)' (\mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i}) - \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}))| \\ &\leq \| -\nabla_i\pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i}) + \nabla_i\pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m)}, \alpha_\mathbf{i})\| \|\amalg_i^{(m+1)}(\alpha_\mathbf{i}) - \amalg_i^{(m)}(\alpha_\mathbf{i})\| \end{aligned}$$

Hence, by Lipschitz continuity of $-\nabla_i\pi_i$ w.r.t $\mathbf{q_{-i}}$, is bounded above by $\mathcal{L}_{\pi_i,q_{-i}}(\alpha_\mathbf{i})\|\mathbf{q_{-i}}^{(m)} - \mathbf{q_{-i}}^{(m-1)}\| \|\mathbf{q_i}^{(m+1)}(\alpha_\mathbf{i}) - \mathbf{q_i}^{(m)}(\alpha_\mathbf{i})\|$. By strong monotonicity, the right hand side is bounded below by $\mathcal{M}_{\pi_i,q_i}(\alpha_\mathbf{i})\|\mathbf{q_i}^{(m+1)}(\alpha_\mathbf{i}) - \mathbf{q_i}^{(m)}(\alpha_\mathbf{i})\|^2$.
Hence, the above inequality becomes:

$$\mathcal{M}_{\pi_i,q_i}(\alpha_\mathbf{i})\|\mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i}) - \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i})\| \leq \mathcal{L}_{\pi_i,q_{-i}}(\alpha_\mathbf{i})\|\mathbf{q_{-i}}^{(m)} - \mathbf{q_{-i}}^{(m-1)}\| \qquad (5.18)$$

Furthermore, the best-response problem $\mathcal{BR}_i(\mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i}^{(m)})$ is equivalent to the following variational inequality:

$$-\nabla_i\pi_i(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i}^{(m)})'(\mathbf{q_i} - \mathbf{q_i}^{(m)}) \geq 0 \qquad , \ \forall\ \mathbf{q_i} \in \mathfrak{Q}_i$$

We apply it to $\mathbf{q_i} = \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i})$, we obtain:

$$-\nabla_i\pi_i(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i}^{(m)})'(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}) - \mathbf{q_i}^{(m)}) \geq 0 \qquad \forall\ \mathbf{q_i} \in \mathfrak{Q}_i$$

The variational inequality corresponding to best-response problem $\mathcal{BR}_i(\mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i})$, applied at $\mathbf{q_i} = \mathbf{q_i}^{(m)}$ is:

$$-\nabla_i\pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i})'(\mathbf{q_i}^{(m)} - \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i})) \geq 0 \qquad \forall\ \alpha_\mathbf{i},\ \forall\ \mathbf{q_i} \in \mathfrak{Q}_i$$

Summing the two inequalities, adding and subtracting $\nabla_i\pi_i(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i})$ and rearranging terms gives:

$$\begin{aligned} &\left(-\nabla_i\pi_i(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i}^{(m)}) + \nabla_i\pi_i(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i})\right)' (\mathbf{q_i}^{(m)}(\alpha_\mathbf{i}) - \mathbf{q_i}^{(m)}) \geq \\ &\quad \left(-\nabla_i\pi_i(\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}), \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i}) + \nabla_i\pi_i(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)}, \alpha_\mathbf{i})\right)' (\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}) - \mathbf{q_i}^{(m)}) \end{aligned}$$

Using Lipschitz continuity of $-\nabla_i\pi_i$ w.r.t $\alpha_\mathbf{i}$ and strong monotonicity of $-\nabla_i\pi_i$ w.r.t $\mathbf{q_i}$ in a similar manner as above yields:

$$\mathcal{M}_{\pi_i,q_i}(\alpha_\mathbf{i})\|\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}) - \mathbf{q_i}^{(m)}\| \leq \mathcal{L}_{\pi_i,\alpha_i}(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)})\|\alpha_\mathbf{i}^{(m)} - \alpha_\mathbf{i}\| \qquad (5.19)$$

Reiterating the process for step $m+1$ instead of m yields:

$$\mathcal{M}_{\pi_i,q_i}(\alpha_\mathbf{i})\|\mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i}) - \mathbf{q_i}^{(m+1)}\| \leq \mathcal{L}_{\pi_i,\alpha_i}(\mathbf{q_i}^{(m+1)}, \mathbf{q_{-i}}^{(m)})\|\alpha_\mathbf{i}^{(m+1)} - \alpha_\mathbf{i}\| \quad (5.20)$$

We have the following:

$$\|\mathbf{q_i}^{(m+1)} - \mathbf{q_i}^{(m)}\| \leq \|\mathbf{q_i}^{(m+1)} - \mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i})\| + \|\mathcal{Q}_i^{(m+1)}(\alpha_\mathbf{i}) - \mathcal{Q}_i^{(m)}(\alpha_\mathbf{i})\| + \|\mathcal{Q}_i^{(m)}(\alpha_\mathbf{i}) - \mathbf{q_i}^{(m)}\|$$

Hence, using Inequalities (5.14), (5.15) and (5.16) we have:

$$\begin{aligned} &\|\mathbf{q_i}^{(m+1)} - \mathbf{q_i}^{(m)}\| \\ &\quad \leq \quad \frac{\mathcal{L}_{\pi_i,\alpha_i}(\mathbf{q_i}^{(m+1)}, \mathbf{q_{-i}}^{(m)})}{\mathcal{M}_{\pi_i,q_i}(\alpha_\mathbf{i})}\|\alpha_\mathbf{i}^{(m+1)} - \alpha_\mathbf{i}\| \end{aligned}$$

$$+ \quad \frac{\mathcal{L}_{\pi_i,q_{-i}}(\alpha_{\mathbf{i}})}{\mathcal{M}_{\pi_i,q_i}(\alpha_{\mathbf{i}})}\|\mathbf{q_{-i}}^{(m)} - \mathbf{q_{-i}}^{(m-1)}\|$$
$$+ \quad \frac{\mathcal{L}_{\pi_i,\alpha_i}(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)})}{\mathcal{M}_{\pi_i,q_i}(\alpha_{\mathbf{i}})}\|\alpha_{\mathbf{i}}^{(m)} - \alpha_{\mathbf{i}}\|$$

Using the constants we introduced, we get:

$$\begin{aligned} &\|\mathbf{q_i}^{(m+1)} - \mathbf{q_i}^{(m)}\| \\ \leq\;& \mathcal{K}'_i(\mathbf{q_i}^{(m+1)}, \mathbf{q_{-i}}^{(m)}, \alpha_{\mathbf{i}})\|\alpha_{\mathbf{i}}^{(m+1)} - \alpha_{\mathbf{i}}\| \\ +\;& \mathcal{K}_i(\alpha_{\mathbf{i}})\|\mathbf{q_{-i}}^{(m)} - \mathbf{q_{-i}}^{(m-1)}\| \\ +\;& \mathcal{K}'_i(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)}, \alpha_{\mathbf{i}})\|\alpha_{\mathbf{i}}^{(m)} - \alpha_{\mathbf{i}}\| \end{aligned}$$

The above inequality holds for every feasible $\alpha_{\mathbf{i}}$, in particular for $\alpha_{\mathbf{i}} = \alpha_{\mathbf{i}}^{(m+1)}$:

$$\begin{aligned} &\|\mathbf{q_i}^{(m+1)} - \mathbf{q_i}^{(m)}\| \\ \leq\;& \mathcal{K}_i(\alpha_{\mathbf{i}}^{(m+1)})\|\mathbf{q_{-i}}^{(m)} - \mathbf{q_{-i}}^{(m-1)}\| + \mathcal{K}'_i(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)}, \alpha_{\mathbf{i}}^{(m+1)})\|\alpha_{\mathbf{i}}^{(m)} - \alpha_{\mathbf{i}}^{(m+1)}\| \end{aligned}$$

Since $\mathcal{K}_i(\alpha_{\mathbf{i}}^{(m+1)}) \leq \mathcal{K}_i$ and $\mathcal{K}'_i(\mathbf{q_i}^{(m)}, \mathbf{q_{-i}}^{(m-1)}, \alpha_{\mathbf{i}}^{(m+1)}) \leq \mathcal{K}'_i$, we have:

$$\|\mathbf{q_i}^{(m+1)} - \mathbf{q_i}^{(m)}\| \leq \mathcal{K}_i\|\mathbf{q_{-i}}^{(m)} - \mathbf{q_{-i}}^{(m-1)}\| + \mathcal{K}'_i\|\alpha_{\mathbf{i}}^{(m)} - \alpha_{\mathbf{i}}^{(m+1)}\|$$

According to Assumption 5.4, $\|\alpha_{\mathbf{i}}^{(m)} - \alpha_{\mathbf{i}}^{(m+1)}\| \leq \mathcal{K}''_i\|\mathbf{p_{-i}}^{(m-1)} - \mathbf{p_{-i}}^{(m)}\|$, hence we have:

$$\|\mathbf{q_i}^{(m+1)} - \mathbf{q_i}^{(m)}\| \leq (\mathcal{K}_i + \mathcal{K}'_i\mathcal{K}''_i)\|\mathbf{q_{-i}}^{(m)} - \mathbf{q_{-i}}^{(m-1)}\|$$

According to 5.4, we also have:

$$\mathcal{K}_i + \mathcal{K}'_i\mathcal{K}''_i \leq \frac{1-\eta}{N-1}$$

Summing the inequality over all i yields:

$$\|\mathbf{q}^{(m+1)} - \mathbf{q}^{(m)}\| \leq (1-\eta)\|\mathbf{q}^{(m)} - \mathbf{q}^{(m-1)}\|$$

Hence, the map of the learning process is contracting. As a result, the process converges.

Denote by $\mathbf{q}^*$, α^* the limits of the learning process. By definition, for all i, $\mathbf{q_i^*}$ is the best-response to strategy $\mathbf{q_{-i}^*}$, when parameters are $\alpha_{\mathbf{i}}^*$. Hence, $\mathbf{q}^*$ are the Nash equilibrium quantities, corresponding to sensitivities α^*. □

Chapter 6

Computational Results

In this chapter, we address the implementation of the approach. We first focus on the best-response problem and Nash equilibrium problem when the parameters of the price-demand relationship are known. The best-response problems under Bertrand and Cournot competition are convex optimization problems which can be solved using traditional convex optimization algorithms. Furthermore, we proved in Chapter 3, Proposition 3.20 that the Cournot equilibrium problem is equivalent to a variational inequality, for which efficient computation methods exist. We discuss relaxation algorithms, which provide efficient ways to compute variational inequality solutions, and their applications to solving the Cournot equilibrium problem. We then discuss computation of the solution to the Bertrand equilibrium. We showed in Chapter 3 Proposition 3.15 that the Bertrand generalized equilibrium problem is equivalent to a quasi-variational inequality. In this Chapter, we show that the quasi-variational inequality can be approximated by a sequence of penalized variational inequalities, which converge to the quasi-variational inequality.

We then turn to the joint dynamic pricing (resp. allocation) with learning problem. We propose an iterative solution method based on the Gauss-Newton method for non-linear least squares problem. This method takes advantage of the special structure of the upper-level estimation problem to efficiently compute solutions to the bilevel program under Cournot competition, and to the MPEC under Bertrand competition.

6.1 Algorithms When the Parameters of the Price-Demand Relationship Are Known

In this section, we assume that the parameters in the price-demand relationship are known and we therefore focus on the lower-level problem, i.e the best-response problem of a firm in the disequilibrium state, and the market equilibrium in the equilibrium state. In disequilibrium, the best-response problem under Bertrand or Cournot competition is a concave maximization problem. Hence, its solution does not present any particular issue.

Under Bertrand competition, the equilibrium problem was formulated as a quasi variational inequality (QVI) in Chapter 3 Proposition 3.15. The difficulty of solving such a QVI resides in the fact that it is defined on a feasible mapping, which depends on the solution to the QVI. This is due to the fact that the capacity constraint of each firm involves competitors' capacity. To address the issue, we approximate the QVI in Section 6.1.1 as a variational inequality by penalizing the capacity constraint. More precisely, we move the capacity constraint into the objective in the form of a penalty. We show that the sequence of penalized variational inequalities converges to the QVI. Under Cournot competition, the equilibrium problem was shown to be equivalent to a variational inequality in Chapter 3, Proposition 3.20, which can be solved efficiently using relaxation algorithms. We describe these algorithms in Section 6.1.2. These algorithms can also be used to solve the penalized variational inequality under Bertrand competition. Relaxation algorithms solve the variational inequality as a sequence of subproblems which are in general nonlinear optimization problems. In particular, since the variational inequality corresponding to the Nash equilibrium decomposes into N coupled variational inequalities corresponding to the firms' best-response problems, we can use decomposition methods, which are specialized relaxation methods for this type of problems. Another feature that can be exploited is the special network structure of the Nash equilibrium problem. Indeed, the problem can be considered as a traffic equilibrium problem in a network, as explained in this section.

6.1.1 Penalty Method for the Bertrand Equilibrium

We focus on solving the lower-level problem under Bertrand competition when the market is in equilibrium. By Proposition 3.15, the generalized Nash equilibrium under Bertrand competition is equivalent to the following quasi variational inequality: find $\mathbf{p} \in \mathfrak{P}(\mathbf{p}, \beta)$ such that for all $\mathbf{p}' \in \mathfrak{P}(\mathbf{p}, \beta)$, we have:

$$\mathbf{f}(\mathbf{p}', \beta)'(\mathbf{p}' - \mathbf{p}) \geq 0$$

where we defined:

- $\mathbf{f}(\mathbf{p}, \beta)$ is the vector-valued function with components $\mathbf{f_i}(\mathbf{p}, \beta) = -\nabla_i \pi_i(\mathbf{p_i}, \mathbf{p_{-i}}, \beta_\mathbf{i})$;
- $\mathfrak{P}(\mathbf{p}, \beta) = \{\mathbf{0} \leq \mathbf{p}' \leq \mathbf{p}^{max} : g_i(\mathbf{p_i}, \mathbf{p}'_\mathbf{-i}, \beta_\mathbf{i}\ \forall\ i = 1, \ldots, N) \leq 0\}$ is the feasible mapping;
- $g_i(\mathbf{p_i}, \mathbf{p}'_\mathbf{-i}, \beta_\mathbf{i}) = \sum_{t=1}^{T} q_i(p_i(t), p_{-i}(t), \beta_i(t)) - C_i$ is the capacity constraint.

Let us define by $\mathbf{g}(\mathbf{p}, \mathbf{p}', \beta)$ the vector-valued function with components $g_i(\mathbf{p_i}, \mathbf{p}'_\mathbf{-i}, \beta_\mathbf{i})$, and denote $[\mathbf{0}, \mathbf{p}^{max}] = [0, p_i^{max}]^{NT}$. We now introduce the penalized variational inequality:

Definition 6.1. *Let $\{\rho^k\}$ be a nonnegative increasing sequence of scalars such that $\lim_{k\to\infty} \rho^k = \infty$ and u_i^k be a bounded sequence of scalars. The penalized variational inequality is: find $\mathbf{p^k} \in [\mathbf{0}, \mathbf{p}^{max}]$ such that for all $\mathbf{p} \in [\mathbf{0}, \mathbf{p}^{max}]$:*

$$\mathbf{f^k}(\mathbf{p^k}, \beta)'(\mathbf{p} - \mathbf{p^k}) \geq 0$$

where $\mathbf{f^k}(\mathbf{p}, \beta) = \mathbf{f}(\mathbf{p}, \beta) + \max\{0, \mathbf{u^k} + \rho^k \mathbf{g}(\mathbf{p}, \mathbf{p}, \beta)\} \nabla \mathbf{g}(\mathbf{p}, \mathbf{p}, \beta)$

We now establish our result.

Theorem 6.1. *Any limit point of the sequence of prices generated by the sequential penalty method converges to the solution of the generalized Nash equilibrium.*

The proof was established in Fukushima and Pang [44] under the Mangasarian-Fromovitz constraint qualification for the penalized variational inequality and an additional assumption that we verify below.

Proof. Let $\mathbf{p^k}$ be the solution to the penalized variational inequality k.
Let us check that the two assumptions under which the result from Fukushima and Pang [44] was proved, hold for the Bertrand equilibrium problem:

- Mangasarian-Fromovitz constraint qualification for the penalized variational inequality: if the following equality holds for some $\bar{\mu}, \underline{\mu} \geq \mathbf{0}$:

$$-\sum_{i=1}^{N}\sum_{t\in\underline{\mathcal{I}}_i^k}\underline{\mu}_i(t)\mathbf{e}(\mathbf{t}) + \sum_{i=1}^{N}\sum_{t\in\bar{\mathcal{I}}_i^k}\bar{\mu}_i(t)\mathbf{e}(\mathbf{t}) = 0$$

  then $\bar{\mu}, \underline{\mu} = \mathbf{0}$, where $\underline{\mathcal{I}}_i^k = \{t : p_i(t) = 0\}$, and $\bar{\mathcal{I}}_i^k = \{t : p_i(t) = p_i^{max}\}$.
  Since $\underline{\mathcal{I}}_i^k \bigcap \bar{\mathcal{I}}_i^k = \emptyset$, we have:

$$-\sum_{i=1}^{N}\sum_{t\in\underline{\mathcal{I}}_i^k}\underline{\mu}_i(t)\mathbf{e}(\mathbf{t}) + \sum_{i=1}^{N}\sum_{t\in\bar{\mathcal{I}}_i^k}\bar{\mu}_i(t)\mathbf{e}(\mathbf{t}) = \begin{cases} -\underline{\mu}_i(t) & \forall\ t \in \underline{\mathcal{I}}_i^k \\ \bar{\mu}_i(t) & \forall\ t \in \bar{\mathcal{I}}_i^k \end{cases}$$

  Hence, the vector is equal to zero iff $\underline{\mu}_i(t) = 0$ for all $t \in \underline{\mathcal{I}}_i^k$ and $\bar{\mu}_i(t) = 0$ for all $t \in \bar{\mathcal{I}}_i^k$. Therefore, the constraint qualification holds.

- Let $\mathbf{p}^\infty$ be a limit point of the sequence $\mathbf{p}^k$. Let $\mathcal{J}^\infty = \{i : g_i(\mathbf{p_i}^\infty, \mathbf{p_{-i}}^\infty, \beta_\mathbf{i} \geq 0)\}$ be the set of indices for which the capacity constraint is violated. We need to check that if the following equality holds for some $\lambda, \underline{\mu}, \bar{\mu} \geq \mathbf{0}$:

$$\sum_{i\in\mathcal{J}^\infty}\lambda_i\nabla_{p'}\mathbf{g_i}(\mathbf{p_i}^\infty, \mathbf{p_{-i}}^\infty, \beta_\mathbf{i}) - \sum_{i=1}^{N}\sum_{\underline{\mathcal{I}}_i^\infty}\underline{\mu}_i(t)\mathbf{e_i}(\mathbf{t}) + \sum_{i=1}^{N}\sum_{\bar{\mathcal{I}}_i^\infty}\bar{\mu}_i(t)\mathbf{e_i}(\mathbf{t}) = 0$$

  then $\lambda, \underline{\mu}, \bar{\mu} = \mathbf{0}$.
  $\nabla_{p'}\mathbf{g_i}(\mathbf{p_i}, \mathbf{p_{-i}}, \beta_\mathbf{i})$ is the vector of components:

$$\frac{\partial g_i}{\partial p'_j(t)} = \begin{cases} 0 & \text{if } j \neq i \\ \frac{\partial q_i}{\partial p_i(t)} & \text{s.t } j = i \end{cases}$$

We have:

$$\sum_{i\in\mathcal{J}^\infty}\lambda_i\nabla_{p'}\mathbf{g_i}(\mathbf{p_i}^\infty,\mathbf{p_{-i}}^\infty,\beta_\mathbf{i})-\sum_{i=1}^{N}\sum_{\underline{\mathcal{I}}_i^\infty}\underline{\mu}_i(t)\mathbf{e_i}(\mathbf{t})+\sum_{i=1}^{N}\sum_{\bar{\mathcal{I}}_i^\infty}\bar{\mu}_i(t)\mathbf{e_i}(\mathbf{t})=$$

$$\begin{cases} \lambda_i\frac{\partial q_i}{\partial p_i(t)}-\underline{\mu}_i(t) & \forall\ i\in\mathcal{J}^\infty,\ \forall\ t\in\underline{\mathcal{I}}_i^\infty \\ \lambda_i\frac{\partial q_i}{\partial p_i(t)}+\bar{\mu}_i(t) & \forall\ i\in\mathcal{J}^\infty,\ \forall\ t\in\bar{\mathcal{I}}_i^\infty \\ \lambda_i\frac{\partial q_i}{\partial p_i(t)} & \forall\ i\in\mathcal{J}^\infty,\ \forall\ t\notin\underline{\mathcal{I}}_i^\infty\bigcup\bar{\mathcal{I}}_i^\infty \end{cases}$$

Since $\frac{\partial q_i}{\partial p_i(t)}<0$, and $\lambda_i,\underline{\mu}_i(t)\geq 0$, the first type of components above is equal to zero implies $\lambda_i=\underline{\mu}_i(t)=0$. Since $\lambda_i=0$, the second type of components is equal to zero iff $\underline{\mu}_i(t)=0$. The third type of components is equal to zero iff $\lambda_i=0$. Therefore, the assumption holds.

$\square$

Note that the penalized variational inequality can also be written as:

$$\sum_{i=1}^{N}\left(-\nabla_i\pi_i(\mathbf{p_i}^k,\mathbf{p_{-i}}^k,\beta_\mathbf{i})+\max\{0,\mathbf{u_i^k}+\rho^k\mathbf{g_i}(\mathbf{p_i}^k,\mathbf{p_{-i}}^k,\beta_\mathbf{i})\}\nabla_p\mathbf{g_i}(\mathbf{p_i}^k,\mathbf{p_{-i}}^k,\beta_\mathbf{i})\right)'(\mathbf{p_i}-\mathbf{p_i}^k)\geq 0$$

Similarly to Chapter 3, it is easy to prove that the above penalized variational inequality is equivalent to N penalized variational inequalities, one for each firm:

$$\left(-\nabla_i\pi_i(\mathbf{p_i}^k,\mathbf{p_{-i}}^k,\beta_\mathbf{i})+\max\{0,\mathbf{u_i^k}+\rho^k\mathbf{g_i}(\mathbf{p_i}^k,\mathbf{p_{-i}}^k,\beta_\mathbf{i})\}\nabla_p\mathbf{g_i}(\mathbf{p_i}^k,\mathbf{p_{-i}}^k,\beta_\mathbf{i})\right)'(\mathbf{p_i}-\mathbf{p_i}^k)\geq 0$$

Furthermore, we can prove as in Chapter 3 that the above penalized variational inequality is equivalent to the following optimization problem:

$$\begin{array}{ll}\max\limits_{\mathbf{p_i}} & \pi_i(\mathbf{p_i},\mathbf{p_{-i}^k},\beta_\mathbf{i})-\frac{1}{2\rho^k}\max\left\{0,u_i^k+\rho^k g_i(\mathbf{p_i},\mathbf{p_{-i}}^k,\beta_\mathbf{i})\right\}^2 \\ \text{s.t} & \qquad\qquad\qquad \mathbf{0}\leq\mathbf{p_i}\leq\mathbf{p_i}^{max}\end{array}$$

Hence, the penalized variational inequality has the interpretation of each firm solving a penalized revenue maximization problem, where the revenue function is penalized when the capacity constraint is violated.
In summary, the penalty method applied to the generalied Nash equilibrium yields the following algorithm:

Algorithm 3: Penalty algorithm
Input: Initial prices $\mathbf{p}^0$, sequence of scalars $\{\rho^k\}$, and $\{u_i^k\}$
Output: The equilibrium prices $\mathbf{p}*$
(1) **for** $i = 1$ **to** N
(2) Initialize $\mathbf{q_i} \leftarrow \mathbf{q_i}^{(0)}$
(3) Initialize $k = 1$
(4) **repeat**
(5) Compute $\mathbf{p}^k$ solution to VI with objective function $\mathbf{f}^k$
(6) $k \leftarrow k + 1$
(7) **until** $\|\mathbf{p}^{k-1} - \mathbf{p}^k\| \leq \eta$

In this section, we do not enter into details about how to solve the variational inequality. Section 6.1.2 below is dedicated to algorithms for solving variational inequalities.

6.1.2 Decomposition Methods for Variational Inequalities

Decomposition methods are particularly adapted to problems where the variational inequality is defined over a set that is a cartesian product. In that case, the variational inequality decomposes into N coupled variational inequality, one for each component of the cartesian product. In Chapter 3, we saw that the Cournot Nash equilibrium corresponds to such a problem, and we proved in the above Section that the penalized Bertrand Nash equilibrium is such a problem as well.
In what follows, we focus on the Cournot Nash equilibrium for the sake of brevity. The results apply readily to the penalized Nash equilibrium. We present the Gauss-Seidel method of decomposition. This method serially updates the value of the iterate for each firm, and the Jacobi decomposition, which simultaneously updates the value of the iterate for all the firms.
Both methods rely on the equivalence between the following:

- Variational inequality $VI(\mathbf{F}, \mathfrak{Q})$:
  find $\mathbf{q}* \in \mathfrak{Q}$ with $\mathfrak{Q} = \mathfrak{Q}_1 \times \ldots \times \mathfrak{Q}_N$ such that for all $\mathbf{q} \in \mathfrak{Q}$:

  $$\mathbf{F}(\mathbf{q}^*)'(\mathbf{q} - \mathbf{q}*) \geq 0$$

- N coupled variational inequalities $VI(\mathbf{F_i}, \mathfrak{Q}_i)$: find $\mathbf{q}*_\mathbf{i} \in \mathfrak{Q}_i$ such that for all $\mathbf{q_i} \in \mathfrak{Q}_i$:

  $$\mathbf{F_i}(\mathbf{q_i}*)'(\mathbf{q_i} - \mathbf{q_i}*) \geq 0$$

A Serial Decomposition Algorithm

The serial decomposition algorithm works the following way: at each iteration k, $VI(\mathbf{F_i}, \mathfrak{Q}_i)$ is solved for each firm i, where competitors' quantities are fixed at their current value, i.e $\mathbf{q_1}^k, \ldots, \mathbf{q_{i-1}}^k$ and $\mathbf{q_{i+1}}^{k-1}, \ldots, \mathbf{q_N}^{k-1}$; the value of the iterate $\mathbf{q_i}^k$

is updated to its new value, and the next variational inequality $VI(\mathbf{F_{i+1}}, \mathfrak{Q}_{i+1})$ is solved for $\mathbf{q_{i+1}}$. When all N variational inequalities are solved, the algorithm moves to the next iteration $k+1$. The algorithm terminates when the convergence criterion $\|\mathbf{q}^{k-1} - \mathbf{q}^k\| \leq \eta$ is met.

Algorithm 4: Serial algorithm
Input: A set of initial values for the quantities $\mathbf{p}^0$
Output: The equilibrium quantities $\mathbf{q}*$

for $i = 1$ **to** N
  Initialize $\mathbf{q_i} \leftarrow \mathbf{q_i}^{(0)}$
  Initialize $k = 1$
  repeat
    for $i = 1$ **to** N
      Fix $\mathbf{q_j} \leftarrow \mathbf{q_j}^k,\ j < i$ and $\mathbf{q_j} \leftarrow \mathbf{q_j}^{k-1},\ j > i$
      Compute $\mathbf{q_i}^k$ solution of $VI(\mathbf{F_i}, \mathfrak{Q}_i)$
      If $i = N$ then $k \leftarrow k+1$
    until $\|\mathbf{q}^{k-1} - \mathbf{q}^k\| \leq \eta$

A Parallel Decomposition Algorithm

The parallel decomposition algorithm works as follows: at each iteration k, $VI(\mathbf{F_i}, \mathfrak{Q}_i)$ is solved for each firm i, where competitors' quantities are fixed at $\mathbf{q_{-i}}^{k-1}$. The solution to this variational inequality is denoted $\mathbf{q_i}^k$ and is set aside until the N variational inequalities of iteration k are solved. Then the algorithm moves to iteration $k+1$. Hence, the update of the quantities is done simultaneously for all firms. Notice that contrary to the serial method, at iteration k, the variational inequality corresponding to firm i is solved, without quantities $\mathbf{q_1}, \ldots, \mathbf{q_{i-1}}$ being updated to their new value $\mathbf{q_1}^k, \ldots, \mathbf{q_{i-1}}^k$. The update is performed once and for all at the end of iteration k, once the N variational inequalities have been solved. As a result, the N variational inequalities can be solved in parallel. The algorithm stops when the convergence criterion $\|\mathbf{q}^{k-1} - \mathbf{q}^k\| \leq \eta$ is reached.

Algorithm 5: Parallel algorithm
Input: A set of initial values for the quantities $\mathbf{p}^0$
Output: The equilibrium quantities $\mathbf{q}*$

for $i = 1$ **to** N
  Initialize $\mathbf{q_i} \leftarrow \mathbf{q_i}^{(0)}$
  Initialize $k = 1$
  repeat
    for $i = 1$ **to** N
      Fix quantities $\mathbf{q_{-i}} \leftarrow \mathbf{q_{-i}}^{k-1}$
      Compute $\mathbf{q_i}^k$ solution of $VI(\mathbf{F_i}, \mathfrak{Q}_i)$
      If $i = N$ then $k \leftarrow k+1$
    until $\|\mathbf{q}^{k-1} - \mathbf{q}^k\| \leq \eta$

Convergence of the Relaxation Algorithms

For the sake of brevity, we focus on the parallel version of the algorithm. The proof for the serial version uses similar ideas. We prove that under Assumptions 5.3 bullets 1 and 2, the algorithm map is a contraction.
Recall the following constants:

- $\mathcal{L}_{\pi_i,q_{-i}}(\alpha_{\mathbf{i}})$: Lipschitz continuity constant;
- $\mathcal{M}_{\pi,q_i}(\alpha_{\mathbf{i}})$: strong monotonicity constant
- $\mathcal{K}_i(\alpha_{\mathbf{i}}) = \frac{\mathcal{L}_{\pi_i,q_{-i}}(\alpha_{\mathbf{i}})}{\mathcal{M}_{\pi,q_i}(\alpha_{\mathbf{i}})}$

Theorem 6.2. *Assume Assumptions 5.3 bullets 1 and 2 hold.*
Assume the constants are such that there exists $\eta > 0$:

$$\max_i \mathcal{K}_i(\alpha_{\mathbf{i}}) \leq \frac{1-\eta}{N-1}$$

Then the relaxation algorithm converges.

We prove the theorem for the parallel relaxation algorithm, since the proofs for the serial and parallel versions are almost identical.

Proof. At iteration k, for each firm i, the following variational inequality is solved:

$$\mathbf{F_i}(\mathbf{q_i}^k, \mathbf{q_{-i}}^{k-1})'(\mathbf{q_i} - \mathbf{q_i}^k) \geq 0 \tag{6.1}$$

Similarly, at iteration $k+1$, we have:

$$\mathbf{F_i}(\mathbf{q_i}^{k+1}, \mathbf{q_{-i}}^{k})'(\mathbf{q_i} - \mathbf{q_i}^{k+1}) \geq 0 \tag{6.2}$$

We apply the former inequality to $\mathbf{q_i} = \mathbf{q_i}^{k+1}$ and the latter to $\mathbf{q_i} = \mathbf{q_i}^k$, and sum them:

$$\left(\mathbf{F_i}(\mathbf{q_i}^k, \mathbf{q_{-i}}^{k-1}) - \mathbf{F_i}(\mathbf{q_i}^{k+1}, \mathbf{q_{-i}}^{k})\right)' (\mathbf{q_i}^{k+1} - \mathbf{q_i}^k) \geq 0$$

We add and subtract $\mathbf{F_i}(\mathbf{q_i}^k, \mathbf{q_{-i}}^{k-1})$, and get:

$$\begin{aligned} &\left(\mathbf{F_i}(\mathbf{q_i}^k, \mathbf{q_{-i}}^{k-1}) - \mathbf{F_i}(\mathbf{q_i}^{k}, \mathbf{q_{-i}}^{k})\right)' (\mathbf{q_i}^{k+1} - \mathbf{q_i}^k) \\ &\quad \geq \quad \left(\mathbf{F_i}(\mathbf{q_i}^{k+1}, \mathbf{q_{-i}}^{k}) - \mathbf{F_i}(\mathbf{q_i}^{k}, \mathbf{q_{-i}}^{k})\right)' (\mathbf{q_i}^{k+1} - \mathbf{q_i}^k) \end{aligned}$$

By strong monotonicity, the right hand side is bounded below by:

$$\mathcal{M}_{\pi,q_i}(\alpha_{\mathbf{i}}) \|\mathbf{q_i}^{k+1} - \mathbf{q_i}^k\|^2$$

Hence, the left hand side is nonnegative, and is equal to its absolute value. By Cauchy-Schwartz inequality, it is bounded above by $\|\mathbf{F_i}(\mathbf{q_i}^k, \mathbf{q_{-i}}^{k-1}) - \mathbf{F_i}(\mathbf{q_i}^k, \mathbf{q_{-i}}^k)\|.\|\mathbf{q_i}^{k+1} - \mathbf{q_i}^k\|$. By Lipschitz continuity, it is therefore bounded above by:

$$\mathcal{L}_{\pi_i,q_{-i}}(\alpha_{\mathbf{i}}) \|\mathbf{q_{-i}}^k - \mathbf{q_{-i}}^{k-1}\|.\|\mathbf{q_i}^{k+1} - \mathbf{q_i}^k\|$$

As a result, we have the following inequality:

$$\|\mathbf{q_i}^{k+1} - \mathbf{q_i}^k\| \leq \frac{\mathcal{L}_{\pi_i,q_{-i}}(\alpha_{\mathbf{i}})}{\mathcal{M}_{\pi,q_i}(\alpha_{\mathbf{i}})} \|\mathbf{q_{-i}}^k - \mathbf{q_{-i}}^{k-1}\|$$

Summing over all i, we get:

$$\|\mathbf{q}^{k+1} - \mathbf{q}^k\| \leq (N-1) \max_i \mathcal{K}_i(\alpha_\mathbf{i}) \|\mathbf{q}^k - \mathbf{q}^{k-1}\|$$

where $\mathcal{K}_i(\alpha_\mathbf{i}) = \frac{\mathcal{L}_{\pi_i, q_{-i}}(\alpha_\mathbf{i})}{\mathcal{M}_{\pi, q_i}(\alpha_\mathbf{i})}$ Since $\max_i \mathcal{K}_i(\alpha_\mathbf{i}) \leq \frac{1-\eta}{N-1}$, we conclude that:

$$\|\mathbf{q}^{k+1} - \mathbf{q}^k\| \leq (1-\eta) \|\mathbf{q}^k - \mathbf{q}^{k-1}\|$$

Hence, the mapping of the algorithm is a contraction, and the algorithm converges.

□

6.2 Algorithms When the Parameters of the Price-Demand Relationship Are Unknown

In this section, we discuss a solution method for solving the bilevel problem in disequilibrium, and the MPEC in equilibrium. Since the upper level estimation problem corresponds to the minimization of the squared norm of a vector-valued function, we choose to adapt the Gauss-Newton to the problem. This method takes advantage of the special form of the upper level. It transforms the bilevel problem into a series of simpler bilevel problems, and the MPEC into a series of simpler MPECs. We first discuss why a more simple and intuitive procedure fails to solve the problem, and then turn to the study of the Gauss-Newton based method. In this section, we focus on the Cournot problem in disequilibrium, and show how the results can be applied to the Cournot problem in equilibrium, as well as the Bertrand disequilibrium and equilibrium problems.

6.2.1 A Simple Iterative Method with a Counterexample

An intuitive way to solve the joint dynamic control with learning problem for one firm is to iterate the following steps:

Algorithm 6: Iterative algorithm
Input: A set of initial values for the parameters α^0
Output: The equilibrium quantities $\mathbf{q}*$ and true parameters α^*
(1) Initialize $\alpha_\mathbf{i} \leftarrow {\alpha_\mathbf{i}}^{(0)}$
(2) Initialize $k = 1$
(3) **repeat**
(4) Fix parameters $\alpha_\mathbf{i} \leftarrow {\alpha_\mathbf{i}}^{k-1}$
(5) Compute ${\mathbf{q}_\mathbf{i}}^k$ solution of $VI(\mathbf{F}_\mathbf{i}(., {\alpha_\mathbf{i}}^{k-1}), \mathfrak{Q}_i)$
(6) Compute ${\alpha_\mathbf{i}}^k$ solution to the estimation problem for $\mathbf{q}_\mathbf{i} \leftarrow {\mathbf{q}_\mathbf{i}}^k$
(7) **until** $\|{\mathbf{q}_\mathbf{i}}^{k-1} - {\mathbf{q}_\mathbf{i}}^k\| \leq \eta$ and $\|{\alpha_\mathbf{i}}^{k-1} - {\alpha_\mathbf{i}}^k\| \leq \eta$

However, it turns out that the above procedure does not converge to the solution of the bilevel problem. Consider indeed the following counterexample:

- Upper level problem:
  Upper level variable $\alpha \in [-1, 1]$; objective function $F(\alpha, d) = (\alpha - d)^2$.
- Lower level problem:
  Lower level variable $d \in [-2, 2]$; objective function $f(\alpha, d) = -(\alpha - d)^2$

Hence, the bilevel problem is:

$$\begin{aligned} \min_{\alpha \in [-1,1]} \quad & F(\alpha, d) \\ \text{s.t} \quad & \min_{d \in [-2,2]} f(\alpha, d) \end{aligned}$$

The bilevel problem has two optimal solutions: $(\alpha^*, d^*) = (-1, 2)$ and $(\alpha^{**}, d^{**}) = (1, -2)$.
However, when one applies the iterative procedure to the above bilevel problem, it converges to $(1, 2)$ or $(-1, -2)$. Note that these are the solutions to the following bilevel problem:

$$\begin{aligned} \min_{d \in [-2,2]} \quad & f(\alpha, d) \\ \text{s.t} \quad & \min_{\alpha \in [-1,1]} \quad F(\alpha, d) \end{aligned}$$

6.2.2 Adaptation of the Gauss-Newton method to the Cournot problem in disequilibrium

In this section, we investigate how to adapt the Gauss-Newton method to the bilevel problem and the MPEC under Cournot competition. The Gauss Newton method for a smooth optimization problem solves a series of approximations of the problem, relying on first order information on the objective function. It is especially well suited for problems whose objective is the squared norm of a vector, as is the case in the estimation problem. For the Cournot problem in disequilibrium and equilibrium, the objective function is directionally differentiable only, since the best-response (resp. Nash equilibrium) function is directionally differentiable in the parameter. Hence, each approximated problem requires the computation of the directional derivative of the best-response (resp. equilibrium) function. As a result, the approximated problem is a bilevel problem (resp. MPEC) in disequilibrium (resp. in equilibrium) where the upper level objective is the approximation based on first order information of the estimation function, and the lower level problem is the computation of the directional derivative of the best-response (resp. Nash equilibrium) function.

Overview of the Gauss-Newton Method for Nonlinear Programs

The Gauss-Newton method is designed to solve problems of the form:

$$\min_{\mathbf{x} \in \mathcal{X}} \frac{1}{2} \|\mathbf{f}(\mathbf{x})\|^2$$

where $\mathbf{f}$ is a continuously differentiable vector-valued function. It is an iterative procedure, which works as follows: at iteration k, we start with the current value

of the iterate $\mathbf{x}^k$, and solve an approximation of the original problem, called the *direction finding problem*, to find a direction $\mathbf{d}^k$. We then compute the next iterate $\mathbf{x}^{k+1} = \mathbf{x}^k + t^k\mathbf{d}^k$, where t^k is some step size. The iteration is repeated until a convergence criterion is reached.
The approximation of the original problem is the following. At iteration k, let us take the first order Taylor expansion of $\mathbf{f}$ around $\mathbf{x}^k$:

$$\begin{aligned}&\mathbf{f}(\mathbf{x})\\ &= \mathbf{f}(\mathbf{x}^k) + \nabla\mathbf{f}(\mathbf{x}^k)(\mathbf{x}-\mathbf{x}^k) + o(\|\mathbf{x}-\mathbf{x}^k\|)\end{aligned}$$

where $o(x)$ is a function such that $\lim_{x\to 0}\frac{o(x)}{x} = 0$ and $\nabla\mathbf{f}(\mathbf{x}^k)$ is the jacobian matrix of $\mathbf{f}$.
Hence, the squared norm $\|\mathbf{f}(\mathbf{x})\|^2$ is approximated by:

$$\begin{aligned}&\frac{1}{2}\|\mathbf{f}(\mathbf{x})\|^2\\ &= \frac{1}{2}\|\mathbf{f}(\mathbf{x}^k)\|^2 + \mathbf{f}(\mathbf{x}^k)'\nabla\mathbf{f}(\mathbf{x}^k)(\mathbf{x}-\mathbf{x}^k)\\ &+ \frac{1}{2}(\mathbf{x}-\mathbf{x}^k)'[\nabla\mathbf{f}(\mathbf{x}^k)'\nabla\mathbf{f}(\mathbf{x}^k)](\mathbf{x}-\mathbf{x}^k) + \frac{1}{2}o(\|\mathbf{x}-\mathbf{x}^k\|)^2\end{aligned}$$

The Gauss-Newton method at iteration k tries to find the feasible direction $\mathbf{d}^k$ which minimizes the function:

$$\begin{aligned}\min_{\mathbf{d}} \quad & \mathbf{f}(\mathbf{x}^k)'\nabla\mathbf{f}(\mathbf{x}^k)\mathbf{d} + \tfrac{1}{2}\mathbf{d}'[\nabla\mathbf{f}(\mathbf{x}^k)'\nabla\mathbf{f}(\mathbf{x}^k)]\mathbf{d}\\ \text{s.t} \quad & \mathbf{x}^k + \mathbf{d} \in \mathcal{X}\end{aligned}$$

The above problem has a convex quadratic objective function (the matrix in the quadratic term is indeed positive semi definite) and is therefore easily solvable. For instance, when $\mathcal{X} = \mathbb{R}^n$ is the euclidian space, and when the matrix $[\nabla\mathbf{f}(\mathbf{x}^k)'\nabla\mathbf{f}(\mathbf{x}^k)]$ is positive definite, the closed form solution of the above problem is:

$$\mathbf{d}^k = -[\nabla\mathbf{f}(\mathbf{x}^k)'\nabla\mathbf{f}(\mathbf{x}^k)]^{-1}\nabla\mathbf{f}(\mathbf{x}^k)'\mathbf{f}(\mathbf{x}^k)$$

Then, after choosing a step size t^k, the next iterate is:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + t^k\mathbf{d}^k$$

The method does not usually converge in a finite number of iterations. Therefore, a convergence criterion is used to terminate the method. Typically, the method terminates when:

$$\|\nabla\mathbf{f}(\mathbf{x}^k)\| \leq \eta$$

In summary, the Gauss-Newton method works as follows:

Algorithm 7: Gauss-Newton method
Input: Initial value $\mathbf{x}^0$
Output: Optimal value $\mathbf{x}^* = \arg\min \|\mathbf{f}(\mathbf{x})\|^2$

(1) Initialize $\mathbf{x} \leftarrow \mathbf{x}^{(0)}$
(2) Initialize $k = 0$
(3) **repeat**
(4) Find direction $\mathbf{d}^k = \arg\min \mathbf{f}(\mathbf{x}^k)'\nabla\mathbf{f}(\mathbf{x}^k)\mathbf{d} + \frac{1}{2}\mathbf{d}'[\nabla\mathbf{f}(\mathbf{x}^k)'\nabla\mathbf{f}(\mathbf{x}^k)]\mathbf{d}$
(5) Find step size t^k
(6) $\tilde{\mathbf{x}}^k \leftarrow \mathbf{x}^k + t^k\mathbf{d}^k$
(7) **if** $\|\nabla\mathbf{f}(\mathbf{x}^k)\| \leq \eta$
(8) **return** $\tilde{\mathbf{x}}^k$
(9) **else**
(10) $\mathbf{x}^k \leftarrow \tilde{\mathbf{x}}^k$, $k \leftarrow k+1$, and go to (4)

The Gauss-Newton method is known to converge geometrically. See for instance Bertsekas [12] for a proof under the Armijo step size rule.

Adaptation of the Gauss-Newton Method to the Cournot Bilevel Program

Recall that the estimation problem corresponding to the joint allocation policy with learning in disequilibrium has the following objective:

$$\sum_{h=-H}^{-1}\{\sum_{s=1}^{T}\left(\widehat{q_i}^h(s) - \mathcal{Q}_i(s, \mathbf{p}_{-\mathbf{i}}(\mathbf{h}), \alpha_\mathbf{i})\right)^2\} + \sum_{s=1}^{T}\left(\widehat{q_i}^0(s) - \mathcal{Q}_i(s, \overline{\mathbf{q}_{-\mathbf{i}}}, \alpha_\mathbf{i})\right)^2$$

It is therefore the squared norm of a vector with components $\widehat{q_i}^h(s) - \mathcal{Q}_i(s, \mathbf{p}_{-\mathbf{i}}(\mathbf{h}), \alpha_\mathbf{i})$ and $\widehat{q_i}^0(s) - \mathcal{Q}_i(s, \overline{\mathbf{q}_{-\mathbf{i}}}, \alpha_\mathbf{i})$. Let us denote by $\mathbf{F}_i(\alpha_\mathbf{i}, \mathbf{q}_\mathbf{i})$ the vector with components $\widehat{q_i}^h(s) - q_i(s)$ and $\widehat{q_i}^0(s) - q_i(s)$. Hence, the estimation function is the minimization of the squared norm of the vector:

$$\mathbf{F}_\mathbf{i}(\alpha_\mathbf{i}, \mathcal{Q}_i(\overline{\mathbf{q}_{-\mathbf{i}}}, \alpha_\mathbf{i}))$$

Let us denote by $\tilde{\mathbf{F}}_\mathbf{i}(\alpha_\mathbf{i}) = \mathbf{F}_\mathbf{i}(\alpha_\mathbf{i}, \mathcal{Q}_i(\overline{\mathbf{q}_{-\mathbf{i}}}, \alpha_\mathbf{i}))$. Since $\mathcal{Q}_i(\overline{\mathbf{q}_{-\mathbf{i}}}, \alpha_\mathbf{i})$ is a Bouligand differentiable function of $\alpha_\mathbf{i}$, so is $\tilde{\mathbf{F}}_\mathbf{i}$. As a result, it is directionally differentiable, the directional derivative being:

$$\tilde{\mathbf{F}}_\mathbf{i}'(\alpha_\mathbf{i}; \mathbf{d}_\mathbf{i}) = \nabla_{\alpha_i}\mathbf{F}_\mathbf{i}(\alpha_\mathbf{i}, \mathcal{Q}_i(\overline{\mathbf{q}_{-\mathbf{i}}}, \alpha_\mathbf{i}))\mathbf{d}_\mathbf{i} + \nabla_{q_i}\mathbf{F}_\mathbf{i}(\alpha_\mathbf{i}, \mathcal{Q}_i(\overline{\mathbf{q}_{-\mathbf{i}}}, \alpha_\mathbf{i}))\mathcal{Q}_i'(\overline{\mathbf{q}_{-\mathbf{i}}}, \alpha_\mathbf{i}; \mathbf{d}_\mathbf{i})$$

where $\mathcal{Q}_i'(\overline{\mathbf{q}_{-\mathbf{i}}}, \alpha_\mathbf{i}; \mathbf{d}_\mathbf{i})$ is the directional derivative of the best-response function.
As seen in Chapter 3, the directional derivative of the best-response function is solution to the convex quadratic optimization problem $\mathcal{QP}(\alpha_\mathbf{i}; \mathbf{d}_\mathbf{i})$.
Therefore, each iteration k of the Gauss-Newton method applied to the joint dynamic allocation with learning problem can be decomposed into the following two steps:

1. Assuming the direction ${\mathbf{d}_\mathbf{i}}^k$ is given, find the directional derivative of the best-response function by solving $\mathcal{QP}({\alpha_\mathbf{i}}^k; {\mathbf{d}_\mathbf{i}}^k)$;

2. Find the direction ${\mathbf{d_i}}^k$ by solving the direction finding problem:

$$\min_{\mathbf{d_i}} \tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\tilde{\mathbf{F}}'_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k;\mathbf{d_i}) + \frac{1}{2}[\tilde{\mathbf{F}}'_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k;\mathbf{d_i})'\tilde{\mathbf{F}}'_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k;\mathbf{d_i})]$$

$\tilde{\mathbf{F}}'_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k;\mathbf{d_i})$ is given by:

$$\tilde{\mathbf{F}}'_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k;\mathbf{d_i}) = \nabla_{\alpha_i}\mathbf{F_i}({\alpha_{\mathbf{i}}}^k,\mathcal{Q}_i^k)\mathbf{d_i} + \nabla_{q_i}\mathbf{F_i}({\alpha_{\mathbf{i}}}^k,\mathcal{Q}_i^k)\mathcal{Q}'_i(\overline{\mathbf{q_{-i}}},{\alpha_{\mathbf{i}}}^k;\mathbf{d_i})$$

where $\mathcal{Q}'_i(\overline{\mathbf{q_{-i}}},{\alpha_{\mathbf{i}}}^k;\mathbf{d_i})$ is the result of Step 1 and $\mathcal{Q}_i^k = \mathcal{Q}_i(\overline{\mathbf{q_{-i}}},{\alpha_{\mathbf{i}}}^k)$.

We see that Step 2 takes the result of Step 1 as input. Consequently, Step 2 is the upper level of a bilevel problem, whose lower level is Step 1. Since both the upper and lower level problem are convex quadratic problems, each iteration of the Gauss-Newton method applied to the joint dynamic allocation with learning problem is a quadratic bilevel problem.

$$\begin{aligned}
\min_{\mathbf{d_i}} \quad & \tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\tilde{\mathbf{F}}'_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k;\mathbf{d_i}) + \tfrac{1}{2}[\tilde{\mathbf{F}}'_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k;\mathbf{d_i})'\tilde{\mathbf{F}}'_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k;\mathbf{d_i})] \\
\text{s.t} \quad & \tilde{\mathbf{F}}'_{\mathbf{i}}(\alpha_{\mathbf{i}};\mathbf{d_i}) = \nabla_{\alpha_i}\mathbf{F_i}({\alpha_{\mathbf{i}}}^k,\mathcal{Q}_i^k)\mathbf{d_i} + \nabla_{q_i}\mathbf{F_i}({\alpha_{\mathbf{i}}}^k,\mathcal{Q}_i^k)\mathcal{Q}'_i(\overline{\mathbf{q_{-i}}},{\alpha_{\mathbf{i}}}^k;\mathbf{d_i}) \\
& \mathcal{Q}'_i(\overline{\mathbf{q_{-i}}},{\alpha_{\mathbf{i}}}^k;\mathbf{d_i}) = \arg\min_{\mathbf{x_i}} -\tfrac{1}{2}\mathbf{x_i}'\nabla^2_{i,i}\pi_i(\mathcal{Q}_i^k,\overline{\mathbf{p_{-i}}},{\alpha_{\mathbf{i}}}^k)\mathbf{x_i} + \mathbf{x_i}'\nabla^2_{i,\alpha_i}\pi_i(\mathcal{Q}_i^k,\overline{\mathbf{p_{-i}}},{\alpha_{\mathbf{i}}}^k)\mathbf{d_i} \\
\text{s.t} \quad & \mathbf{x_i} \in \mathfrak{Q}_i\perp
\end{aligned} \tag{6.3}$$

where $\mathfrak{Q}_i\perp$ is the following polyhedral cone ($\mathcal{T}^k = \{t : \mathcal{Q}_i^k(t) = 0\}$):

- If the capacity constraint is tight, i.e $\sum_{t=1}^T \mathcal{Q}_i^k(t) = C_i$:

$$\mathfrak{Q}_i\perp = \{\mathbf{x_i} \in \mathbb{R}^T \ : \ \nabla_i\pi_i(\mathcal{Q}_i^k,\overline{\mathbf{q_{-i}}},{\alpha_{\mathbf{i}}}^k)'\mathbf{x_i} = 0, \quad \sum_{t=1}^T x_i(t) \le 0, \quad \forall\ t \in \mathcal{T}^k,\ x_i(t) \ge 0\}$$

- If the capacity constraint is not tight, i.e $\sum_{t=1}^T \mathcal{Q}_i^k(t) < C_i$:

$$\mathfrak{Q}_i\perp = \{\mathbf{x_i} \in \mathbb{R}^T \ : \ \nabla_i\pi_i(\mathcal{Q}_i^k,\overline{\mathbf{q_{-i}}},{\alpha_{\mathbf{i}}}^k)'\mathbf{x_i} = 0, \quad \forall\ t \in \mathcal{T}^k,\ x_i(t) \ge 0\}$$

Hence, the lower level quadratic problem can be replaced by its equivalent KKT conditions:

- If the capacity constraint is tight at $\mathcal{Q}_i^k$, the lower level quadratic problem is equivalent to the following KKT conditions:

$$-\nabla^2_{i,i}\pi_i(\mathcal{Q}_i^k,\overline{\mathbf{p_{-i}}},{\alpha_{\mathbf{i}}}^k)\mathbf{x_i} + \nabla^2_{i,\alpha_i}\pi_i(\mathcal{Q}_i^k,\overline{\mathbf{p_{-i}}},{\alpha_{\mathbf{i}}}^k)\mathbf{d_i} + \lambda_i\mathbf{e} - \sum_{t\in\mathcal{T}^k}\mu_i(t)\mathbf{e}(t) = 0$$

$$\nabla_i \pi_i(\mathcal{Q}_i^k, \overline{\mathbf{q_{-i}}}, {\alpha_\mathbf{i}}^k)' \mathbf{x_i} = 0$$
$$\lambda_i(\sum_{t=1}^{T} x_i(t)) = 0$$
$$\mu_i(t) x_i(t) = 0 \ \ \forall\, t \in \mathcal{T}^k$$
$$\sum_{t=1}^{T} x_i(t) \leq 0,\ x_i(t) \geq 0 \ \ \forall\, t \in \mathcal{T}^k$$

- If the capacity constraint is tight at $\mathcal{Q}_i^k$, the lower level quadratic problem is equivalent to the following KKT conditions:

$$-\nabla^2_{i,i}\pi_i(\mathcal{Q}_i^k, \overline{\mathbf{p_{-i}}}, {\alpha_\mathbf{i}}^k)\mathbf{x_i} + \nabla^2_{i,\alpha_i}\pi_i(\mathcal{Q}_i^k, \overline{\mathbf{p_{-i}}}, {\alpha_\mathbf{i}}^k)\mathbf{d_i} - \sum_{t \in \mathcal{T}^k} \mu_i(t)\mathbf{e}(t) = 0$$
$$\nabla_i \pi_i(\mathcal{Q}_i^k, \overline{\mathbf{q_{-i}}}, {\alpha_\mathbf{i}}^k)' \mathbf{x_i} = 0$$
$$\mu_i(t) x_i(t) = 0 \ \ \forall\, t \in \mathcal{T}^k$$
$$x_i(t) \geq 0 \ \ \forall\, t \in \mathcal{T}^k$$

The constraints are linear, except for the complementary constraints. Since there are at most N such constraints, one can easily enumerate all the cases.
As a result, each Gauss-Newton iteration is equivalent to solving a series of single-level optimization problems with quadratic objective function and linear constraints.

Convergence of the Gauss-Newton Method Applied to the Cournot Bilevel Problem

For the sake of brevity, we make some simplifying assumptions which are without loss of generality.

Assumptions 6.1.
- *We assume that the capacity constraint in the Cournot best-response problem is an equality constraint: $\sum_{t=1}^{T} q_i(t) = C_i$;*
- *We assume that the set of feasible parameters is the whole euclidian space: $\mathcal{A}_i = \mathbb{R}^{(N+1)T}$.*

Under these assumptions, the direction-finding problem is an unconstrained problem and can therefore be solved in closed-form. Convergence of the method is proved in two steps. We first prove that the direction computed at each iteration of the Gauss-Newton method by the direction finding problem is a direction of descent, i.e of cost decrease. We then prove that the sequence of iterates computed by the Gauss-Newton method converges to a stationary point.
Without Assumptions 6.1, the direction-finding problem becomes a constrained optimization problem. It can easily be shown that the problem is equivalent to its KKT conditions. Hence, the Gauss-Newton direction is the solution to the KKT system of equalities and inequalities. Convergence of the method is proved in two steps: first,

we prove that the Gauss-Newton direction which solves the KKT conditions induces a cost decrease, and second that the sequence of iterates generated by the method converges to a stationary point.

We now state the convergence result of the Gauss-Newton method under Assumptions 6.1.

Proposition 1. *Let ${\alpha_{\mathbf{i}}}^k$ be the current iterate of the Gauss-Newton method. Let ${\mathbf{d_i}}^k$ be the Gauss-Newton direction. If ${\mathbf{d_i}}^k \neq 0$ then ${\mathbf{d_i}}^k$ is a direction of cost decrease, i.e there is $\bar{\tau} > 0$ such that for all $0 \leq \tau \leq \bar{\tau}$:*

$$\|\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k + \tau {\mathbf{d_i}}^k)\|^2 < \|\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)\|^2$$

Proof. Since $\mathcal{Q}_i$ is PC1, recall that for each $\alpha_{\mathbf{i}}$, there is m such that $\mathcal{Q}_i(\alpha_{\mathbf{i}}) = \psi_m(\alpha_{\mathbf{i}})$, where ψ_m is continuously differentiable. Under Assumption 6.1, m is the same for all $\alpha_{\mathbf{i}}$.

As a result, $\tilde{\mathbf{F}}_{\mathbf{i}}(.)$ is continuously differentiable, and we have:

$$\begin{aligned}
&\tilde{\mathbf{F}}'_{\mathbf{i}}(\alpha_{\mathbf{i}}; \mathbf{d_i}) \\
&= \nabla \tilde{\mathbf{F}}_{\mathbf{i}}(\alpha_{\mathbf{i}})\mathbf{d_i}) \\
&= \nabla_{\alpha_i}\mathbf{F}_{\mathbf{i}}(\alpha_{\mathbf{i}}, \mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}))\mathbf{d_i} + \nabla_{q_i}\mathbf{F}_{\mathbf{i}}(\alpha_{\mathbf{i}}, \mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}))\nabla\psi_m(\overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}})\mathbf{d_i}
\end{aligned}$$

In other words, the Gauss-Newton direction is:

$${\mathbf{d_i}}^k = -[\mathbf{D}^k]^{-1}\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)$$

where the matrix $[\mathbf{D}^k]$ is defined as follows. Let $[\mathbf{H}^k] = [\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)]$:

$$[\mathbf{D}^k] = \begin{cases} [\mathbf{H}^k] & \text{if it is invertible} \\ \mathbf{H}^k] + [\triangle^k] + [\triangle^k] & \text{otherwise} \end{cases}$$

$[\triangle^k]$ is a diagonal matrix such that $[\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)] + [\triangle^k]$ is invertible.

In what follows, we assume without loss of generality that $[\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)]$ is invertible.

Let $\tau > 0$:

$$\begin{aligned}
&\frac{1}{2}\|\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k + \tau{\mathbf{d_i}}^k)\|^2 \\
&= \frac{1}{2}\|\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k) + \tau\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k){\mathbf{d_i}}^k + \tau o(\|{\mathbf{d_i}}^k\|)\|^2 \\
&= \frac{1}{2}\|\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)\|^2 + \tau\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k){\mathbf{d_i}}^k + \tau o(\|{\mathbf{d_i}}^k\|) \\
&= \frac{1}{2}\|\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)\|^2 - \tau\left(\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)\right)[\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)]^{-1}\left(\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)\right) \\
&+ \tau o(\|{\mathbf{d_i}}^k\|)
\end{aligned}$$

The first equality uses the first order Taylor series expansion of $\tilde{\mathbf{F}}_{\mathbf{i}}$. The second equality uses bilinearity of the norm operator. In the third equality, ${\mathbf{d_i}}^k$ was replaced by its value.

The matrix $[\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)]$ is positive semi definite, and is assumed invertible without loss of generality. Hence, $[\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)'\nabla\tilde{\mathbf{F}}_{\mathbf{i}}({\alpha_{\mathbf{i}}}^k)]^{-1}$ is positive definite. Thus for any non zero vector $\mathbf{x}$, $\mathbf{x}'[\nabla\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)'\nabla\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)]^{-1}\mathbf{x} > 0$ and it is equal to zero iff $\mathbf{x} = \mathbf{0}$. As a result, $\left(\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)'\nabla\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)\right)[\nabla\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)'\nabla\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)]^{-1}\left(\nabla\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)'\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)\right) > 0$ unless the vector $\nabla\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)'\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k) = 0$. But this would imply that ${\mathbf{d}_\mathbf{i}}^k = 0$. Hence, if ${\mathbf{d}_\mathbf{i}}^k \neq 0$, then there is $\bar{\tau} > 0$ such that for all $0 \leq \tau \leq \bar{\tau}$, we have:

$$\frac{1}{2}\|\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k + \tau{\mathbf{d}_\mathbf{i}}^k)\|^2 < \frac{1}{2}\|\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k)\|^2$$

In conclusion, if the Gauss-Newton direction ${\mathbf{d}_\mathbf{i}}^k$ is non zero, then it is a direction of cost decrease. □

Theorem 6.3. *Assume that the feasible set $\mathcal{A}_i$ is closed bounded.*
Let $\{{\alpha_\mathbf{i}}^k\}$ be the sequence of iterates generated by the Gauss-Newton method, and $\{{\mathbf{d}_\mathbf{i}}^k\}$ the corresponding sequence of Gauss-Newton directions.
The sequence $\{{\alpha_\mathbf{i}}^k\}$ converges to a stationary point, i.e any limit point $\bar{\alpha}_\mathbf{i}$ of the sequence is such that:

$$\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})'\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i}) = 0$$

Proof. Since $\mathcal{A}_i$ is compact, then by Weierstrass' s theorem, $\{{\alpha_\mathbf{i}}^k\}$ has at least one converging subsequence. Without loss of generality, we assume that the entire sequence $\{{\alpha_\mathbf{i}}^k\}$ converges. Let us denote by $\bar{\alpha}_\mathbf{i}$ its limit and $\bar{\mathbf{d}}_\mathbf{i}$ the corresponding Gauss-Newton direction, i.e:

$$\bar{\mathbf{d}}_\mathbf{i} = -[\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})'\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})]^{-1}\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})'\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})$$

Suppose that the limit is not a stationary point, i.e $\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})'\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i}) \neq 0$.
Since the matrix $[\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})'\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})]^{-1}$ is positive definite, then:

$$\left(\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})'\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})\right)'[\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})'\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})]^{-1}\nabla\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})'\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i}) > 0$$

As a result, similarly to the proof of Proposition .., there is a $\bar{\tau} > 0$ such that for all $0 \leq \tau \leq \bar{\tau}$, we have $\bar{\alpha}_\mathbf{i} + \tau\bar{\mathbf{d}}_\mathbf{i} \in \mathcal{A}_i$ and:

$$\xi = \|\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})\|^2 - \|\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i} + \tau\bar{\mathbf{d}}_\mathbf{i})\|^2 > 0$$

Since $\lim_{k\to\infty}({\alpha_\mathbf{i}}^k + \tau{\mathbf{d}_\mathbf{i}}^k) = \bar{\alpha}_\mathbf{i} + \tau\bar{\mathbf{d}}_\mathbf{i}$ and $\tilde{\mathbf{F}}_\mathbf{i}$ is continuous, then we have:

$$\lim_{k\to\infty}\|\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k + \tau{\mathbf{d}_\mathbf{i}}^k)\|^2 = \|\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i} + \tau\bar{\mathbf{d}}_\mathbf{i})\|^2$$

As a result, for k sufficiently large, we have:

$$\|\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k + \tau{\mathbf{d}_\mathbf{i}}^k)\|^2 \leq \|\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i} + \tau\bar{\mathbf{d}}_\mathbf{i})\|^2 + \frac{\xi}{2} = \|\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})\|^2 - \frac{\xi}{2}$$

By definition of the optimal step size t^k, we have ${\alpha_\mathbf{i}}^k + t^k{\mathbf{d}_\mathbf{i}}^k = {\alpha_\mathbf{i}}^{k+1}$, we therefore have:

$$\begin{aligned}
&\|\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})\|^2 \\
&\quad < \|\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k + t^k{\mathbf{d}_\mathbf{i}}^k)\|^2 \\
&\quad \leq \|\tilde{\mathbf{F}}_\mathbf{i}({\alpha_\mathbf{i}}^k + \tau{\mathbf{d}_\mathbf{i}}^k)\|^2 \\
&\quad \leq \|\tilde{\mathbf{F}}_\mathbf{i}(\bar{\alpha}_\mathbf{i})\|^2 - \frac{\xi}{2}
\end{aligned}$$

This contradicts the optimality of $\bar{\alpha}_{\mathbf{i}}$. $\square$

6.2.3 Discussion of the Gauss-Newton Method for the Cournot and Bertrand Equilibrium Problems

We now discuss briefly how the above method can be adapted to solve the Cournot and Bertrand equilibrium problems.

Gauss-Newton method for the Cournot Equilibrium Problem

In the Cournot equilibrium problem, the lower level problem is the variational inequality equivalent to the Nash equilibrium problem. The upper level takes the form of the squared norm of a vector-valued function:

$$\sum_{i=1}^{N}\left\{\sum_{h=-H}^{-1}\{\sum_{s=1}^{T}(\widehat{q_i}^h(s)-\mathcal{Q}_i(s,\alpha_{\mathbf{i}},\alpha_{-\mathbf{i}}))^2\}+\sum_{s=1}^{T}(\widehat{q_i}^0(s)-\mathcal{Q}_i(s,\alpha_{\mathbf{i}},\alpha_{-\mathbf{i}}))^2\right\}$$

Let us denote by $\mathbf{F}(\alpha,\mathcal{Q}(\alpha))$ the vector-valued function with components $(\widehat{q_i}^h(s)-\mathcal{Q}_i(s,\alpha_{\mathbf{i}},\alpha_{-\mathbf{i}}))$ and $(\widehat{q_i}^0(s)-\mathcal{Q}_i(s,\alpha_{\mathbf{i}},\alpha_{-\mathbf{i}}))$. Hence, the estimation problem in equilibrium state is:

$$\min_{\alpha\in\alpha}\frac{1}{2}\|\mathbf{F}(\alpha,\mathcal{Q}(\alpha))\|^2$$

Let us denote by $\tilde{\mathbf{F}}(\alpha)=\mathbf{F}(\alpha,\mathcal{Q}(\alpha))$.
As a result, we may apply the Gauss-Newton method, provided that we can compute directional derivatives for the Nash equilibrium quantity function. We have established in Proposition 3.23 that the Nash equilibrium quantities as functions of the parameters are Bouligand differentiable and the directional derivative $\mathcal{Q}'(\alpha;\mathbf{d})$ are solutions to the affine variational inequality given in 3.23.
Therefore, each iteration k of the Gauss-Newton method applied to the joint dynamic allocation with learning problem can be decomposed into the following two steps:

1. Assuming the direction $\mathbf{d}^k$ is given, find the directional derivative of the best-response function by solving the affine VI 3.23;

2. Find the direction $\mathbf{d}^k$ by solving the direction finding problem:

$$\min_{\mathbf{d}}\tilde{\mathbf{F}}(\alpha^k)'\tilde{\mathbf{F}}'(\alpha^k;\mathbf{d})+\frac{1}{2}[\tilde{\mathbf{F}}'(\alpha^k;\mathbf{d})'\tilde{\mathbf{F}}'(\alpha^k;\mathbf{d})]$$

$\tilde{\mathbf{F}}'(\alpha^k;\mathbf{d})$ is given by:

$$\tilde{\mathbf{F}}'(\alpha^k;\mathbf{d})=\nabla_\alpha\mathbf{F}(\alpha^k,\mathcal{Q}^k)\mathbf{d}+\nabla_q\mathbf{F}(\alpha^k,\mathcal{Q}^k)\mathcal{Q}'(\alpha^k;\mathbf{d})$$

where $\mathcal{Q}'(\alpha^k;\mathbf{d})$ is the result of Step 1 and $\mathcal{Q}^k=\mathcal{Q}(\alpha^k)$.

Since Step 2 takes the result of Step 1 as input, we conclude that Steps 1 and 2 together form an MPEC, with a convex quadratic upper level problem, and a linear lower level VI.

$$\begin{array}{ll} \min_{\mathbf{d}} & \tilde{\mathbf{F}}(\alpha^k)'\tilde{\mathbf{F}}'(\alpha^k;\mathbf{d}) + \frac{1}{2}[\tilde{\mathbf{F}}'(\alpha^k;\mathbf{d})'\tilde{\mathbf{F}}'(\alpha^k;\mathbf{d})] \\ \text{s.t} & \tilde{\mathbf{F}}'(\alpha^k;\mathbf{d}) = \nabla_\alpha \mathbf{F}(\alpha^k,\mathcal{Q}^k)\mathbf{d} + \nabla_q \mathbf{F}(\alpha^k,\mathcal{Q}^k)\mathcal{Q}'(\alpha^k;\mathbf{d}) \\ & \forall\ \mathbf{x} \in \mathfrak{Q}\perp : \\ & \left(\nabla_{\mathbf{q}}\mathbf{F}(\bar{\mathcal{Q}},\alpha)\mathcal{Q}'(\alpha^k;\mathbf{d}) + \nabla_\alpha \mathbf{F}(\bar{\mathcal{Q}},\alpha)\mathbf{d}\right)'(\mathbf{x} - \mathcal{Q}'(\alpha^k;\mathbf{d})) \geq 0 \end{array}$$

Similarly to Chapter 4, we can reformulate the above MPEC by replacing the lower-level variational inequalities by their equivalent first order optimality conditions. Then, by reformulating the complementary constraints using binary variables, we obtain an equivalent mixed integer program with convex quadratic upper-level and linear lower-level.
As a result, we need to solve at each Gauss-Newton iteration mixed integer program with convex quadratic upper-level and linear lower-level.This class of problems is known as Mixed Integer Quadratic Programs (MIQP) and has been the subject to a lot of research.

Gauss-Newton Method for the Bertrand Equilibrium Problem

The estimation problem in the Bertrand equilibrium problem is also the minimization of the squared norm of a vector-valued function that we can denote by $\mathbf{F}(\beta, \mathcal{P}(\beta))$, where $\mathcal{P}(\beta)$ is the generalized equilibrium price. Hence, provided that this function has some differentiability properties, the Gauss-Newton method can be applied to the Bertrand problem in equilibrium. The difficulty stems from the fact that that to our knowledge, there are no known sensitivity results concerning generalized equilibria.
We have nevertheless established in Section 1 above that the generalized Nash equilibrium can be obtained as the limit of a sequence of equilibria corresponding to a penalized version of the generalized Nash equilibrium. We can establish as in Chapter 3 for the Cournot equilibrium, that each penalized equilibrium is Bouligand differentiable, with directional derivative obtained as the solution to an affine variational inequality.
As a result, it appears that if used together, the Gauss-Newton method and the penalty scheme can be applied to the Bertrand problem in equilibrium in order to solve the corresponding MPEC.

6.3 Implementation of the Approach

We have implemented the approach on an IBM Thinkpad, with a processor of 1.6 GHz, and 256MB or RAM, using Matlab 7.0. In order for the method to be implemented on a computer, several technical issues must be addressed. We review these technical issues in Subsection 6.3.1. Subsection 6.3.2 discusses the results obtained on various examples.

6.3.1 Challenges of the Implementation

Evaluation of the Learning

In Chapter 4, we showed that the learning approach is a contraction mapping, i.e the norm of the difference between two parameter vectors at consecutive time periods $\|\beta^{(m)} - \beta^{(m-1)}\|$ (resp. $\|\alpha^{(m)} - \alpha^{(m-1)}\|$) in the approach shrinks as time increases. In order to evaluate the learning of the parameters, we can track the evolution of the norm differences $\|\beta^{(m)} - \beta^{(m-1)}\|$ (resp. $\|\alpha^{(m)} - \alpha^{(m-1)}\|$) as the number of iterations m increases: learning of the parameters of the price-demand relationship is achieved if the norm differences decrease throughout the approach and converge to 0 at the end of the approach.
Similarly, the approach is a contraction mapping for the prices (resp. quantities): the norm differences $\|\mathbf{p}^{(m)} - \mathbf{p}^{(m-1)}\|$ (resp. $\|\mathbf{q}^{(m)} - \mathbf{q}^{(m-1)}\|$) shrink throughout the approach, and converge to 0. Hence, in order to evaluate the learning of the equilibrium prices (resp. quantities), one can track the evolution of $\|\mathbf{p}^{(m)} - \mathbf{p}^{(m-1)}\|$ (resp. $\|\mathbf{q}^{(m)} - \mathbf{q}^{(m-1)}\|$) as the number m of iterations increases.

Historical Data

In disequilibrium stage, each firm computes its best-response to its belief concerning its competitors's pricing policy. More precisely, we have made the realistic assumption that firms form beliefs regarding their competitors' strategy using historical data from past selling horizons.
Furthermore, since the parameters of the price-demand relationship differ at each time period of the approach, one must address the issue of potential over-fitting in the estimation problem. Indeed, one must ensure that there are enough data available at each period for each firm to compute the parameters of its demand (resp. price) function. For instance, in the linear, loglinear, and constant elasticity demand (resp. inverse demand) functions, the price-demand relationship at each period and for each firm is determined by $N+1$ parameters. As a result, we must ensure that the database of historical prices contains at least prices from $N+1$ past selling horizons.
The historical prices are generated according to the parametric model that we infer demand to follow. For instance, if we assume a linear demand model for the market, we generate historical prices as follows: let $\overline{d_i}^h(s)$, $i = 1, \ldots, N$, $s = 1, \ldots, T$ be a random sequence of demands such that $\overline{d_i}(s) \geq 0$ and $\sum_{s=1}^{T} \widehat{d_i}(s) \leq C_i$. Let $\bar{\beta} = (\overline{\beta_{ij}}(s))$, $i = 1, \ldots, N$, $j = 0, \ldots, N$, $s = 1, \ldots, T$ be a fixed vector of price-demand parameters. We first generate prices $\overline{p_i}(s)$ verifying:

$$\overline{d_i}(s) = \overline{\beta_{i0}}(s) - \overline{\beta_{ii}}(s)\overline{p_i}(s) + \sum_{j \neq i} \overline{\beta_{ij}}(s)\overline{p_j}(s)$$

In other words, in vector notation:

$$\overline{\mathbf{d}}(s) = \overline{\beta_{\mathbf{0}}}(s) - [\overline{\beta}(s)]\overline{\mathbf{p}}(s)$$

where $[\overline{\beta_{\mathbf{i}}}(s)]$ is the matrix such that:

$$\begin{pmatrix} \beta_{11}(s) & \dots & -\beta_{1N}(s) \\ \vdots & \ddots & \vdots \\ -\beta_{N1}(s) & \dots & \beta_{NN}(s) \end{pmatrix}$$

Hence, we generate prices such that:

$$\overline{\mathbf{p_i}}(s) = [\overline{\beta}(s)]^{-1}(\overline{\beta_{\mathbf{0}}}(s) - \overline{\mathbf{d}}(s))$$

The historical prices are then generated according to the stochastic model:

$$\widehat{p_i}^h(s) = \overline{\mathbf{p_i}}(s) + \epsilon_i(s)$$

where $\epsilon_i(s)$ are independent, identically distributed and follow the normal distribution with mean 0 and standard deviation σ_i.

6.3.2 Results of the Computations

The figures referred to below are to be found in part B of the appendix.

Computations in disequilibrium stage

- **An Airline Example**:
  Consider $N = 2$ airlines competing for a single leg. Market demand is heterogeneous: the leisure travelers tend to purchase early and are more sensitive to price, whereas business travelers tend to purchase closer to departure of the flight and are less sensitive to price. Since price sensitivity is decreasing over time, the parameter value decreases with time. Figure B-1 displays the values of the parameters as a function of time. Figures B-2, B-3, B-4 show respectively the norm differences $\|\beta^{(m)} - \beta^{(m-1)}\|$, $\|\mathbf{p}^{(m)} - \mathbf{p}^{(m-1)}\|$ and $\|\mathbf{p}^{(m)} - \mathbf{p}^{(m-1)}\|$ as functions of m for a selling horizon $T = 10$. We observe that the norm differences decrease as the end of the selling horizon approaches. Figures B-5, B-6, B-7 are the norm differences when the selling horizon is $T = 20$. They exhibit the same decreasing trend as the $T = 10$ case.
  Hence, the numerical results in disequilibrium state confirm the predictions of the mathematical analysis.

- **A Fashion Retail Example**
  Consider now $N = 2$ fashion retailers who dynamically price their inventory over a selling season of $T = 10$ time periods. Market demand is heterogeneous: customers who purchase early in the season tend to be less price sensitive than shoppers who wait until the end of the season for discounts. Here, the price sensitivity increases with time, and hence, so do the parameters. Figure B-8 shows the evolution of the parameters as a function of time. Figures B-9, B-10, B-11 show respectively the norm differences $\|\beta^{(m)} - \beta^{(m-1)}\|$, $\|\mathbf{p}^{(m)} - \mathbf{p}^{(m-1)}\|$ and $\|\mathbf{p}^{(m)} - \mathbf{p}^{(m-1)}\|$ as functions of m.

The results in the case of increasing price sensitivities over time are similar to those obtained in the airline example above, and corroborate the predictions of the mathematical analysis.

Computations in Equilibrium Setting

Here we investigate the airline example in equilibrium setting. The parameters are still those of Figure B-1. Figures B-12, B-13 and B-14 in appendix display the logarithm of the norm differences : $\log \|\beta^{(m)} - \beta^{(m-1)}\|$ and $\log \|\mathbf{p}^{(m)} - \mathbf{p}^{(m-1)}\|$. We show logarithmic differences because the convergence speed in equilibrium is much faster than in disequilibrium.

Evaluation of the Learning

As suggested by theory and supported by computational results, the following behavior was observed:

- In disequilibrium, learning is achieved for both the parameters and the pricing strategies in a relatively small number of time periods.

- Learning speed, as measured by the number of iterations required to satisfy the convergence criterion, is much faster in equilibrium than disequilibrium: it is achieved after just a few time periods. This is due to the fact that in equilibrium, only the parameters need to be learned.

- The learning speed is affected by the choice of historical prices. More precisely, the larger the standard deviation σ_i of the distribution according to which the prices were generated, the longer it takes to learn the parameters.

- The learning speed decreases as the number of firms in the market increases, all else being equal, as shown in Figure B-15: this is to be expected, since the more firms there are in the market, the more parameters there are to estimate at each time period.

- The learning speed decreases as the length of the horizon increases, all else being equal, as shown in Figure B-16: this is because the longer the horizon, the more parameters need to be estimated.

Behavior of the Algorithm

Numerical experience led us to the following conclusions regarding the practical convergence of the algorithm:

- At each time period of the approach, the number of iterations it takes for the Gauss-Newton method to converge increases as the size of the problem increases, i.e as the number of firms N, or the selling horizon T increases.

- The convergence of the algorithm is greatly affected by the choice of the starting point. This is because the problem at hand is highly non convex, and the method used only finds local solutions. Hence, to guarantee convergence, the method randomizes the starting point within the feasible region.

- At each time period of the approach, the algorithm takes much longer to converge in equilibrium than disequilibrium. This is because the lower level which needs to be solved repeatedly at each iteration is a variational inequality, whereas in disequilibrium, the lower level best-response problem can be solved in closed form.

- At each time period of the approach, the algorithm converges much faster when capacity is scarce. Indeed, in that case, firms can maximize each period's revenue independently of other periods, rendering the lower-level problem separable and much easier to solve, and the upper-level problem becomes separable as a result.

Chapter 7

Closed-Loop Dynamic Pricing

In previous chapters, we have considered open-loop policies: the policy is computed once and for all for the entire horizon. We now consider closed-loop policies. In a closed-loop setting, the firm postpones the decision of its optimal price or quantity for period t until the last possible moment, i.e at t. It can therefore take advantage of the information gathered up to time t, i.e the prices or quantities set by the firm and its competitors in past periods. In a deterministic setting, closed-loop and open-loop policies are equivalent. This means that there is no extra value in the additional information that can be gathered. Indeed, in deterministic problems, where no random disturbances can affect the market, the future behavior of the market is completely determined by its initial state. In the Bertrand and Cournot oligopoly models that were considered in previous chapters, the market is totally determined at equilibrium, once we know the total capacities of all the firms.
In this chapter, we introduce a stochastic component to the model. Therefore, the capacity available to the firm at each period is not known deterministically, and there is value in postponing the decision until additional information is available.
The model we present is the competitive counterpart to the model in Gallego and Van Ryzin [46]. They focus on a monopolistic market for a single perishable product, for which demand is a Poisson process with price dependent rate. For the exponential intensity function, they establish a closed-form solution for the optimal pricing policy and the corresponding revenue-to-go. They also study the structural properties of the optimal policy and revenue-to-go function, such as monotonicity and convexity properties. A similar model to ours was also introduced in Gallego and Hu [45], but the authors focused in their paper on open-loop strategies related to the model.
In this Chapter, we establish existence and uniqueness of Markov-perfect equilibrium strategies and establish closed-form solutions in the case of an exponential intensity function. We prove that unlike the monopoly model, monotonicity of the equilibrium policies does not hold. We then study approximations of the problem. We investigate a one-step look-ahead equilibrium policy with approximation of the revenue-to-go as a linear or quadratic function of the firms' capacities. We carry out computational experiment in order to compare the revenues and policies generated by the optimal policies, by the one-step look-ahead equilibrium policies, and by the open-loop feedback equilibrium policies. Finally, we investigate an iterative approach to converge

to the Markov-perfect equilibrium. This approach is of the best-response dynamics nature. We study its convergence computationally.

7.1 The Demand Model

We consider an oligopoly market, for single, perishable, substitutable products. The demand of a firm is modeled as a stochastic point process with Markovian intensity. In particular, we consider demand to be a nonhomogeneous Poisson process, with price-dependent intensity. Let $\lambda_i(p_i(s), p_{-i}(s), \alpha_{\mathbf{i}}(s))$ denote the intensity function of firm i at time s. It depends on the prices of the firms at time s, and is parameterized by a vector of parameters denoted $\alpha_{\mathbf{i}}(s)$.
We make the following assumptions on the intensities:

Assumptions 7.1. *1. The function λ_i is continuous in $\mathbf{p} = (p_i, \boldsymbol{p_{-i}})'$ and twice differentiable in p_i;*

2. *The vector-valued function $\boldsymbol{\lambda}(\mathbf{p})$ is an invertible function of $\mathbf{p}$;*

3. *λ_i is strictly decreasing in p_i, and non increasing in p_j, for $j \neq i$;*

4. *The instantaneous revenue function $r_i(\lambda_i, \boldsymbol{\lambda_{-i}}) = \lambda_i p_i(\lambda_i, \boldsymbol{\lambda_{-i}})$ is strictly concave in λ_i and bounded.*

5. *For all $\boldsymbol{\lambda_{-i}}$, there exists a bounded maximizer $\lambda_i^*(\boldsymbol{\lambda_{-i}}) = \arg\max_{\lambda_i} r_i(\lambda_i, \boldsymbol{\lambda_{-i}})$.*

6. *There is a null price $p_i^\infty \in \mathbb{R}_+ \bigcup \{\infty\}$ such that, for all $\boldsymbol{p_{-i}}$:*

$$\lim_{p_i \to p_i^\infty} \lambda_i(p_i, \boldsymbol{p_{-i}}) = 0$$

The set of feasible prices is defined as:

$$\mathcal{P}_i = [0, p_i^\infty]$$

Assumption (7.1.2) states that $\boldsymbol{\lambda}(\mathbf{p})$ is an invertible function of $\mathbf{p}$. Indeed, as is customary in revenue management (see for instance Gallego and Van Ryzin [46]), we will work with the inverse demand function $\mathbf{p}(\boldsymbol{\lambda})$. That is, we will formulate the problem as a competitive intensity control problem, and we will seek policies in terms of intensities. Assumption (7.1.3) restricts the problem to normal goods which are gross substitutes: the demand of firm i decreases when it increases its price, ceteris paribus, and the demand of firm i increases or stays the same when its competitor $-i$ increases its price. Assumption (7.1.4) states that the instantaneous revenue function is strictly concave in λ_i, i.e there are strictly diminishing returns. Assumption (7.1.5) guarantees the existence of a maximizer for the revenue function. As a result of Assumption (7.1.4) and Assumption (7.1.5), the revenue function is increasing in λ_i for $\lambda_i \leq \lambda_i^*(\boldsymbol{\lambda_{-i}})$, and decreasing for $\lambda_i \geq \lambda_i^*(\boldsymbol{\lambda_{-i}})$. Since p_i^∞ is in the set of feasible prices, then Assumption (7.1.6) implies that 0 is in the set of feasible intensities,

when the problem is expressed in variable λ_i, using the inverse rate function. This guarantees that the firm can set its intensity to 0 when it does not have any capacity left.
We denote by $\mathcal{L}_i$ the set of allowable intensities. Since $p_i \in [0, p_i^\infty]$ for all i, and by Assumption (7.1.2), we have that $\lim_{p_i \to p_i^\infty} \lambda_i(p_i, \mathbf{0}) \leq \lambda_i(p_i, \boldsymbol{p_{-i}}) \leq \lim_{\boldsymbol{p_{-i}} \to \boldsymbol{p_{-i}^\infty}} \lambda_i(0, \boldsymbol{p_{-i}})$. By Assumption (7.1.6), we know that the left hand side limit is equal to 0. As a result, $0 \in \mathcal{L}_i$.
We check that the above assumptions are realistic, by checking that they are satisfied by some of the intensity functions most used in the literature.

- Cobb-Douglas rate:

$$\lambda_i(p_i, \boldsymbol{p_{-i}}) = \theta_i \prod_{j=1}^{N} p_j^{-\beta_{ij}}$$

Let $[\boldsymbol{B}]$ be the matrix with coefficients β_{ij}. Let $[\boldsymbol{A}]$ be the inverse of $[\boldsymbol{B}]$ with coefficients α_{ij}. The inverse rate function is:

$$p_i(\lambda_i, \boldsymbol{\lambda_{-i}}) = \prod_{j=1}^{N} \left(\frac{\lambda_j}{\theta_j}\right)^{-\alpha_{ij}}$$

Assumptions (7.1) are satisfied iff: $[\boldsymbol{B}]$ is invertible, $\beta_{ii} > 0$, $\beta_{ij} \leq 0$ for $j \neq i$, $\alpha_{ii} > 0$. $p_i^\infty = \infty$.

- Exponential rate:

$$\lambda_i(p_i, \boldsymbol{p_{-i}}) = \theta_i \exp\{-\sum_{j=1}^{N} \beta_{ij} p_j\}$$

Let $[\boldsymbol{B}]$ be the matrix with coefficients β_{ij}. Let $[\boldsymbol{A}]$ be the inverse of $[\boldsymbol{B}]$ with coefficients α_{ij}. The inverse intensity function is:

$$p_i(\lambda_i, \boldsymbol{\lambda_{-i}}) = -\sum_{j=1}^{N} \alpha_{ij} \ln(\frac{\lambda_j}{\theta_j})$$

Assumptions (7.1) are satisfied iff: $[\boldsymbol{B}]$ is invertible, $\beta_{ii} > 0$, $\beta_{ij} \leq 0$ for $j \neq i$, $\alpha_{ii} > 0$. $p_i^\infty = \infty$.

- Linear rate:

$$\lambda_i(p_i, \boldsymbol{p_{-i}}) = \beta_{i0} - \sum_{j=1}^{N} \beta_{ij} p_j$$

Let $[\boldsymbol{B}]$ be the matrix with coefficients β_{ij}. Let $[\boldsymbol{A}]$ be the inverse of $[\boldsymbol{B}]$ with coefficients α_{ij}. Let $\boldsymbol{\beta_0}$ the vector of coefficients β_{i0}, and let $\boldsymbol{\alpha_0} = [\boldsymbol{B}]^{-1}\boldsymbol{\beta_0}$. The inverse intensity function is:

$$p_i(\lambda_i, \boldsymbol{\lambda_{-i}}) = \alpha_{i0} - \sum_{j=1}^{N} \alpha_{ij}\lambda_j$$

Assumptions (7.1) are satisfied iff: $[\boldsymbol{B}]$ is invertible, $\beta_{ii} > 0$, $\beta_{ij} \leq 0$ for $j \neq i$, $\alpha_{ii} > 0$. However, there does not exist a p_i^{∞} satisfying Assumption (7.1.6).

7.2 The Continuous-Time Problem

Each firm i has initial capacity denoted $C_i(0) \in \mathbb{N}$ at time 0. The selling horizon has length denoted t. Let $N_i(s)$ be the stochastic process corresponding to the number of units of demand for firm i up to time s. A unit of demand is realized at time s if $dN_i(s) = 1$. We denote by $\mathcal{P}(\boldsymbol{C})$ (resp. $\mathcal{L}(\boldsymbol{C})$) the space of allowable Markovian pricing (resp. intensity) policies for the firms, when initial capacities are $\boldsymbol{C} = (C_1, \ldots, C_N)$:

$$\forall\ i,\ \int_0^t dN_i(s) \leq C_i \quad \text{a.s}$$
$$\boldsymbol{p}(s) \in \mathcal{P} \ \Leftrightarrow\ \boldsymbol{\lambda}(s) \in \mathcal{L}$$

We denote by $\mathcal{P}_i(\boldsymbol{p_{-i}}, C_i)$ (resp. $\mathcal{L}_i(\boldsymbol{\lambda_{-i}}, C_i)$) be the space of allowable Markovian pricing (resp. intensity) policies for firm i, when its competitors use policy $\boldsymbol{p_{-i}}$ (resp. $\boldsymbol{\lambda_{-i}}$) and when its initial capacity is C_i. We consider closed-loop intensity strategies, that is, functions prescribing the optimal intensity the firm should set at any time s, as a function of the history of the process. The history should summarize all past information relevant to the decision making of the firm. The history at any time s is the set of all past intensities set by the firm and its competitors: $\{(\bar{\lambda}_1(s'), \ldots, \bar{\lambda}_N(s')), s' < s\}$.
In game theoretical terms, a policy is a mapping from histories to intensities, i.e:

$$\Lambda_i(s, \{\lambda_i(s'), \boldsymbol{\lambda_{-i}}(s')\}_{s'<s}) = \lambda_i(s)$$

Due to the Markovian property of the Poisson demand process, a policy does not depend on the entire history of the process, it only depends on the current capacity level of the firms, i.e $(C_1 - N_1(s), \ldots, C_N - N_N(s))$. We denote the current inventory level of each firm i by $C_i(s)$. Hence, in a Markovian system, the intensity policies at each time s are mappings from the current capacity levels to intensities:

$$\Lambda_i(s, C_i(s), \boldsymbol{C_{-i}}(s)) = \lambda_i(s)$$

The goal of each firm is to maximize its expected revenue over the selling horizon $[0, t]$. For feasible policies $(\lambda_1, \ldots, \lambda_N)$, and initial capacities $C_1, \ldots, C_N$, firm i's expected revenue over the selling horizon is $R_i(t, \lambda_i, \boldsymbol{\lambda_{-i}}, C_i, \boldsymbol{C_{-i}})$:

$$R_i(t, \lambda_i, \boldsymbol{\lambda_{-i}}, C_i, \boldsymbol{C_{-i}}) = \mathbb{E}\left\{\int_0^t p_i(s)dN_i(s)\right\}$$

In other words, the total expected revenue of firm i can be written as:

$$R_i(t, \lambda_i, \boldsymbol{\lambda_{-i}}, C_i, \boldsymbol{C_{-i}}) = \int_0^t \lambda_i(s) p_i(\lambda_i(s), \boldsymbol{\lambda_{-i}}(s)) ds$$

Hence, firm i's best-response problem, for fixed competitors's strategy $\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}$ is:

$$R_i^*(t, \lambda_i, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}, C_i, \boldsymbol{C_{-i}}) = \sup_{\lambda_i \in \mathcal{L}_i(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}, C_i)} R_i(t, \lambda_i, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}, C_i, \boldsymbol{C_{-i}})$$

A Markov-perfect equilibrium policy, is a set of policies for the firms, which are Nash equilibrium policies in every subgame. In other words, for any time s, and any current capacity levels $\mathbf{C}(s) = (C_1(s), \ldots, C_N(s))$, the policies $\Lambda_1^*(s, \mathbf{C}(s)), \ldots, \Lambda_N^*(s, \mathbf{C}(s))$ are Nash equilibrium policies for the game consisting of the firms maximizing their expected revenue from 0 to $s \leq t$. Let $R_i(s, \lambda_i, \boldsymbol{\lambda_{-i}}, C_i, \boldsymbol{C_{-i}})$ denote the expected revenue-to-go from 0 to s, when intensities $\lambda_i, \boldsymbol{\lambda_{-i}}$ are used. As a result, we have:

$$R_i(t, \lambda_i, \boldsymbol{\lambda_{-i}}, C_i, \boldsymbol{C_{-i}}) = \Pi_i(t, \lambda_i, \boldsymbol{\lambda_{-i}}, C_i, \boldsymbol{C_{-i}})$$

We define below the notion of Markov-perfect equilibrium formally.

Definition 1. *Policies $\Lambda_1^*(.), \ldots, \Lambda_N^*(.)$ are Markov-perfect equilibrium policies iff for all $i = 1, \ldots, N$, for all $C_1, \ldots, C_N$, and for all $0 \leq s \leq t$:*

$$R_i(s, \Lambda_i^*, \boldsymbol{\Lambda^*_{-i}}, C_i, \boldsymbol{C_{-i}}) \geq R_i(s, \lambda_i, \boldsymbol{\Lambda^*_{-i}}, C_i, \boldsymbol{C_{-i}}) \quad \forall \lambda_i$$

Note that the state space had dimension N, i.e the number of firms in the market.

7.2.1 Best-Response Problem

Let us focus on firm i, and let us fix its competitors' intensity policy to $\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}$. Given its competitors' policy, firm i seeks to maximize its total expected revenue over the selling horizon. As done in Gallego and van Ryzin [46], one can establish the Hamilton-Jacobi Bellman equation for the expected revenue-to-go.

Proposition 1. *Let $\mathcal{I}_0((C_1, \ldots, C_0) = \{i : C_i = 0\}$, and $\mathcal{I}_+((C_1, \ldots, C_0) = \{i : C_i \geq 1\}$. The best-response revenue-to-go of firm i, for fixed competitor' intensity policy $\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}$ is given by the Hamilton-Jacobi-Bellman (HJB) equation below.*

- *if $i \in \mathcal{I}_+((C_1, \ldots, C_0)$:*

$$\begin{aligned} \frac{\partial R_i(s, C_i, \boldsymbol{C_{-i}})}{\partial s} &= \\ &\max_{\lambda_i} \lambda_i \left(p_i(\lambda_i, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) + R_i(s, C_i - 1, \boldsymbol{C_{-i}}) - R_i(s, C_i, \boldsymbol{C_{-i}}) \right) \\ + &\sum_{j \neq i, j \in \mathcal{I}_+} \bar{\lambda}_j \left(R_i(s, C_i, \boldsymbol{C_{-i}} - \boldsymbol{e_j}) - R_i(s, C_i, \boldsymbol{C_{-i}}) \right) \end{aligned} \tag{7.1}$$

with boundary conditions $R_i(0, C_i, \boldsymbol{C_{-i}}) = 0$.

- *if* $i \in \mathcal{I}_0((C_1,\ldots,C_0)$*:* $R_i(0,\boldsymbol{C_{-i}})=0$

By convention, we set: $\Lambda_i(s,0,\boldsymbol{C_{-i}})=0$*.*

The proof relies on a discrete-time approximation of the problem.

Proof. Let δt be a time interval, such that at most one unit of demand occurs within δt. Let $t=K\delta t$. For any $0\le k\le K$, we call period $n-k$ the interval $[k\delta t,(k+1)\delta t)$. Let $(C_1^k,\ldots,C_N^k)$ be the state at period k. The index sets $\mathcal{I}_+^k,\mathcal{I}_0^k$ form a partition of $\{1,\ldots,N\}$. We denote by $\boldsymbol{C_+^k}$ the subvector of capacities $C_i^k\ge 1$, and $\boldsymbol{C_0^k}$ the subvector of capacities $C_i^k=0$. The state $(C_1^k,\ldots,C_N^k)=(\boldsymbol{C_+^k},\boldsymbol{C_0^k})$. Hence, the state evolution equation is:

$$(\boldsymbol{C_+^{k+1}},\boldsymbol{C_0^k})=\begin{cases}\forall\, i\in\mathcal{I}_+^K:(\boldsymbol{C_+^k}-\boldsymbol{e_i},\boldsymbol{C_0^k}) & \text{w.p } \lambda_i^k\delta t\\ (\boldsymbol{C_+^k},\boldsymbol{C_0^k}) & \text{w.p } 1-\sum_{j\in\mathcal{I}_+^k}\lambda_j^k\delta t\end{cases}$$

where $\boldsymbol{e_i}$ is the vector of components all zero except for component i which is equal to 1. Let us assume that $C_i^k\ge 1$. The backward induction equation for the revenue-to-go of firm i is therefore:

$$\begin{aligned}&R_i(k\delta t,C_i^k,\boldsymbol{C_{-i}^k},\boldsymbol{\bar\lambda_{-i}})=\\&\quad\max_{\lambda_i^k}\lambda_i^k\delta t\left(p_i^k(\lambda_i^k,\bar\lambda_{-i}^k)+R_i((k+1)\delta t,C_i^k-1,C_{-i}^k|\bar\lambda_{-i})\right)\\&+\quad\sum_{j\neq i,j\in\mathcal{I}_+^k}\bar\lambda_j^k\delta tR_i((k+1)\delta t,C_i^k,\boldsymbol{C_{-i}^k}-\boldsymbol{e_j},\boldsymbol{\bar\lambda_{-i}})\\&+\quad(1-(\lambda_i^k+\sum_{j\neq i,j\in\mathcal{I}_+^k}\bar\lambda_j^k)\delta t)R_i((k+1)\delta t,C_i^k,\boldsymbol{C_{-i}^k},\boldsymbol{\bar\lambda_{-i}})\end{aligned}$$

By rearranging the terms in the above equation, we get:

$$\begin{aligned}&R_i(k\delta t,C_i^k,\boldsymbol{C_{-i}^k},\boldsymbol{\bar\lambda_{-i}})-R_i((k+1)\delta t,C_i^k,\boldsymbol{C_{-i}^k},\boldsymbol{\bar\lambda_{-i}})=\\&\quad\max_{\lambda_i^k}\lambda_i^k\delta t\left(p_i^k(\lambda_i^k,\boldsymbol{\bar\lambda_{-i}^k})+R_i((k+1)\delta t,C_i^k-1,\boldsymbol{C_{-i}^k},\boldsymbol{\bar\lambda_{-i}})-R_i((k+1)\delta t,C_i^k,\boldsymbol{C_{-i}^k},\boldsymbol{\bar\lambda_{-i}})\right)\\&+\quad\sum_{j\neq i,j\in\mathcal{I}_+^k}\bar\lambda_j^k\delta t\left(R_i((k+1)\delta t,C_i^k,\boldsymbol{C_{-i}^k}-\boldsymbol{e_j},\boldsymbol{\bar\lambda_{-i}})-R_i((k+1)\delta t,C_i^k,\boldsymbol{C_{-i}^k},\boldsymbol{\bar\lambda_{-i}})\right)\end{aligned}$$

Dividing by δt and letting $\delta t\to 0$, we get the above HJB equation. □

One shortcoming of the above formulation is that it implies firm i knows its competitor's policy in terms of *intensity*. It is more realistic to assume that it knows its competitor's policy in terms of *prices*. The major advantage of this formulation, over a formulation using competitor's price policy is its simplicity, which makes it easier to study. If one assumes that the function $p_j(\lambda_j,\boldsymbol{\lambda_{-j}})$ is invertible in λ_i for fixed $\boldsymbol{\lambda_{-j}}$, then we can write the intensity policy $\bar\lambda_j$ as: $\bar\lambda_j=\lambda_j(\lambda_i,\bar p_{-i})$, where $\bar p_{-i}$ is the competitor's price policy.
We now prove existence and uniqueness of the best-response policy.

Proposition 2. *Assume that there exists a function R_i solving the partial differential equation: for all feasible intensities λ_i,λ_{-i}:*

- *if* $i \in \mathcal{I}_+((C_1, \ldots, C_0)$*:*

$$\begin{aligned}
&\frac{\partial R_i(s, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}})}{\partial s} = \\
&\qquad \lambda_i \left(p_i(\lambda_i, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) + R_i(s, C_i - 1, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}}) - R_i(s, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}})\right) \\
&\quad + \sum_{j \neq i, j \in \mathcal{I}_+} \bar{\lambda}_j \left(R_i(s, C_i, \boldsymbol{C_{-i}} - \boldsymbol{e_j}, \lambda_i, \boldsymbol{\lambda_{-i}}) - R_i(s, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}})\right)
\end{aligned}$$

  with boundary conditions $R_i(t, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}}) = 0$*.*

- *if* $i \in \mathcal{I}_0((C_1, \ldots, C_0)$*:* $R_i(0, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}}) = 0$

Then, there exists a unique best-response intensity policy $\Lambda_i(C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$*.*

Note that we cannot in general prove existence of a function solving the HJB equation, and can only prove existence of the best-response policy, provided such a function exists.
Note, furthermore, that the state space has dimension N, the number of firms in the market, and hence, computations are mired by the curse of dimensionality.

Proof. By Assumptions (7.1.4), the instantaneous revenue function is strictly concave in λ_i. The function to maximize is the sum of the instantaneous revenue $r_i(\lambda_i, \lambda_{-i})$ and a linear function in λ_i. Therefore, it is strictly concave in λ_i Furthermore, due to Assumptions (7.1.4) and (7.1.5), we can restrict the problem to intensities $\lambda_i \in [0, \lambda_i^*(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})]$. Thus, the problem is that of maximizing a strictly concave function over a compact set, and by Weierstrass's theorem, it has a solution which is unique. □

The following holds regarding the best-response policy:

Proposition 3. *For all* $\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}$*, the best-response policy* $\Lambda_i(C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$ *verifies:*

$$\Lambda_i(C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \leq \lambda_i^*(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$$

Proof. Let $\lambda_i > \lambda_i^*(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$. By definition, we have:

$$r_i(\lambda_i, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) < r_i(\lambda_i^*(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}), \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$$

Since R_i is nondecreasing in C_i, then we get:

$$\begin{aligned}
&r_i(\lambda_i, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) + \lambda_i \left(R_i(C_i - 1, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i(C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})\right) \\
&\quad \leq \ r_i(\lambda_i^*(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}), \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) + \lambda_i^*(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \left(R_i(C_i - 1, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i(C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})\right)
\end{aligned}$$

As a result, no $\lambda_i > \lambda_i^*(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$ is optimal, therefore $\Lambda_i(C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \leq \lambda_i^*(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$. □

Furthermore, we can establish the following properties regarding the best-response revenue-to-go function:

Proposition 4. *The best-response revenue-to-go* $R_i(s, C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$ *is strictly increasing in* C_i*, for fixed* C_{-i}*, and strictly increasing in* s *for fixed* $C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}$*.*

Proof. • We first prove that R_i is nondecreasing in C_i, ceteris paribus. Indeed,

$$R_i(s, C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) = \sup_{\lambda_i \in \mathcal{L}_i(\bar{\lambda}_{-i})} \mathbb{E}\left\{\int_0^s\right\}$$

where $\mathcal{L}_i(\bar{\lambda}_{-i}, C_i)$ is the space of allowable intensities for i, such that:

$$\int_0^s dN_i(s) \leq C_i$$

Any policy in $\mathcal{L}_i(\bar{\lambda}_{-i}, C_i)$ is allowable in $\mathcal{L}_i(\bar{\lambda}_{-i}, C_i+1)$. Hence, the policy which is optimal for capacity $C_i + 1$ yields a revenue from 0 to s which is no smaller than the optimal revenue for capacity C_i. This means that the best-response revenue-to-go is nondecreasing in C_i.

• Let $s < s'$: we have:

$$R_i(s', C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i(s, C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) =$$
$$\mathbb{E}_{\lambda_i}\left\{\int_s^{s'} p_i(s) dN_i(s)\right\}$$

The above quantity is positive, hence $R_i(s', C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i(s, C_i, C_{-i}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) > 0$ for $s < s'$.

□

However, we cannot establish monotonicity of the policy or of the revenue-to-go w.r.t the firm's own capacity.

7.2.2 Markov-Perfect equilibrium

A Markov-perfect equilibrium $\Lambda_1^*, \ldots, \Lambda_N^*$ exists if the system of HJB equations for the firms has a solution:

Proposition 5. *Assume that there exist $R_1^*, \ldots, R_N^* : [0, t] \times \mathbb{N}^N \to \mathbb{R}_+$, continuously differentiable, and $\Lambda_1^*, \ldots, \Lambda_N^* : [0, t] \times \mathbb{N}^N \to \mathbb{R}_+$, such that for all i:*

$$\begin{aligned}\frac{\partial R_i^*(s, C_i, \boldsymbol{C_{-i}})}{\partial s} &= \\ &\max_{\lambda_i} \lambda_i \left(p_i(\lambda_i, \boldsymbol{\Lambda^*_{-i}}(\boldsymbol{C_{-i}}, C_i)) + R_i^*(s, C_i - 1, \boldsymbol{C_{-i}}) - R_i^*(s, C_i, \boldsymbol{C_{-i}})\right) \\ + &\sum_{j \in \mathcal{I}_+, j \neq i} \Lambda_j^*(C_j, \boldsymbol{C_{-j}}) \left(R_i^*(s, C_i, \boldsymbol{C_{-i}} - \boldsymbol{e_j}) - R_i^*(s, C_i, \boldsymbol{C_{-i}})\right)\end{aligned} \tag{7.2}$$

with boundary conditions $R_i^*(0, C_i, C_{-i}) = 0$ and $R_i^*(s, 0, C_{-i}) = 0$ for all $s \leq t$. Then $\Lambda_1^*, \ldots, \Lambda_N^*$ are the Markov-perfect equilibrium policies for the firms.

Hence, computing the MPE policies requires solving the system of N HJB equations simultaneously, each of them having a state space of dimension N. As a result of this, the problem is not computationally tractable, except for monopolies or duopolies. For this reason, we will explore in Section 3 some approximations of the problem which aim at reducing the complexity of the best-response and MPE problems.
Using standard tools from Game Theory, it is possible to establish existence of Markov-perfect equilibrium policies.

Proposition 6. *Assume that there exist functions $R_1, \dots, R_N$ simultaneously solving the partial differential equations: for all i, for all feasible intensities $\lambda_i, \boldsymbol{\lambda_{-i}}$:*

- *if $i \in \mathcal{I}_+((C_1, \dots, C_0)$:*

$$\begin{aligned}\frac{\partial R_i(s, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}})}{\partial s} &= \\ &\lambda_i \left(p_i(\lambda_i, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) + R_i(s, C_i - 1, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}}) - R_i(s, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}})\right) \\ + &\sum_{j \neq i, j \in \mathcal{I}_+} \bar{\lambda}_j \left(R_i(s, C_i, \boldsymbol{C_{-i}} - \boldsymbol{e_j}, \lambda_i, \boldsymbol{\lambda_{-i}}) - R_i(s, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}})\right)\end{aligned}$$

  with boundary conditions $R_i(t, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}}) = 0$.

- *if $i \in \mathcal{I}_0((C_1, \dots, C_0)$: $R_i(0, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}}) = 0$*

There exists a set of Markov-perfect equilibrium policies.

Proof. Let us fix $C_i, \boldsymbol{C_{-i}}$. Let $s \in [0, t]$. The MPE policies are those maximizing the following objectives: for all i, we have:

$$\max_{\lambda_i} r_i(\lambda_i, \boldsymbol{\lambda_{-i}}) + \lambda_i(R_i^*(s, C_i - 1, \boldsymbol{C_{-i}}) - R_i^*(s, C_i, \boldsymbol{C_{-i}}))$$

Since r_i is strictly concave in λ_i and the second term in the objective is linear in λ_i, then the objective is strictly concave in λ_i. It is furthermore continuous in $\boldsymbol{\lambda}$. By consequence of Assumptions (7.1.6), we have seen that the set of feasible intensities in compact.
Therefore, we can use Arrow-Debreu-Glicksberg theorem to prove existence of a Nash equilibrium. $\square$

However, even for simple intensity functions, the MPE policies are not available in closed form.
Uniqueness of the MPE policies requires an additional assumption.

Assumption 7.2.

$$|\frac{\partial^2 r_i}{\partial \lambda_i^2}| > \sum_{j \neq i} |\frac{\partial^2 r_i}{\partial \lambda_i \lambda_j}|$$

$$|\frac{\partial^2 r_i}{\partial \lambda_i^2}| > \sum_{j \neq i} |\frac{\partial^2 r_j}{\partial \lambda_i \lambda_j}|$$

Proposition 7. *Assume that there exist functions $R_1, \ldots, R_N$ simultaneously solving the partial differential equations: for all i, for all feasible intensities λ_i, $bm\lambda_{-i}$:*

- *if $i \in \mathcal{I}_+((C_1, \ldots, C_0)$:*

$$\begin{aligned} -\frac{\partial R_i(s, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}})}{\partial s} &= \\ &\lambda_i \left(p_i(\lambda_i, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) + R_i(s, C_i - 1, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}}) - R_i(s, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}})\right) \\ + &\sum_{j \neq i, j \in \mathcal{I}_+} \bar{\lambda}_j \left(R_i(s, C_i, \boldsymbol{C_{-i}} - \boldsymbol{e_j}, \lambda_i, \boldsymbol{\lambda_{-i}}) - R_i(s, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}})\right) \end{aligned}$$

  with boundary conditions $R_i(t, C_i, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}}) = 0$.

- *if $i \in \mathcal{I}_0((C_1, \ldots, C_0)$: $R_i(0, \boldsymbol{C_{-i}}, \lambda_i, \boldsymbol{\lambda_{-i}}) = 0$*

Under Assumption 7.2, the MPE policies are unique.

Proof. Assumption 7.2 implies that the mapping:

$$\boldsymbol{F}(\boldsymbol{\lambda}) = \left(-\frac{\partial r_1}{\partial \lambda_1}, \ldots, -\frac{\partial r_N}{\partial \lambda_N}\right)'$$

is strictly monotone. As a result, the MPE policies are unique. □

Furthermore, the monotonicity property w.r.t time that we established for the best-response problem can be extended to the equilibrium problem.

Proposition 8. *The MPE revenue-to-go is strictly increasing in s.*

The proof is similar to that for the best-response problem, and is thus omitted. However, we cannot establish monotonicity of the equilibrium policy or revenue w.r.t the firm's own capacity.

7.3 The Discrete-Time Problem

In this section, we consider an approximation of the above model in discrete-time. Continuous-time pricing policies have many detractors. Indeed, as discussed in Bitran and Mondschein [15], they give rise to "saw-toothed" policies, where the price drops after every sales. Their implementation is unrealistic, and they tend to raise issues regarding fairness to the consumer. In a discrete-time model, the firms price their product at each time period simultaneously.

Stochastic Dynamic Pricing Problem

We consider a discretization of the interval $[0, t]$, where the time period δt is chosen so that at most one unit of demand can occur during one period. The finite horizon t is thus divided into $n = t/\delta t$ time periods of length δt. Furthermore, we make the following assumption on δt.

Assumption 7.3. *δt is such that there is at most one unit of demand occurring in any interval $[k\delta t, (k+1)\delta t)$.*

We use superscript k to denote the time period $[(n-k-1)\delta t, (n-k)\delta t)$. Let d_i^k denote demand at time k. Let $\mathcal{I}_+^k$ be the set of indices of firms with positive capacity in period k, and $\mathcal{I}_0^k$ the set of indices of firms with zero capacity in period k. By Assumption 7.3, the Poisson purchasing process is approximated by the following Bernoulli process:

$$(\boldsymbol{d}_+^k, \boldsymbol{d}_\mathbf{0}^k) = \begin{cases} (\boldsymbol{e_i}, \mathbf{0}) & \text{w.p } \lambda_i^k \delta t \quad \forall\ i \in \mathcal{I}_+^k \\ (\mathbf{0}, \mathbf{0}) & \text{w.p } 1 - \sum_{i \in \mathcal{I}_+^k} \lambda_i^k \delta t \end{cases}$$

7.3.1 Best-Response Problem

Let us fix competitors' strategy to $\bar{\boldsymbol{\lambda}}^{\boldsymbol{1}}_{\boldsymbol{-i}}, \ldots, \bar{\boldsymbol{\lambda}}^{\boldsymbol{n}}_{\boldsymbol{-i}}$, and consider the revenue management problem of firm i. At each period k, firm i aims at maximizing its expected revenue for the remainder of the horizon, given its belief $\bar{\lambda}_{-\mathbf{i}}$ regarding its competitors' intensities. The best-response policy and best-response revenue-to-go of firm i can be computed by backwards induction: if $i \in \mathcal{I}_+^k$:

$$\begin{aligned} & R_i^k(C_i^k, \boldsymbol{C}^{\boldsymbol{k}}_{\boldsymbol{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \\ = & \max_{\lambda_i^k} \lambda_i^k \delta t \left(p_i^k(\lambda_i^k, \bar{\lambda}_{-i}^k) + R_i^{k+1}(C_i^k - 1, \boldsymbol{C}^{\boldsymbol{k}}_{\boldsymbol{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \right) \\ + & \sum_{j \in \mathcal{I}_+^k, j \neq i} \bar{\lambda}_j^k \delta t R_i^{k+1}(C_i^k, \boldsymbol{C}^{\boldsymbol{k}}_{\boldsymbol{-i}} - \boldsymbol{e_j}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) + (1 - (\lambda_i^k + \sum_{j \in \mathcal{I}_+^k, j \neq i} \bar{\lambda}_j^k)\delta t) R_i^{k+1}(C_i^k, \boldsymbol{C}^{\boldsymbol{k}}_{\boldsymbol{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \end{aligned} \tag{7.3}$$

Furthermore, the boundary conditions are:

$$R_i^n(C_i^n, \boldsymbol{C}^{\boldsymbol{n}}_{\boldsymbol{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) = 0 \tag{7.4}$$

$$R_i^k(0, \boldsymbol{C}^{\boldsymbol{t}}_{\boldsymbol{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) = 0 \tag{7.5}$$

The best-response policy $\Lambda_i^k(C_i^k, \boldsymbol{C}^{\boldsymbol{k}}_{\boldsymbol{-i}}, \boldsymbol{\lambda}_{\boldsymbol{-i}})$ maximizes the right hand side of the backward induction equation:

$$\begin{aligned} & \lambda_i^k(C_i^k, \boldsymbol{C}^{\boldsymbol{k}}_{\boldsymbol{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \\ = & \arg\max_{\lambda_i^k} \lambda_i^k \delta t \left(p_i^k(\lambda_i^k, \bar{\boldsymbol{\lambda}}^{\boldsymbol{k}}_{\boldsymbol{-i}}) + R_i^{k+1}(C_i^k - 1, \boldsymbol{C}^{\boldsymbol{k}}_{\boldsymbol{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i^{k+1}(C_i^k, \boldsymbol{C}^{\boldsymbol{k}}_{\boldsymbol{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \right) \end{aligned} \tag{7.6}$$

Notice that the discretization transforms the HJB equation into a backward induction equation. The only reduction in complexity is therefore to avoid the partial derivative and its integration. However, the dimension of the state space is still N, and thus the discrete-time problem is still subject to the curse of dimensionality.
We now establish existence and uniqueness of the best-response policy, and best-response revenue-to-go.

Proposition 9. *There exist unique best-response intensity policy and best-response revenue-to-go.*

Proof. The proof is by induction. For $k = n-1$, the best-response policy for firm i is obtained by maximizing the instantaneous revenue function $r_i(\lambda_i^n, \boldsymbol{\lambda_{-i}^n})$. By Assumption (7.1.5), the instantaneous revenue has a bounded maximizer. By strict concavity, this maximizer is unique. Hence the best-response policy for period n exists and is unique. The best-response revenue to go in period $n-1$ exists and is therefore:

$$R_i^{n-1}(C_i, \boldsymbol{C_{-i}}, \boldsymbol{\lambda_{-i}}) = r_i(\Lambda_i^{n-1}(C_i, \boldsymbol{C_{-i}}, \boldsymbol{\lambda_{-i}}), \boldsymbol{\lambda_{-i}^{n-1}})\delta t$$

Let us assume that the best-response policy and best-response revenue-to-go exist for periods $n-1$ to $k+1$. The best-response problem in period k is that of maximizing the following function:

$$\begin{aligned} h_i^k(\lambda_i^k) &= r_i^k(\lambda_i^k, \boldsymbol{\bar{\lambda}_{-i}^k})\delta t \\ &+ \lambda_i \delta t \left(R_i^{k+1}(C_i^k - 1, \boldsymbol{C_{-i}^k}, \boldsymbol{\bar{\lambda}_{-i}}) - R_i^{k+1}(C_i^k, \boldsymbol{C_{-i}^k}, \boldsymbol{\bar{\lambda}_{-i}}) \right) \end{aligned}$$

where $R_i^{k+1}(C_i^k - 1, \boldsymbol{C_{-i}^k}, \boldsymbol{\bar{\lambda}_{-i}})$ is the best response revenue-to-go in period $k+1$. By Assumptions (7.1.4) and (7.1.5), the problem can be restricted to intensities $\lambda_i^k \in [0, \lambda_i^*(\boldsymbol{\bar{\lambda}_{-i}})]$. By strict concavity of h_i (sum of r_i which is strictly concave, and a linear function), we conclude that there is a unique maximizer. Hence, the best-response policy for period k exits and is unique. By plugging the best-response policy into the right-hand side of the backward induction equation, we obtain the best-response revenue-to-go for period k.

Therefore, by induction principle, existence and uniqueness of the best-response policy and best-response revenue-to-go holds for all $0 \leq k \leq n-1$. □

Furthermore, similarly to the continuous-time problem, we state the following result:

Proposition 10. *The best-response revenue-to-go is decreasing in k.*

The proof is similar to that of the continuous-time problem (note that s increases corresponds to k decreases and vice-versa).

7.3.2 Markov-Perfect Equilibrium

In equilibrium, all firms want to find their closed-loop optimal policies, knowing that their competitors are doing the same. Hence, the Markov-perfect equilibrium policies solve the backward induction equations for all firms simultaneously. Let $R_i^{*k}(C_i^k, \boldsymbol{C_{-i}^k})$ be the MPE revenue-to-go and Λ_i^* the MPE policy. For all i, $\Lambda_i^{*k}(C_i^k, \boldsymbol{C_{-i}^k})$ solves the following:

$$\begin{aligned} & R_i^{*k}(C_i^k, \boldsymbol{C_{-i}^k}) \qquad (7.7) \\ &= \max_{\lambda_i^k} \lambda_i^k \delta t \left(p_i^k(\lambda_i^k, \Lambda_{-i}^{*k}(\boldsymbol{C_{-i}^k}, C_i^k)) + R_i^{*k+1}(C_i^k - 1, \boldsymbol{C_{-i}^k}) \right) \\ &+ \sum_{j \in \mathcal{I}_+^k, j \neq i} \Lambda_j^{*k} \delta t R_i^{*k+1}(C_i^k, \boldsymbol{C_{-i}^k} - \boldsymbol{e_j}) + (1 - (\lambda_i^k + \sum_{j \in \mathcal{I}_+^k, j \neq i} \Lambda_j^{*k}(C_j^k, \boldsymbol{C_{-j}^k}))\delta t) R_i^{k+1}(C_i^k, \boldsymbol{C_{-i}^k}) \end{aligned}$$

Hence the MPE problem in discrete time is the system of N N-dimensional backward induction equations above. As its continuous-time counterpart, this problem is thus

not computationally tractable in general.
We now state the first result of this subsection:

Proposition 11. *There exists a Markov-perfect equilibrium policy.*

The proof relies on the same arguments as the proof of Proposition 6 and is therefore omitted.

Uniqueness can be obtained under Assumption 7.2.

Proposition 12. *Under Assumption 7.2, the Markov-perfect equilibrium policies are unique.*

The proof is similar to the continuous-time case, and is therefore omitted.

7.3.3 Properties of the Best-response and Equilibrium Problems

We establish the following properties regarding the best-response and MPE revenue-to-go functions and policies:

Proposition 13.

- *If $C_i^k \geq n-k$, then the best-response policy $\Lambda_i^k(C_i^k, C_{-i}^k, \bar{\boldsymbol{\lambda}}_{\boldsymbol{i}})$ and revenue-to-go function $R_i^k(C_i^k, C_{-i}^k, \bar{\boldsymbol{\lambda}}_{\boldsymbol{i}})$ are independent of C_i^k;*
- *Let $j \in \mathcal{I}_+^k$, $j \neq i$. If $C_j^k \geq n-k$, then the best-response policy $\Lambda_i^k(C_i^k, C_{-i}^k, \bar{\boldsymbol{\lambda}}_{\boldsymbol{i}})$ and revenue-to-go function $R_i^k(C_i^k, C_{-i}^k, \bar{\boldsymbol{\lambda}}_{\boldsymbol{i}})$ are independent of C_j^k;*
- *If for all $j \in \mathcal{I}_+^k$, $C_j^k \geq n-k$, then the Markov-perfect policies $(\Lambda_i^{*k}(C_i^k, \boldsymbol{C_{-i}^k})$ and the Markov-perfect equilibrium revenue-to-go functions $(R_i^{*k}(C_i^k, \boldsymbol{C_{-i}^k})$ are independent of $\{C_j^k\}_{j \in \mathcal{I}_+^k}$.*

We prove the proposition for the best-response problem, since the proof is very similar in equilibrium.

Proof. Let $i \in \mathcal{I}_+^{n-1}$. Let $k = n-1$. The best-response policy and revenue-to-go function of firm i are such that:

$$\begin{aligned} & R_i^{*n-1}(C_i^{n-1}, \boldsymbol{C_{-i}^{n-1}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \\ & \quad = \max_{\lambda_i^{n-1}} \lambda_i^{n-1} \delta t p_i^{n-1}(\lambda_i^{n-1}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \end{aligned}$$

Hence firm i's best-response policy does not depend on its capacity and neither does its best-response revenue-to-go function. Assume that the property holds for $k+1$, i.e $\Lambda_i^{k+1}(C_i^{k+1}, \boldsymbol{C_{-i}^{k+1}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$ and $R_i^{k+1}(C_i^{k+1}, C_{-i}^{k+1}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$ is independent of $C_i^{k+1} \geq n-k-1$. Let $C_i^k \geq n-k$. Since $C_i^k - 1, C_i^k, \geq n-k-1$, then the revenue-to-go functions in the right-hand side of Equation (3) are independent of i's capacity. As a result, the best-response policy, and hence, the best-response revenue-to-go function, are independent of i's capacity.
By induction, we conclude that the property holds for all k. $\square$

Furthermore, we establish the following property for the best-response policy and revenue-to-go:

Proposition 14. *The best-response policy and revenue-to-go functions for firm i are independent of competitors' capacity $\boldsymbol{C_{-i}}$.*

Proof. The proof is by induction.
For $k = n-1$, the result holds trivially. Let us assume that the best-response policy $\lambda_i^{k+1}(C_i, \boldsymbol{c_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$ and the best-response revenue-to-go $R_i^{k+1}(C_i, \boldsymbol{C_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$ are independent of $\boldsymbol{C_{-i}}$.
The best-response policy solves the maximization problem:

$$\begin{aligned}
&\lambda_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \\
&\quad = \arg\max_{\lambda_i^k} \lambda_i^k \delta t \left(p_i^k(\lambda_i^k, \bar{\boldsymbol{\lambda}}^{\boldsymbol{k}}_{\boldsymbol{-i}}) + R_i^{k+1}(C_i^k - 1, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i^{k+1}(C_i^k, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \right)
\end{aligned}$$

Hence, the policy is independent of competitors' capacity. Furthermore, $R_i^{k+1}(C_i, \boldsymbol{C_{-i}} - 1, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i^{k+1}(C_i, \boldsymbol{C_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) = 0$. Therefore, the revenue-to-go at period k is independent of $\boldsymbol{C_{-i}}$. □

Next, we would like to establish some monotonicity properties on the policies and revenue-to-go functions in general. Indeed, such properties exist for the monopoly problem, as established in Gallego and van Ryzin [46]. However, as the counterexample below suggests, we cannot establish a monotonic behavior of the equilibrium policies as a function of capacity.

Counterexample 1. *Consider the duopoly market with exponential rate function:*

$$\lambda_i(p-i, p_{-i}) = \theta_i \exp\{-\alpha_i p_i - \beta_i p_{-i}\}$$

Assume $C_{-i}^{n-3} \geq 3$. Then the following holds: for all $\bar{\lambda}_{-i}$,

$$\Lambda_i^{n-3}(3, C_{-i}, \bar{\lambda}_{-i}) \geq \Lambda_i^{n-3}(2, C_{-i}, \bar{\lambda}_{-i})$$

But for $\bar{\lambda}_{-i}^{n-1} \leq \theta_{-i} \left(\frac{e\ln(2)}{\theta_i \delta t}\right)^{\alpha_{-i}/\beta_i}$, we have:

$$\Lambda_i^{n-3}(2, C_{-i}, \bar{\lambda}_{-i}) \leq \Lambda_i^{n-3}(1, C_{-i}, \bar{\lambda}_{-i})$$

Furthermore, we have for all $\bar{\lambda}_{-i}$,

$$R_i^{n-3}(3, C_{-i}, \bar{\lambda}_{-i}) \geq R_i^{n-3}(2, C_{-i}, \bar{\lambda}_{-i})$$

But if the two conditions below are satisfied:

$$\frac{\theta_i \delta t}{2e} \left(\frac{\bar{\lambda}_{-i}^{n-2}}{\theta_{-i}}\right)^{\beta_i/\alpha_{-i}} \leq \frac{\beta_i}{\alpha_{-i}} \ln\left(\frac{\bar{\lambda}_{-i}^{n-3}}{\theta_{-i}}\right) + \ln(\theta_i \delta t) - 1$$

$$\frac{\theta_i}{e} \delta t \left(\right)$$

we have:

$$R_i^{n-3}(2, C_{-i}, \bar{\lambda}_{-i}) \leq R_i^{n-3}(1, C_{-i}, \bar{\lambda}_{-i})$$

Proof. We first notice that if $C^k_{-i} \geq n-k$, since the best-response revenue-to-go does not depend on capacity, the problem becomes:

$$R^k_i(C^k_i, C^k_{-i}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}}) = \\ \max_{\lambda^k_i} \lambda^k_i \delta t \left(p_i(\lambda^k_i, \bar{\lambda}^k_{-i}) + R^k_i(C^k_i - 1, C^k_{-i}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}}) - R^k_i(C^k_i, C^k_{-i}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}})\right)$$

Furthermore, for the exponential rate function, the first order optimality condition is:

$$p_i(\lambda^k_i, \bar{\lambda}^k_{-i}) + R^k_i(C^k_i - 1, C^k_{-i}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}}) - R^k_i(C^k_i, C^k_{-i}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}}) = \frac{\alpha_{-i}}{\Delta}$$

where $\Delta = \alpha_i\alpha_{-i} - \beta_i\beta_{-i}$.
Hence, the best-response revenue-to-go can be written as:

$$R^k_i(C^k_i, C^k_{-i}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}}) = \frac{\alpha_{-i}}{\Delta}\delta t \Lambda_i(C^k_i, C^k_{-i}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}}) + R^{k+1}_i(C^k_i, C^k_{-i}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}})$$

At period $n-1$, the best-response policy solves: for $C^{n-1}_i \geq 1$:

$$\max_{\lambda^{n-1}_i} \lambda^{n-1}_i \delta t p_i(\lambda_i, \bar{\lambda}^{n-1}_{-i})$$

Hence,

$$\Lambda^{n-1}_i(C^{n-1}_i, C^{n-1}_{-i}, \bar{\lambda}^{n-1}_{-i}) = \begin{cases} \frac{\theta_i}{e}\left(\frac{\bar{\lambda}^{n-1}_{-i}}{\theta_{-i}}\right)^{\beta_i/\alpha_{-i}} & C^{n-1}_i \geq 1 \\ 0 & C^{n-1}_i = 0 \end{cases}$$

and the best-response revenue-to-go at $n-1$ is:

$$R^{n-1}_i(C^{n-1}_i, C^{n-1}_{-i}, \bar{\lambda}^{n-1}_{-i}) = \frac{\alpha_{-i}}{\Delta}\delta t \Lambda^{n-1}_i(C^{n-1}_i, C^{n-1}_{-i}, \bar{\lambda}^{n-1}_{-i})$$

At period $n-2$, the best-response policy solves:

- If $C^{n-2}_i \geq 2$:

$$\max_{\lambda^{n-2}_i} \lambda^{n-2}_i \delta t p_i(\lambda^{n-2}_i, \bar{\lambda}^{n-2}_{-i})$$

- If $C^{n-2}_i = 1$:

$$\max_{\lambda^{n-2}_i} \lambda^{n-2}_i \delta t \left(p_i(\lambda^{n-2}_i, \bar{\lambda}^{n-2}_{-i}) - R^{n-1}_i(1, C^{n-2}_{-i}, \bar{\lambda}^{n-1}_{-i}\right)$$

- If $C^{n-2}_i = 0$: then $\Lambda^{n-2}_i(0, C^{n-2}_{-i}, \bar{\lambda}^{n-2}_{-i}) = 0$.

We thus obtain the closed-form solutions for the best-response policy at period $n-2$:

$$\Lambda^{n-2}_i(C^{n-2}_i, C^{n-2}_{-i}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}}) = \\ \begin{cases} \frac{\theta_i}{e}\left(\frac{\bar{\lambda}^{n-2}_{-i}}{\theta_{-i}}\right)^{\beta_i/\alpha_{-i}} & C^{n-2}_i \geq 2 \\ \frac{\theta_i}{e}\left(\frac{\bar{\lambda}^{n-2}_{-i}}{\theta_{-i}}\right)^{\beta_i/\alpha_{-i}} \exp\{-\delta t \frac{\theta_i}{e}\left(\frac{\bar{\lambda}^{n-1}_{-i}}{\theta_{-i}}\right)^{\beta_i/\alpha_{-i}}\} & C^{n-2}_i = 1 \\ 0 & C^{n-2}_i = 0 \end{cases}$$

As a result, the best-response revenue-to-go function is:

$$R_i^{n-2}(C_i^{n-2}, C_{-i}^{n-2}, \bar{\lambda}_{-i}) = \frac{\alpha_{-i}}{\Delta}\delta t \Lambda_i^{n-2}(C_i^{n-2}, C_{-i}^{n-2}, \bar{\lambda}_{-i}) + R_i^{n-1}(C_i^{n-2}, C_{-i}^{n-2}, \bar{\lambda}_{-i})$$

Finally, for period $n-3$:

- If $C_i^{n-3} \geq 3$:

$$\max_{\lambda_i^{n-3}} \lambda_i^{n-3} \delta t p_i(\lambda_i^{n-3}, \bar{\lambda}_{-i}^{n-3})$$

- If $C_i^{n-2} = 2$:

$$\max_{\lambda_i^{n-3}} \lambda_i^{n-3} \delta t p_i(\lambda_i^{n-3}, \bar{\lambda}_{-i}^{n-3}) + \\ \lambda_i^{n-3} \delta t \left(R_i^{n-2}(1, C_{-i}^{n-3}, \bar{\lambda}_{-i}^{n-2}) - R_i^{n-2}(2, C_{-i}^{n-3}, \bar{\lambda}_{-i}^{n-2}) \right)$$

- If $C_i^{n-3} = 1$:

$$\max_{\lambda_i^{n-3}} \lambda_i^{n-3} \delta t \left(p_i(\lambda_i^{n-3}, \bar{\lambda}_{-i}^{n-3}) - R_i^{n-2}(1, C_{-i}^{n-3}, \bar{\lambda}_{-i}^{n-2} \right)$$

- If $C_i^{n-3} = 0$: then $\Lambda_i^{n-3}(0, C_{-i}^{n-3}, \bar{\lambda}_{-i}^{n-3}) = 0$.

We thus obtain the closed-form solutions:

$$\Lambda_i^{n-3}(C_i^{n-3}, C_{-i}^{n-3}, \boldsymbol{\bar{\lambda}_{-i}}) = \begin{cases} \frac{\theta_i}{e} \left(\frac{\bar{\lambda}_{-i}^{n-3}}{\theta_{-i}} \right)^{\beta_i/\alpha_{-i}} \\ \frac{\theta_i}{e} \left(\frac{\bar{\lambda}_{-i}^{n-3}}{\theta_{-i}} \right)^{\beta_i/\alpha_{-i}} \exp\left\{ -\delta t \frac{\theta_i}{e} \left(\frac{\bar{\lambda}_{-i}^{n-2}}{\theta_{-i}} \right)^{\beta_i/\alpha_{-i}} \left(1 - \exp\{ -\delta t \frac{\theta_i}{e} \left(\frac{\bar{\lambda}_{-i}^{n-1}}{\theta_{-i}} \right)^{\beta_i/\alpha_{-i}} \right. \right. \\ \frac{\theta_i}{e} \left(\frac{\bar{\lambda}_{-i}^{n-3}}{\theta_{-i}} \right)^{\beta_i/\alpha_{-i}} \exp\left\{ -\delta t \frac{\theta_i}{e} \left(\frac{\bar{\lambda}_{-i}^{n-2}}{\theta_{-i}} \right)^{\beta_i/\alpha_{-i}} \exp\{ -\delta t \frac{\theta_i}{e} \left(\frac{\bar{\lambda}_{-i}^{n-1}}{\theta_{-i}} \right)^{\beta_i/\alpha_{-i}} \} \right. \\ 0 \end{cases}$$

We now compute the ratio:

$$\ln\left(\frac{\Lambda_i^{n-3}(2, C_{-i}^{n-3}, \boldsymbol{\bar{\lambda}_{-i}})}{\Lambda_i^{n-3}(1, C_{-i}^{n-3}, \boldsymbol{\bar{\lambda}_{-i}})} \right) = \\ \frac{\theta_i}{e} \left(\frac{\bar{\lambda}_{-i}^{n-2}}{\theta_{-i}} \right)^{\beta_i/\alpha_{-i}} \left(1 - 2\exp\{ -\delta t \frac{\theta_i}{e} \left(\frac{\bar{\lambda}_{-i}^{n-1}}{\theta_{-i}} \right)^{\beta_i/\alpha_{-i}} \} \right)$$

Hence, if $\bar{\lambda}_{-i}^{n-1} \leq \theta_{-i} \left(\frac{e \ln(2)}{\theta_i \delta t} \right)^{\alpha_{-i}/\beta_i}$, the above is non positive, and $\Lambda_i^{n-3}(2, C_{-i}^{n-3}, \boldsymbol{\bar{\lambda}_{-i}}) \leq \Lambda_i^{n-3}(1, C_{-i}^{n-3}, \boldsymbol{\bar{\lambda}_{-i}})$.
Furthermore, for the revenue-to-go functions at period $n-3$, we have:

$$R_i^{n-3}(3, C_{-i}^{n-3}, \boldsymbol{\bar{\lambda}_{-i}}) - R_i^{n-3}(2, C_{-i}^{n-3}, \boldsymbol{\bar{\lambda}_{-i}}) = \\ \frac{\alpha_{-i}}{\Delta}\delta t \left(\Lambda_i^{n-3}(3, C_{-i}^{n-3}, \boldsymbol{\bar{\lambda}_{-i}}) - \Lambda_i^{n-3}(2, C_{-i}^{n-3}, \boldsymbol{\bar{\lambda}_{-i}}) \right)$$

$$+ \quad \left(R_i^{n-2}(3, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i^{n-2}(2, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})\right)$$

The last term in the above equation is equal to zero, and the first term was just proved to be nonnegative. Hence, the above quantity is nonnegative.

$$R_i^{n-3}(2, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i^{n-3}(1, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) =$$
$$\frac{\alpha_{-i}}{\Delta}\delta t \left(\Lambda_i^{n-3}(2, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - \Lambda_i^{n-3}(1, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})\right)$$
$$+ \quad \left(R_i^{n-2}(2, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i^{n-2}(1, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})\right)$$

Let us define the following:

$$\begin{aligned} x_{-i}^{n-k} &= \frac{\theta_i}{e}\left(\frac{\bar{\lambda}_{-i}^k}{\theta_{-i}}\right)^{\beta_i/\alpha_{-i}} \\ y_{-i}^{n-k} &= \exp\{-\delta t x_{-i}^k\} \end{aligned}$$

We have:

$$R_i^{n-3}(2, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i^{n-3}(1, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) =$$
$$\frac{\alpha_{-i}}{\Delta}\left(x_{-i}^3\delta t(\exp\{-x_{-i}^2\delta t(1-y_{-i}^1)\} - \exp\{x_{-i}^2\delta t y_{-i}^1\}) + x_{-i}^2\delta t(1-y_{-i}^1)\right)$$

Let $f(y) = x_{-i}^3\delta t(\exp\{-x_{-i}^2\delta t(1-y)\} - \exp\{x_{-i}^2\delta t y\}) + x_{-i}^2\delta t(1-y)$. We have:

$$f'(y) = x_{-i}^3\delta t x_{-i}^2\delta t(\exp\{-x_{-i}^2\delta t(1-y_{-i}^1)\} + \exp\{x_{-i}^2\delta t y_{-i}^1\}) - x_{-i}^2\delta t$$

$$x_{-i}^3\delta t(x_{-i}^2\delta t)^2(\exp\{-x_{-i}^2\delta t(1-y_{-i}^1)\} - \exp\{x_{-i}^2\delta t y_{-i}^1\})$$

f'' is negative for $y < 1/2$ and positive for $y > 1/2$. Hence, f' is decreasing for $y \in [0, 1/2)$ and increasing for $y \in (1/2, \infty)$. Its minimum is attained at $y = 1/2$, and we have:

$$f'(1/2) = 2x_{-i}^3\delta t x_{-i}^2\delta t \exp\{-x_{-i}^2\delta t/2\} - x_{-i}^2\delta t$$

As a result, if $\delta t/2x_{-i}^2 \geq \ln\left(x_{-i}^3\delta t\right)$, then $f'(1/2) \leq 0$.
In that case, since f' is decreasing for $y < 1/2$ and increasing for $y > 1/2$, there exist unique $\underline{y} < 1/2$ and $\bar{y} > 1/2$ such that $f'(\underline{y}) = f'(\bar{y}) = 0$.
Therefore, f' is positive for $y < \underline{y}$ or $y > \bar{y}$ and negative for $y \in (\underline{y}, \bar{y})$.
Furthermore, if the second condition in the Proposition is satisfied, then $f(0) \leq 0$, therefore guaranteeing the existence of a range of values y for which $f(y) \leq 0$. This implies that: $R_i^{n-3}(2, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \leq R_i^{n-3}(1, C_{-i}^{n-3}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$. □

This is unlike the monopoly case, for which Gallego, Van Ryzin [46] proved that the optimal revenue-to-go function and the optimal intensity policy are monotonically increasing in the firm's capacity.

7.4 Suboptimal Policies

In this section, we investigate approximations of the above problem that enable us to break the curse of dimensionality. Indeed, we have seen in Section 1 and 2 that both the continuous-time and discrete=time problems are in general ont computationally tractable. First, we consider the one-step look-ahead policy. The one-step look-ahead policy is such that next period's revenue-to-go in the backward induction equation is replaced by an approximation. We explore different approximation architectures, in particular the linear architecture and the quadratic architecture. Second, we consider the open-loop feedback policy. At each time period k, it solves the deterministic open-loop problem for the firm from k to n. It then applies the first component of the vector of optimal intensities as policy for time k. We investigate the approximate best-response problem, when one player applies the open-loop feedback policy and his competitors' policy are fixed. In equilibrium, we consider the case where all players are using the open-loop feedback policy, as well as the case where only some players use open-loop feedback, and the other use the optimal problem.

7.4.1 One-Step Look-Ahead Policy

In a one-step look-ahead policy, the backward induction equation corresponding to the best-response problem of the firm is solved, where next period's revenue-to-go is an approximation. Let $\tilde{R}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}})$. Let $\widehat{\Lambda}_i(C_i^k, C_{-i}^k, \bar{\lambda}_{-i})$. It solves the following backward induction equation:

$$\begin{aligned} &\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) && (7.8)\\ &= \arg\max_{\lambda_i^k} \lambda_i^k \delta t \left(p_i^k(\lambda_i^k, \boldsymbol{\bar{\lambda}^k_{-i}}) + \tilde{R}_i^{k+1}(C_i^k - 1, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) - \tilde{R}_i^{k+1}(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) \right) \end{aligned}$$

In order to choose the approximating function for the revenue-to-go, we exploit ideas from neuro-dynamic programming. We choose approximations which are functions of the state space. In particular, we investigate the linear and quadratic approximations:

$$\tilde{R}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) = a_i^k C_i^k$$

$$\tilde{R}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) = a_i^k C_i^k + \frac{1}{2} b_{ii}^k (C_{-i}^k)^2 + \sum_{j \neq i} b_{ij}^k C_i^k C_j^k$$

We choose the approximations so that they satisfy the boundary conditions $\tilde{R}_i^n(C_i^n, \boldsymbol{C^n_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) = 0$ (the coefficients at time n are thus zero) and $\tilde{R}_i^k(0, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) = 0$.
The revenue-to-go corresponding to the implementation of the one-step look-ahead policy is:

$$\begin{aligned} &\widehat{R}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) && (7.9)\\ &= \widehat{\lambda}_i^k \delta t \left(p_i^k(\widehat{\lambda}_i^k, \boldsymbol{\bar{\lambda}^k_{-i}}) + \widehat{R}_i^{k+1}(C_i^k - 1, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) - \widehat{R}_i^{k+1}(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) \right) \\ &+ \boldsymbol{\bar{\lambda}^k_{-i}} \delta t \left(\widehat{R}_i^{k+1}(C_i^k, \boldsymbol{C^k_{-i}} - 1, \boldsymbol{\bar{\lambda}_{-i}}) - \widehat{R}_i^{k+1}(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) \right) && (7.10) \end{aligned}$$

Best-Response Look-Ahead Policy with Linear Approximation

The linear approximation for the revenue-to-go function is:

$$\tilde{R}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) = a_i^k C_i^k$$

Hence, the corresponding one-step look-ahead policy solves:

$$\begin{aligned}
&\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) \\
&= \arg\max_{\lambda_i^k} \lambda_i^k \delta t \left(p_i^k(\lambda_i^k, \boldsymbol{\bar{\lambda}^k_{-i}}) - a_i^{k+1}\right)
\end{aligned} \tag{7.11}$$

It therefore maximizes the following:

$$r_i^k(\lambda_i^k, \boldsymbol{\bar{\lambda}^k_{-i}}) - a_i^{k+1}\lambda_i^k$$

We can establish the following properties regarding the one-step look-ahead policy:

Proposition 15. *The one-step look-ahead policy $\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}})$ is independent of the firms' capacity levels.*
Furthermore, it verifies the following:

- $a_i^{k+1} \geq 0$, *iff* $\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) \leq \lambda_i^*(\boldsymbol{\bar{\lambda}_{-i}})$, *where* $\lambda_i^*(\bar{\lambda}_{-i})$ *denotes the unconstrained maximizer of the instantaneous revenue* $r_i^k(\lambda_i^k, \boldsymbol{\bar{\lambda}^k_{-i}})$.
- $a_i^{k+1} \leq R_i^{k+1}(C_i^k, \boldsymbol{C_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) - R_i^{k+1}(C_i^k - 1, \boldsymbol{C_{-i}}, \boldsymbol{\bar{\lambda}_{-i}})$ *iff* $\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) \geq \Lambda_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}})$.

Proof.

- In the linear approximation, the maximand of the one-step look-ahead policy is independent of the state space. Hence, the policy is independent of the capacity levels.
  The first-order optimality condition is:

  $$\frac{\partial r_i^k}{\partial \lambda_i^k}(\lambda_i^k, \boldsymbol{\bar{\lambda}^k_{-i}}) = a_i^{k+1}$$

  whereas $\lambda_i^*(\bar{\lambda}_{-i})$ verifies the following:

  $$\frac{\partial r_i^k}{\partial \lambda_i^k}(\lambda_i^k, \boldsymbol{\bar{\lambda}^k_{-i}}) = 0$$

  Hence, if $a_i^{k+1} \geq 0$, i.e $\frac{\partial r_i^k}{\partial \lambda_i^k}(\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}), \boldsymbol{\bar{\lambda}^k_{-i}}) \geq \frac{\partial r_i^k}{\partial \lambda_i^k}(\lambda_i^*(\bar{\lambda}_{-i}), \boldsymbol{\bar{\lambda}^k_{-i}})$.
  By strict concavity of r_i^k in λ_i^k, the above derivative is strictly decreasing in λ_i^k. As a result, the above inequality implies:

  $$\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \boldsymbol{\bar{\lambda}_{-i}}) \leq \lambda_i^*(\bar{\lambda}_{-i})$$

- The first-order optimality condition for the approximation is:

  $$\frac{\partial r_i^k}{\partial \lambda_i^k}(\lambda_i^k, \boldsymbol{\bar{\lambda}^k_{-i}}) = a_i^{k+1}$$

whereas for the optimal problem, it is:

$$\frac{\partial r_i^k}{\partial \lambda_i^k}(\lambda_i^k, \bar{\boldsymbol{\lambda}}^{\boldsymbol{k}}_{\boldsymbol{-i}}) = R_i^{k+1}(C_i^k, \boldsymbol{C_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i^{k+1}(C_i^k - 1, \boldsymbol{C_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$$

Hence, if $a_i^{k+1} \leq R_i^{k+1}(C_i^k, \boldsymbol{C_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) - R_i^{k+1}(C_i^k - 1, \boldsymbol{C_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$, due to strict concavity of r_i in λ_i, its first derivative is strictly decreasing in λ_i, and hence, $\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \geq \Lambda_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$.

□

In other words, by choice of coefficient a_i^k, we can predict whether the approximate policy is underestimating, or overestimating the optimal policy.

Best-Response Look-Ahead Policy with Quadratic Approximation

The approximating revenue-to-go function in the quadratic case is:

$$\tilde{R}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) = a_i^k C_i^k + \frac{1}{2} b_{ii}^k (C_{-i}^k)^2 + \sum_{j \neq i} b_{ij}^k C_i^k C_j^k$$

The look-ahead policy thus solves the following optimization problem:

$$\begin{aligned} & \widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \\ = \quad & \arg\max_{\lambda_i^k} \lambda_i^k \delta t \left(p_i^k(\lambda_i^k, \bar{\boldsymbol{\lambda}}^{\boldsymbol{k}}_{\boldsymbol{-i}}) - a_i^{k+1} - \sum_{j \neq i} b_{ij}^{k+1} C_j^k - b_{ii}^{k+1}(C_i^k - \frac{1}{2}) \right) \end{aligned} \tag{7.12}$$

The first-order optimality condition is thus:

$$\frac{\partial r_i}{\partial \lambda_i}(\lambda_i^k, \bar{\boldsymbol{\lambda}}^k_{\boldsymbol{-i}}) = a_i^{k+1} - \frac{1}{2} b_{ii}^{k+1} + \sum_{j \neq i} b_{ij}^{k+1} C_j^k + b_{ii}^{k+1} C_i^k$$

As a result, we can establish the following property:

Proposition 16.

- *$b_{ij}^{k+1} \geq 0$ iff $\widehat{\lambda}_i$ is non increasing in C_j;*
- *$b_{ii}^{k+1} \geq 0$ iff $\widehat{\lambda}_i$ is non increasing in C_i.*
- *$(C_i^k, \boldsymbol{C^k_{-i}})$ verifies:*

$$b_{ii}^{k+1} C_i^k + \sum_{j \neq i} b_{ij}^{k+1} C_j^k \geq \frac{1}{2} b_{ii}^{k+1} - a_i^{k+1}$$

iff $\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}}) \leq \lambda_i^*(\bar{\boldsymbol{\lambda}}_{\boldsymbol{-i}})$.

Proof. By strict concavity of r_i, the derivative $\frac{\partial r_i}{\partial \lambda_i}$ is strictly decreasing in λ_i. If $b_{ij}^{k+1} \geq 0$, the right-hand side of the first order optimality condition is non decreasing in C_j^k. Hence, $\widehat{\lambda}_i$ is non increasing in C_j^k.
Furthermore, if $(C_i^k, \boldsymbol{C^k_{-i}})$ verifies:

$$b_{ii}^{k+1} C_i^k + \sum_{j \neq i} b_{ij}^{k+1} C_j^k \geq \frac{1}{2} b_{ii}^{k+1} - a_i^{k+1}$$

then, the right-hand side of the first order optimality condition is nonnegative, and as a result, $\widehat{\Lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}})$ is at most equal to the unconstrained maximum. □

7.4.2 Open-Loop Feedback Policy

The open-loop feedback problem consists of solving at each period k the deterministic open-loop revenue-to-go problem from k to n. This is the problem of maximizing the total expected revenue from k to n, subject to the constraint of the expected demand not exceeding the remaining capacity.

$$\begin{array}{cc} \max\limits_{\lambda_i^k, \ldots, \lambda_i^n} & \sum_{m=k}^n \lambda_i^m p_i^m (\lambda_i^m, \bar{\boldsymbol{\lambda}}^{\boldsymbol{m}}_{-\boldsymbol{i}}) \\ \text{s.t} & \sum_{m=k}^n \lambda_i^m \leq \frac{C_i^k}{\delta t} \end{array}$$

Let us denote by $(\lambda_i^k)^o(\bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}}), \ldots, (\lambda_i^n)^o(\bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}})$ the solution to the above problem. At period k, the open-loop feedback policy consists of using the first of the above set of controls:

$$\widehat{\lambda}_i^k(C_i^k, \boldsymbol{C^k_{-i}}, \bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}}) = (\lambda_i^k)^o(\bar{\boldsymbol{\lambda}}_{-\boldsymbol{i}})$$

7.5 Computational Performance Analysis on the Approximations

We perform computations in a duopoly market, in disequilibrium. We compare the revenue-to-go functions for the one-step look-ahead policy with linear and quadratic architectures, for various values of the parameters characterizing the policy, for the open-loop feedback policy, and for the optimal policy.

7.5.1 The Data

We perform computations, on a monopoly model with $n = 10$ periods, and the following rate function: $\lambda_i(p_i, p_{-i}) = e^{-p_i + 1/2 p_{-i}}$.
We fix competitor's policy, and implement the optimal best-response policy, as well as the open-loop feedback policy, and the one-step look ahead policy with linear and quadratic architectures for various values of the coefficients of the approximations. For simplicity of implementation, we choose look-ahead policies with coefficients which

are constant over time.
In order to compare the methods, we track the percentage gap in the revenue-to-go between the suboptimal policies and the optimal policy: if $R^{*k}(C)$ denotes the revenue-to-go of the optimal policy, and $\tilde{R}^k(C)$ the revenue-to-go of a suboptimal policy, we compute the following ratio:

$$\frac{(R^{*k}(C) - \tilde{R}^k(C))}{R^{*k}(C)}$$

We plot the ratio as a function of the time period k, for various values of the capacity, and various values of the coefficients of the one-step look-ahead policy. Recall that the approximation for the one-step look-ahead policy is: $a.C + b.C^2$ in the quadratic case, and $a.C$ in the linear case. We also compute the gap between the suboptimal policies and the optimal policy for the total revenue:

$$\frac{(R^{*0}(C) - \tilde{R}^0(C))}{R^{*0}(C)}$$

We plot the gap in total revenue as a function of capacity, for various values of the coefficients of the approximation.
Since the values we plot are percentage gaps, the smaller the gap in revenue, the better the performance of the policy as compared to the optimal policy. As far as the gap in intensity, if the gap is negative, then the suboptimal intensity is larger than the optimal intensity, hence the suboptimal policy tends to sell at a rate that is too high. If the gap in intensity is positive, then the suboptimal policy sells at a rate that is too low. Similarly, if the gap in prices is negative, the suboptimal policy overprices as compared to the optimal policy; if it is positive, then it underprices. The larger the gap in absolute value, the larger the difference between the policy or price of the suboptimal policy, compared to the optimal one.
The figures we refer to in this section are to be found in appendix.

7.5.2 The Results

Linear Look-Ahead Policy

Figures B-17 through B-28 are the relevant figures for the comparison of the linear look-ahead policies, as well as the open-loop feedback policies. Figures B-17 through B-25 display the percentage gap in revenue-to-go, intensity, and price as a function of the time period: $k = 1$ represents the first time period, hence the corresponding revenue-to-go is equal to the total revenue for the selling horizon, whereas $k = 10$ is the last period for which there needs to be made a decision. In each plot, the capacity level has been fixed. Each colored line in the graphs represent the value of the gap for a specific value of the coefficient a of the linear approximation, and the value of the gap for the open-loop feedback policy. Finally, Figures B-26 to B-28 correspond to the percentage gap for a fixed time period, as a function of capacity.

- Let us first focus on the gap in revenue-to-go, for fixed capacity, as a function of the time period. In Figures B-17 to B-19, we see that the open-loop feedback revenue-to-go is worse than the linear one-step lookahead revenue-to-go, for all capacity levels and all time periods. For medium and high capacity levels (Figures B-18 and B-19), the open-loop feedback revenue-to-go is far worse. Furthermore, there is no improvement in the performance of the policy over time. For small capacity (Figure B-17), there is an improvement from $k = 10$ to $k = 1$ for the open-loop feedback policy and the one-step lookahead policy for large ($a = 1$) value of the coefficient. For smaller values of the coefficient, the one-step lookahead policy gets worse from $k = 10$ to $k = 1$. The value of the coefficient for which the policy performs best seems to depend both on the time period and the capacity level. Figure B-20 displays the total revenue, as a function of capacity, for the various approximations. The performance of the open-loop feedback, as well as the one-step lookahead policies with $a = 0.6$ and $a = 1$ gets worse for higher capacities. The performance of the one-step lookahead policies with $a = 0.1, 0.2, 0.4$ improves with capacity.

- As far as the comparison of the intensities is concerned, the open-loop feedback intensity is always larger than the optimal one. For small capacity (Figure B-21), the one-step lookahead intensity overestimates the optimal intensity for k high, and is smaller for k low. For medium capacity (Figure B-22), the one-step lookahead intensities with small value of a underestimate the optimal ones for small k, for small values of a only. For large values of a (Figure B-23), the one-step lookahead and open-loop feedback intensities always overestimate the optimal ones. The behavior with respect to capacity is best observed in Figure B-24.

- As far as prices are concerned, the behavior is the opposite as that of intensities: the suboptimal policies tend to underestimate the optimal ones (Figures B-25 to B-28).

Quadratic Look-Ahead Policy:

Figures B-29 to B-40 display the revenue-to-go, intensity and price gaps between the optimal policy and the approximate policies, when the parameter b of the quadratic term of the one-step lookahead policy is fixed at $b = 0.5$, and the parameter a of the linear term varies.

- In Figures B-29 and B-30, we display the gap in revenue-to-go, as a function of time. For small capacity (Figure B-29), the open-loop feedback policy performs worse than the one-step lookahead policies except when the coefficient a is high ($a = 1$). The smaller the value of a, the better the performance of the lookahead policy. Furthermore, the quality of the policy improves from $k = 10$ to $k = 1$. For high capacity (Figure B-30), all policies perform as poorly. The policies perform better for small capacities than large capacities (Figure B-31).

- Figures B-32 to B-35 show the gap in intensities. The suboptimal policies all underestimate the optimal intensity. For small capacity (Figure B-32), the gap with the optimal intensity decreases from $k = 10$ to $k = 1$, but it remains constant for medium and large capacities (Figures B-33 and B-34). All policies perform equally poorly for high capacities. The gap in policies improves a lot as capacity decreases, as exhibited in Figure B-35.

- Finally,the behavior of the pricing policy is displayed in Figures B-36 to B-39. The suboptimal policies underestimate the prices, as compared to the optimal policies, for all time periods, and all capacity levels. For small capacity (Figure B-36), the gap improves from $k = 10$ to $k = 1$, but remains constant for medium and high capacities (Figures B-37 and B-38). The price gap gets worse as the capacity increases, as seen on Figure B-39.

Let us now fix $a = 0.5$ and display the gap in revenue-to-go, intensity, price between the optimal policy on the one hand, and either the open-loop feedback policy, or the one-step lookahead policy with quadratic architecture, for various values of the quadratic coefficient b. The results are displayed in Figures B-40 through B-51.

- Let us first focus on the gap in revenue-to-go. The first noticeable element is that the gap increases as the value of b increases. Furthermore, the open-loop feedback policy performs worse than all the one-step lookahead policies. The behavior is roughly the same for all capacity levels, as shown in Figures B-41 to B-43. Figure B-44 shows the gap in total revenue as a function of capacity. We observe that all policies perform better for small capacity,except for $b = 0$, which corresponds to a linear architecture.

- As far as the gap in intensity is concerned, all suboptimal policies overestimate the optimal intensity, with open-loop feedback and one-step lookahead with high values of a overestimating it most (Figures B-45 to B-47). In Figure B-48, we see that the performance of the policies decreases as the capacity increases.

- Finally, the suboptimal prices underestimate the optimal prices, with the worst performance for the open-loop feedback policy and the one-step lookahead policies with high values of a, the performance being worse for high capacities. These behaviors can be observed in Figures B-49 to B-52.

7.5.3 Conclusions Regarding Approximate Policies

The computations have shown that the open-loop feedback policy is consistently outperformed by the one-step lookahead policies. The best results are actually obtained by the linear look-ahead policy. Indeed, whatever the capacity level, the worst gap in revenue achieved by the linear one-step lookahead policy is of 30%, whereas for high values of b, the quadratic lookahead policy can perform as poorly as the open-loop feedback policy (gap of 100% for large capacities).

Furthermore, the performance of the suboptimal policies depends on the capacity level, as well as on the time period. This suggests that better performances of the one-step lookahead policy can be obtained by choosing different values of the coefficients for different time periods.

7.6 Closed-Loop Best-response Dynamics

We implement a closed-loop version of a best-response dynamics approach. In evolutionary game theory, best response dynamics represents a class of strategy updating rules, where players strategies in the next round are determined by their best-response to a belief regarding their competitors' strategies, which depends on their strategies in previous rounds. The best-response dynamics scheme that we implement solves at each iteration the closed-loop best-response problem for each firm, for a value of the competitors' strategy equal to last iteration's strategy. We show computationally that the best-response dynamics scheme converges to the Markov-perfect equilibrium.

7.6.1 Description of the Best-Response Dynamics

The best-response dynamics scheme that we consider in this paper takes the following form: at each iteration (k), each firm i solves its closed-loop best-response problem, for competitors' strategy fixed at last iteration's value ${}^{(k)}\boldsymbol{\lambda}_{-i}^{1}, \ldots, {}^{(k)}\boldsymbol{\lambda}_{-i}^{T}$ for all time periods $t = 1, \ldots, n$. Each firm's policy is then updated to its new value: $\forall\ t = 1, \ldots, n, \ \ \forall\ i = 1, \ldots, N$:

$$
{}^{(k+1)}\lambda_i{}^{t} = \Lambda_i^t(C_i, \boldsymbol{C_{-i}}, {}^{(k)}\boldsymbol{\lambda}_{-i})
$$

We terminate the approach when the gap between two consecutive policy iterates crosses a convergence threshold ϵ:

$$
\|{}^{(k+1)}\boldsymbol{\lambda} - {}^{(k)}\boldsymbol{\lambda}\| \leq \epsilon
$$

Below is the algorithm for the best-response dynamics, written for $N = 2$ for simplicity.

Algorithm 8: Best-Response Dynamics
Input: A set of initial values for the policies ${}^{(0)}\boldsymbol{\lambda}$
Output: The closed-loop equilibrium policies $\boldsymbol{\Lambda}*$

```
for i = 1 to N
    Initialize (λ_i^1, ..., λ_i^n) ← ((0)λ_i^1, ..., (0)λ_i^n)
    Initialize k = 1
    repeat
        for i = 1 to 2
            foreach C_i, C_{-i}
                Fix:
```

$$(\lambda^1_{-i}(C_i, C_{-i}), \ldots, \lambda^n_{-i}(C_i, C_{-i})) \leftarrow ({}^{(k+1)}\lambda_i^1(C_i, C_{-i}), \ldots, {}^{(k+1)}\lambda_i^n(C_i, C_{-i})) \text{ if } -i < i$$

$$(\lambda^1_j(\boldsymbol{C}), \ldots, \lambda^n_j(\boldsymbol{C})) \leftarrow ((k)^{\lambda}{}^1_j(\boldsymbol{C}), \ldots, (k)^{\lambda}{}^n_j(\boldsymbol{C})) \text{ if } -i > i$$

```
                for t = 1 to n
                    Compute (k+1)λ_i^t(C_i, C_{-i}) solution of:
```

$$\max_{\lambda_i^t} \lambda_i^t \delta t \left(p_i^t(\lambda_i^t, {}^{(k)}\lambda_{-i}(C_i, C_{-i})) + {}^{(k+1)}R_i^{t+1}(C_i - 1, C_{-i}) - {}^{(k+1)}R_i^{t+1}(C_i - 1, C_{-i}) \right)$$

where ${}^{(k+1)}R_i^t(C_i, C_{-i})$ solves the backward induction equation:

$$\begin{aligned}
{}^{(k+1)}R_i^t(C_i, C_{-i}) &= {}^{(k+1)}R_i^{t+1}(C_i, C_{-i} - 1) \\
&+ \ {}^{(k+1)}\lambda_i^t(C_i, C_{-i})\delta t p_i^t({}^{(k+1)}\lambda_i^t(C_i, C_{-i}), {}^{(k)}\lambda_{-i}^t(C_i, C_{-i})) \\
&+ \ {}^{(k+1)}\lambda_i^t(C_i, C_{-i})\delta t \left({}^{(k+1)}R_i^{t+1}(C_i - 1, C_{-i}) - {}^{(k+1)}R_i^{t+1}(C_i, C_{-i}) \right) \\
&+ \ {}^{(k)}\lambda_i^t(C_i, C_{-i}) \left({}^{(k+1)}R_i^{t+1}(C_i, C_{-i} - 1) - {}^{(k+1)}R_i^{t+1}(C_i, C_{-i}) \right)
\end{aligned}$$

```
                    If i = 2 then k ← k + 1
                until  Σ_{i=1}^2 ||(k+1)λ_i − (k)λ_i|| ≤ ε
```

7.6.2 Computational Results

We implement the approach on a duopoly, for selling horizons of up to $n = 9$, for various values of the initial capacity. We also study the convergence, as the accuracy level ϵ is increased. The results are displayed on Figures B-52 through B-55.
Figures B-52 and B-53 show the gap (i.e $\sum_{i=1}^{2} \|{}^{(k+1)}\boldsymbol{\lambda_i} - {}^{(k)}\boldsymbol{\lambda_i}\|$) as a function of the iteration, for two horizon lengths: $n = 6$ and $n = 8$. What is striking is the fact that for $n = 8$, the gap first surges, and then subsides. Hence, it is not monotonically decreasing. Were the algorithm to stop after a fixed number of iterations, it might stop very far from equilibrium.
Figure B-54 shows the number of iterations needed to reach convergence, as a function

of the horizon length n. The number of iteration seems to grow exponentially with the horizon length. This is corroborated by the fact that the algorithm takes very long to converge for $n = 10$, whereas convergence requires only 63 iterations for $n = 9$. Finally, Figure B-55 shows the number of iterations required for convergence, as a function of the accuracy level: the dependence is logarithmic in the accuracy ϵ.

7.7 Conclusions on Closed-Loop Policies

In this Chapter, we modeled an oligopolistic market for perishable products under Bertrand (price) competition, where demand for each firm is modeled as a nonhomogeneous Poisson process whose rate depends on the prices of all firms in the market. The price sensitivities of the demand are time-dependent. Due to the difficulties, both theoretical and practical, pertaining to the continuous-time model, we focus on a discrete-time version of the model. For the discrete-time version of the model, we prove existence and uniqueness of Markov-perfect equilibrium strategies, and study properties of the equilibrium.

Furthermore, we investigate approximate policies. We consider the open-loop feedback policy, and one-step lookahead policies, with linear and quadratic architectures. In our computational section, we compare performance of the above policies. We establish that the one-step lookahead policy with linear architecture outperforms both the one-step lookahead policy with quadratic architecture, and the open-loop feedback policy. Our analysis suggests that, to achieve better performance for the one-step lookahead policy, one should choose coefficients with different values in different time periods.

Finally, we propose a closed-loop best-response dynamics approach to compute the subgame-perfect equilibrium policies, and study computationally its convergence.

Chapter 8

Conclusions and Future Research Directions

In this thesis, we studied oligopolistic markets for single, perishable products, operating under Bertrand or Cournot competition. We proposed a data-driven approach to joint dynamic pricing or allocation and learning of the price-demand relationship. We studied this problem in two states of the market: a state of disequilibrium, where firms' strategies are best-response to their belief regarding competitors' strategies, and a state of equilibrium, where firms' strategies are the Nash equilibrium strategies. Extending the results to multiple perishable products does not seem to present much issues from the theoretical perspective, but would make the problem more complex and computationally intensive. Such an extension could be an area of future research.

For the problem of dynamic pricing or allocation without learning, we proved existence and uniqueness of best-response strategies. We also proved existence of equilibrium strategies, as well as their uniqueness under additional assumptions. We also performed sensitivity analysis on the solutions, and proved that the solutions are Bouligand differentiable, and that the directional derivative is the unique solution to a convex quadratic problem, or a linear variational inequality. Such an analysis is useful, not only from a practical point of view - it is key for the sellers to know how their strategies would change, would their estimation of the price-demand relationship differ - but also from a theoretical point of view: we utilize these results in the design of solution methods.

For the problem with learning, we showed that we can reformulate the joint dynamic pricing (resp. allocation) with learning problem as a bilevel problem in equilibrium, and as an MPEC in equilibrium. We gave alternative formulations as a mixed integer problem. Furthermore, we proved that learning of the price-demand relationship and of the equilibrium strategies is achieved in the long run. To achieve this result, we imposed an additional constraint on the choice of the parameters: we impose that the estimated vector of parameters for the horizon does not vary from one period to the next more than the vector of prices vary from one period to the next. This makes the map of the approach a contraction. It would be interesting to investigate ways to relax this assumption, while still guaranteeing convergence of the learning approach.

We investigated solution methods for the approach. For the equilibrium problem under Bertrand competition, which turns out to be a generalized equilibrium problem, we proposed a penalization method to compute the generalized Nash equilibrium. For the joint dynamic pricing (or allocation) with learning problem, we proposed a method based on the Gauss-Newton method, which takes advantage of the fact that the estimation problem is a sum of squares. As future research, it would be interesting to investigate other solution methods, and compare their performance.
Finally, we looked into closed loop strategies in the framework of a duopoly with stochastic demand. We established existence and uniqueness of Markov perfect policies. We proposed approximations of the problem based on limited look-ahead policies. Extensions of this work would include looking at demand models which are more general than the Poisson demand model we consider here. It would also be important to establish bounds on the performance of approximate policies, and perform more in-depth computational study.

Appendix A

Notations of the Thesis

Table A.1: General Notations

N	number of firms in the market
T	number of periods in the selling horizon
$h = -H, \ldots, 0$	superscript indicating the selling horizon
$h = 0$	current selling horizon
$h < 0$	historical selling horizons
C_i	total capacity of firm i

Table A.2: Notations pertaining to the Bertrand model

$\widehat{p_i}^0(t)$	price set by firm i at period t in the current selling horizon
$\widehat{p_i}^h(t)$	historical price set by firm i at period t in past horizon h
$\overline{\mathbf{p_{-i}}}$	firm i's belief regarding its competitors' pricing strategy.
$\mathcal{BR}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$	best response problem of firm i, for for belief $\overline{\mathbf{p_{-i}}}$ regarding its competitors, and when parameters are $\beta_\mathbf{i}$
$\mathcal{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$	best response price vector of firm i, for belief $\overline{\mathbf{p_{-i}}}$ regarding its competitors, and when parameters are $\beta_\mathbf{i}$
$\mathcal{P}(\beta)$	Nash equilibrium price when parameters are β
$\mathfrak{P}_i(\overline{\mathbf{p_{-i}}}, \beta_\mathbf{i})$ (resp. $\mathfrak{P}$)	Feasible set of the best response (resp. equilibrium) problem
$\mathcal{B}_i$ (resp. $\mathcal{B}$	Feasible set of firm i's (the equilibrium) estimation problem
$\underline{\mathcal{I}}_i(\mathbf{p_i})$ (resp. $\bar{\mathcal{I}}_i(\mathbf{p_i})$)	$\{t : p_i(t) = 0\}$ (resp. $\{t : p_i(t) = p_i^{max}\}$)
$\underline{\mathcal{I}}_i^0(\beta_\mathbf{i})$ (resp. $\bar{\mathcal{I}}_i^0(\beta_\mathbf{i})$)	$\{t \in \underline{\mathcal{I}}_i(\beta_\mathbf{i}) : \underline{\mu}_i(t) = 0\}$ (resp. $\{t \in \bar{\mathcal{I}}_i(\beta_\mathbf{i}) : \bar{\mu}_i(t) = 0\}$)
$\underline{\mathcal{I}}_i^+(\beta_\mathbf{i})$ (resp. $\bar{\mathcal{I}}_i^+(\beta_\mathbf{i})$)	$\{t \in \underline{\mathcal{I}}_i(\beta_\mathbf{i}) : \underline{\mu}_i(t) > 0\}$ (resp. $\{t \in \bar{\mathcal{I}}_i(\beta_\mathbf{i}) : \bar{\mu}_i(t) > 0\}$)
$\mathcal{L}_i(\mathbf{p_i}, \overline{\mathbf{p_{-i}}}, \beta_\mathbf{i}, \mathcal{M}_i)$	Lagrangian function of the best response problem
$\mathcal{C}(\bar{\beta}_\mathbf{i}; \mathbf{d_i})$	critical cone at $\mathcal{P}_i(\bar{\beta}_i)$ in direction $\mathbf{d_i}$

Table A.3: Notations pertaining to the Cournot model

$\widehat{q_i}^0(t)$	quantity set by firm i at period t in the current selling horizon
$\widehat{q_i}^h(t)$	historical quantity set by firm i at period t in past horizon h
$\overline{\mathbf{q_{-i}}}$	firm i's belief regarding its competitors' allocation strategy.
$\mathcal{BR}_i(\overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}})$	best response problem of firm i, for for belief $\overline{\mathbf{q_{-i}}}$ regarding its competitors, and when parameters are $\alpha_{\mathbf{i}}$
$\mathcal{Q}_i(\overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}})$	best response quantity vector of firm i, for belief $\overline{\mathbf{q_{-i}}}$ regarding its competitors, and when parameters are $\alpha_{\mathbf{i}}$
$\mathcal{Q}(\alpha)$	Nash equilibrium quantity when parameters are α
$\mathfrak{Q}_i(\alpha_{\mathbf{i}})$ (resp. $\mathfrak{Q}$)	feasible set of the best response (resp. equilibrium) problem
$\mathcal{A}_i$ (resp. $\mathcal{A}$)	feasible set of firm i's (resp. the equilibrium) estimation problem
$\mathcal{I}_i(\mathbf{q_i})$	index set of active nonnegativity constraints at $\mathbf{q_i}$
$\mathcal{L}_i(\mathbf{q_i}, \overline{\mathbf{q_{-i}}}, \alpha_{\mathbf{i}}, \mathcal{M}_i)$	Lagrangian function of the best response problem
$\mathfrak{Q}_i\perp(\bar{\alpha}_{\mathbf{i}})$	Polyhedral cone at $\bar{\alpha}_{\mathbf{i}}$

Appendix B

Figures

Figure B-1: Price Sensitivities for the Airline Example

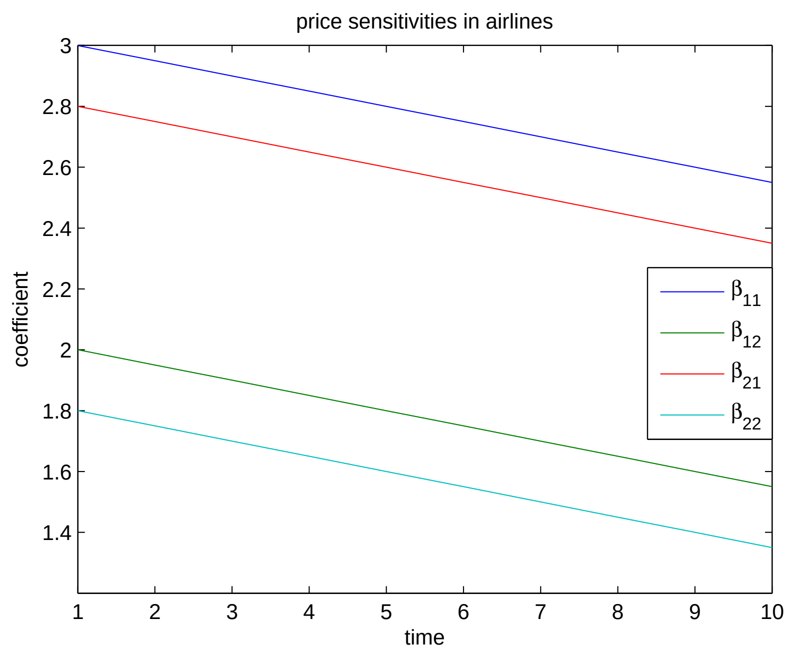


Figure B-2: Evolution of the Change in Parameters for the Airline Example ($T = 10$)

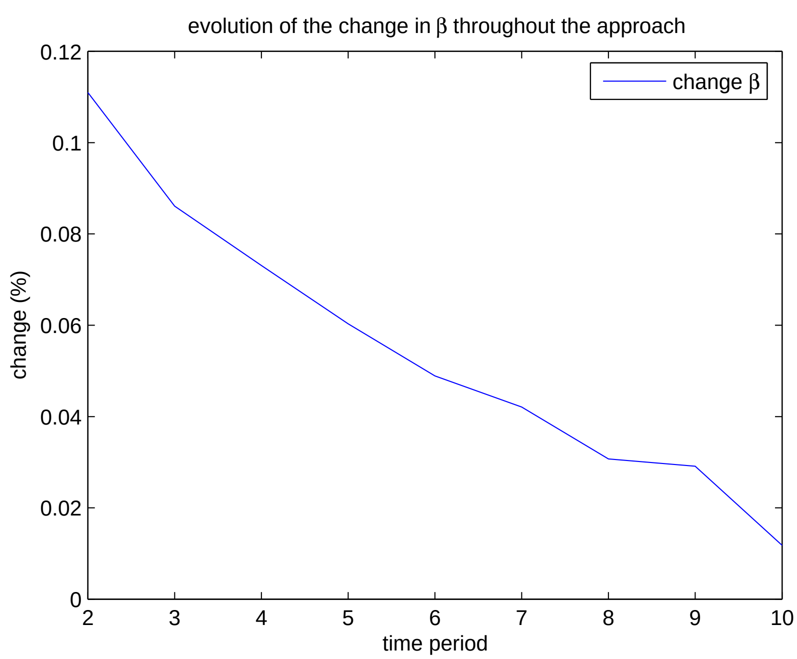


Figure B-3: Evolution of the Change in Prices for the Airline Example ($T = 10$)

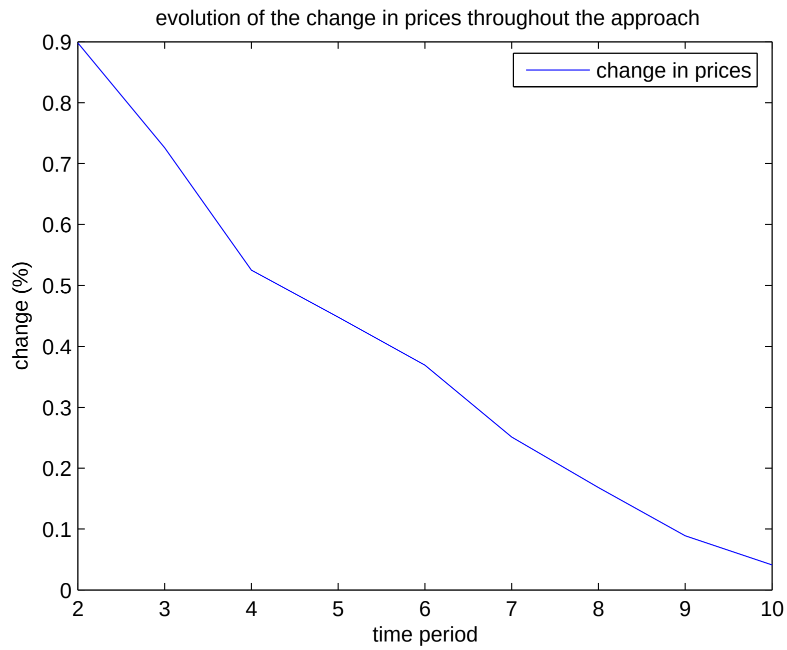


Figure B-4: Evolution of the Change in Demands for the Airline Example ($T = 10$)

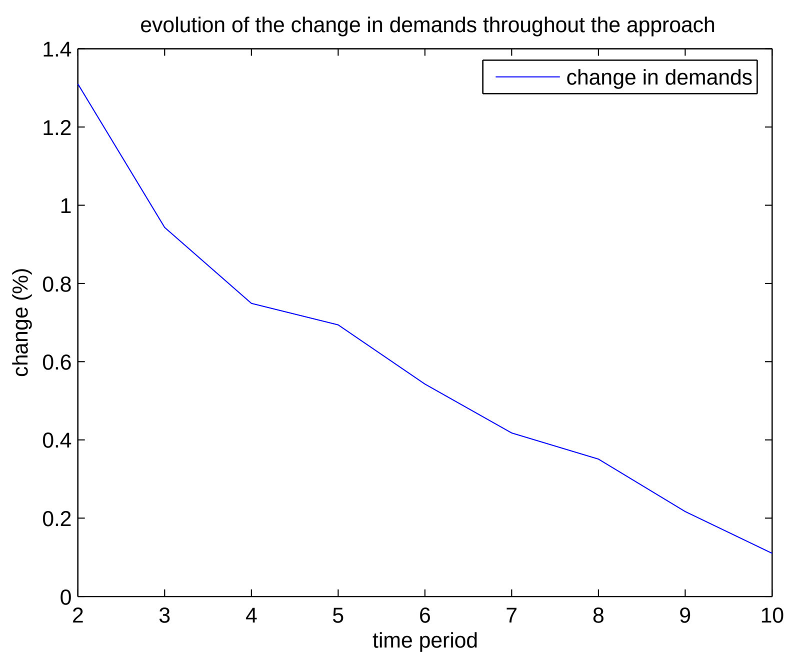


Figure B-5: Evolution of the Change in Parameters for the Airline Example ($T = 20$)

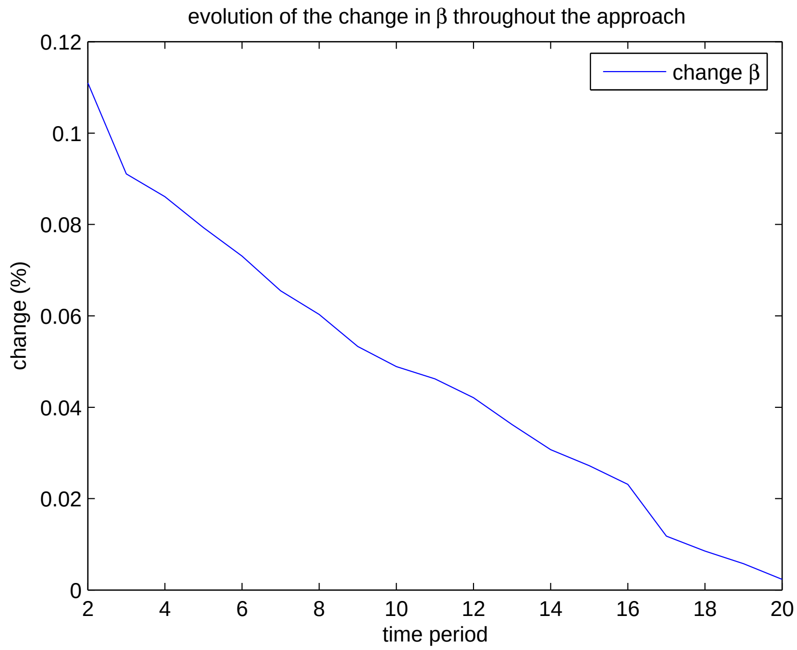


Figure B-6: Evolution of the Change in Prices for the Airline Example ($T = 20$)

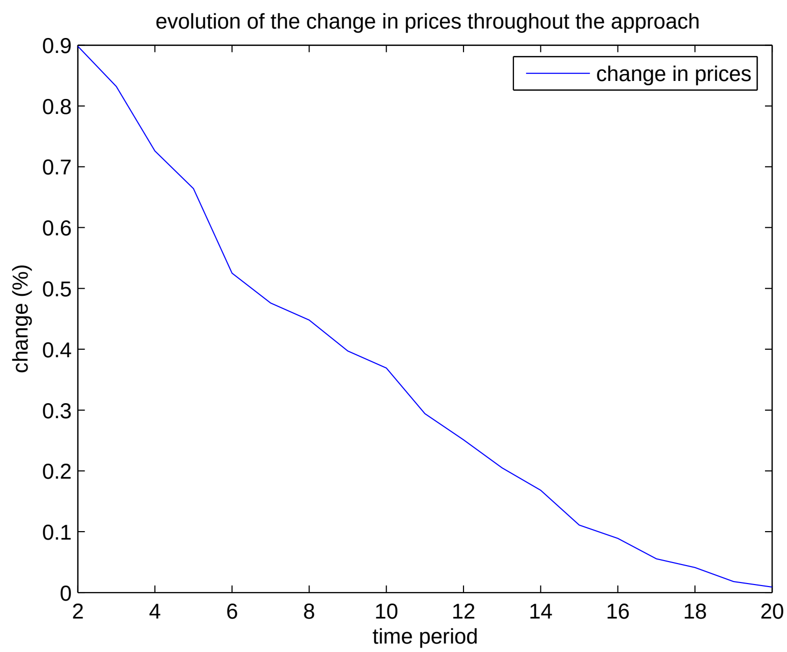


Figure B-7: Evolution of the Change in Demands for the Airline Example ($T = 20$)

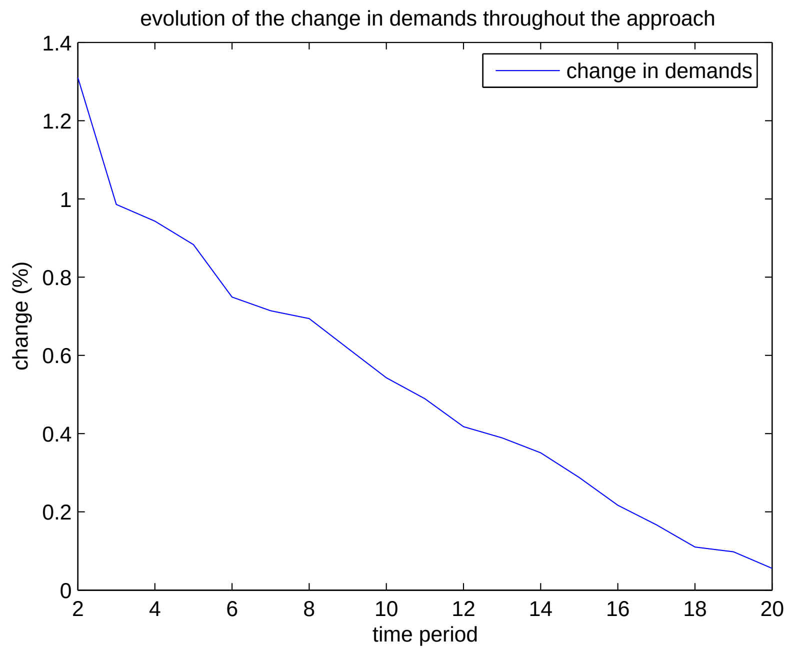


Figure B-8: Price Sensitivities for the Retail Example

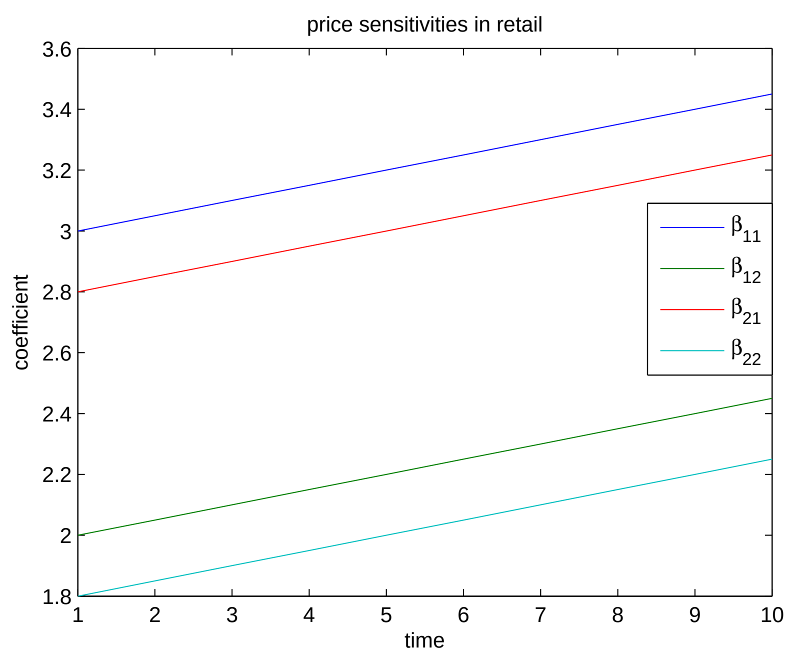


Figure B-9: Evolution of the Change in Parameters for the Retail Example ($T = 10$)

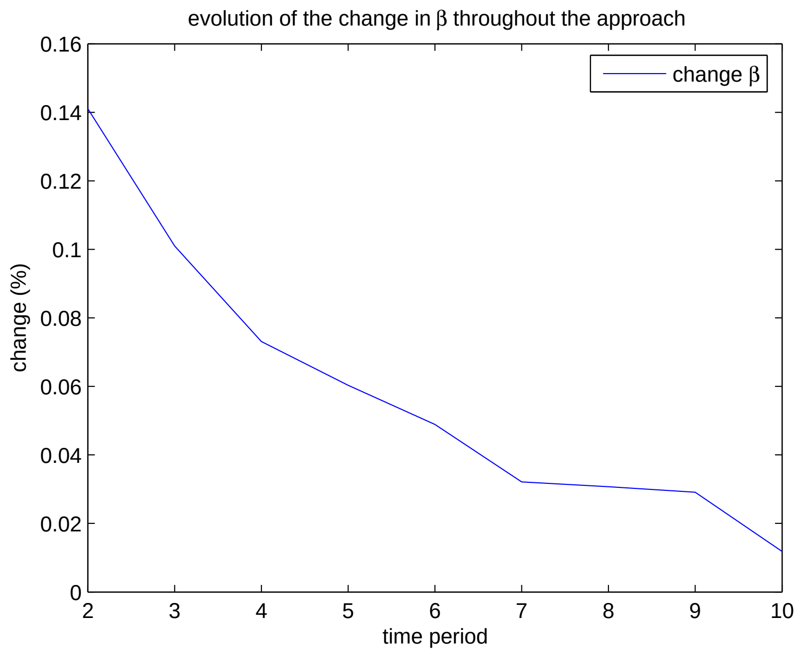


Figure B-10: Evolution of the Change in Prices for the Retail Example ($T = 10$)

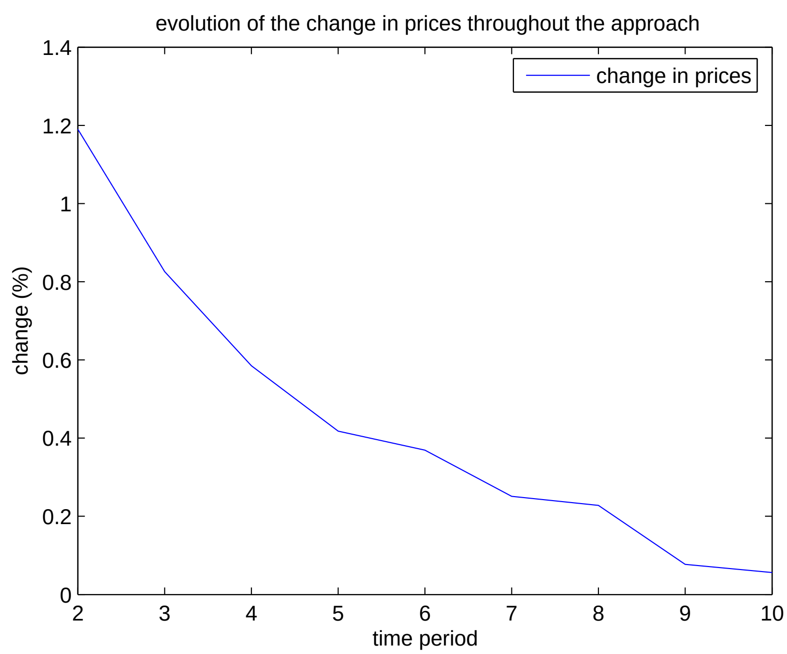


Figure B-11: Evolution of the Change in Demands for the Retail Example ($T = 10$)

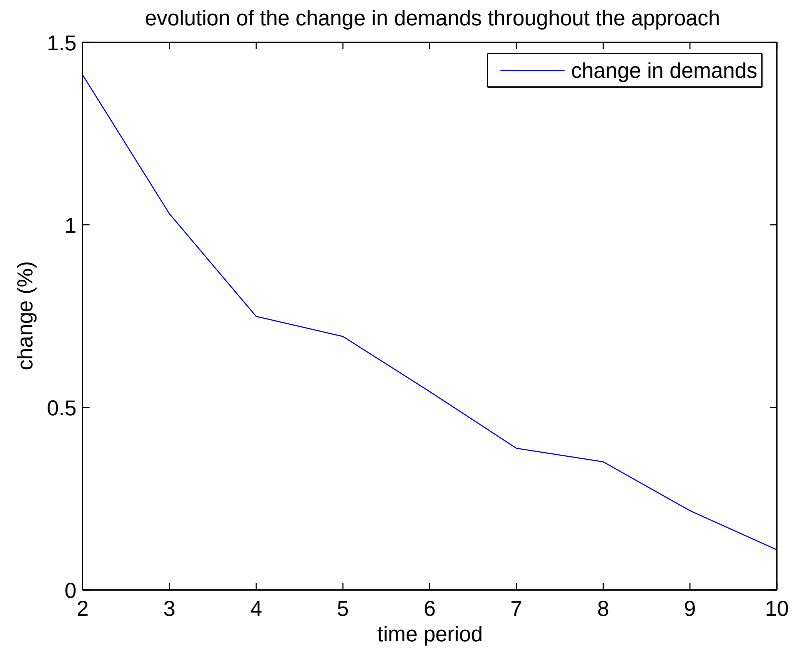


Figure B-12: Evolution of the Change in Parameters in Equilibrium

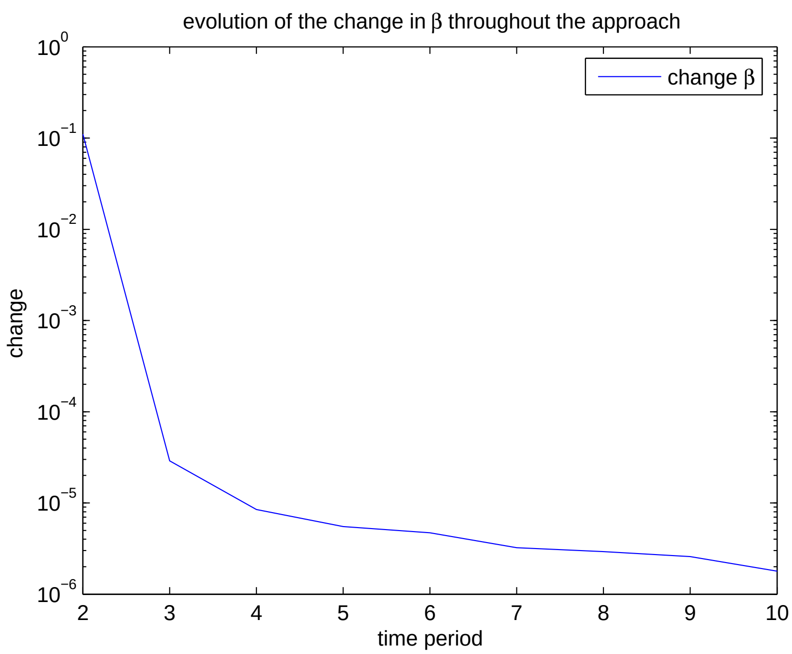


Figure B-13: Evolution of the Change in Prices in Equilibrium

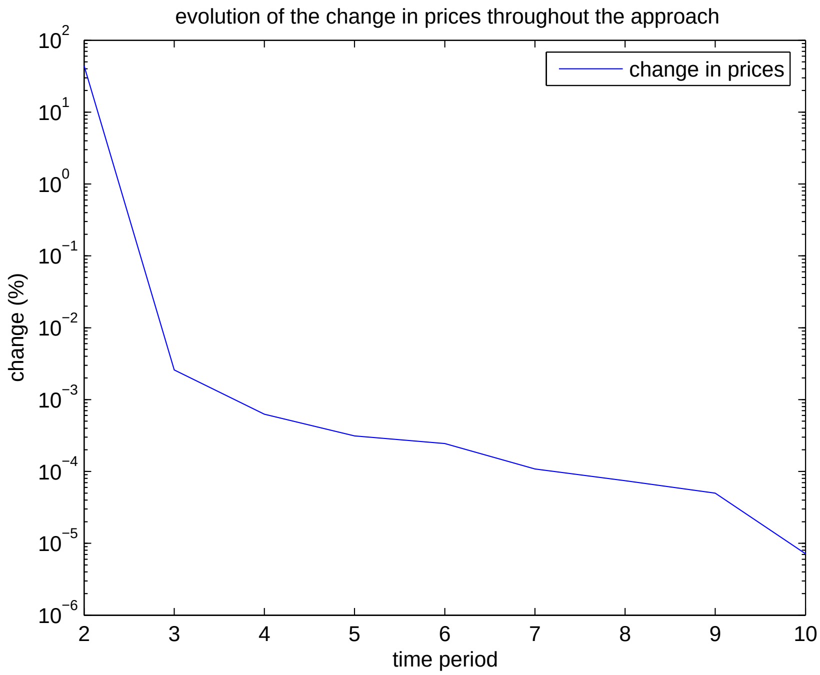


Figure B-14: Evolution of the Change in Demands in Equilibrium

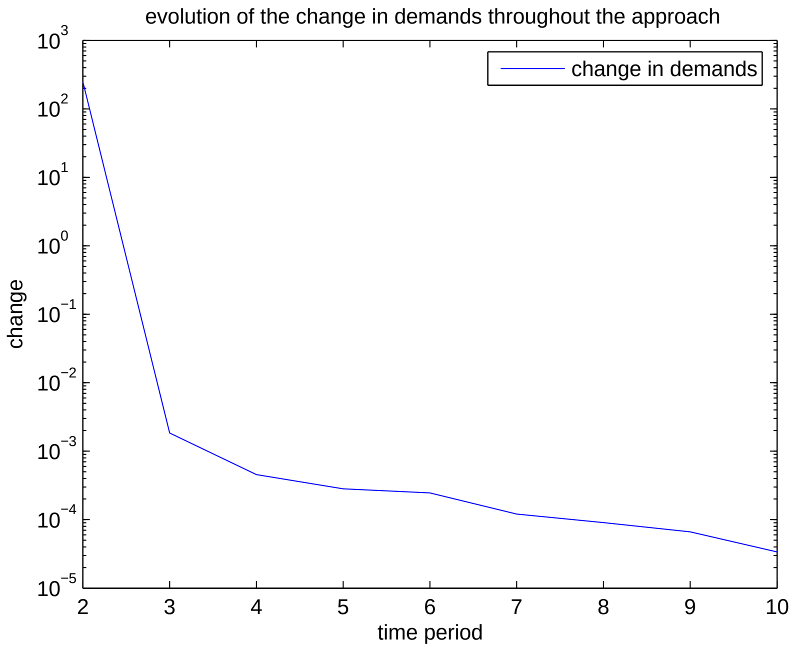


Figure B-15: Comparison of the Learning Speeds as a Function of the Number of Firms

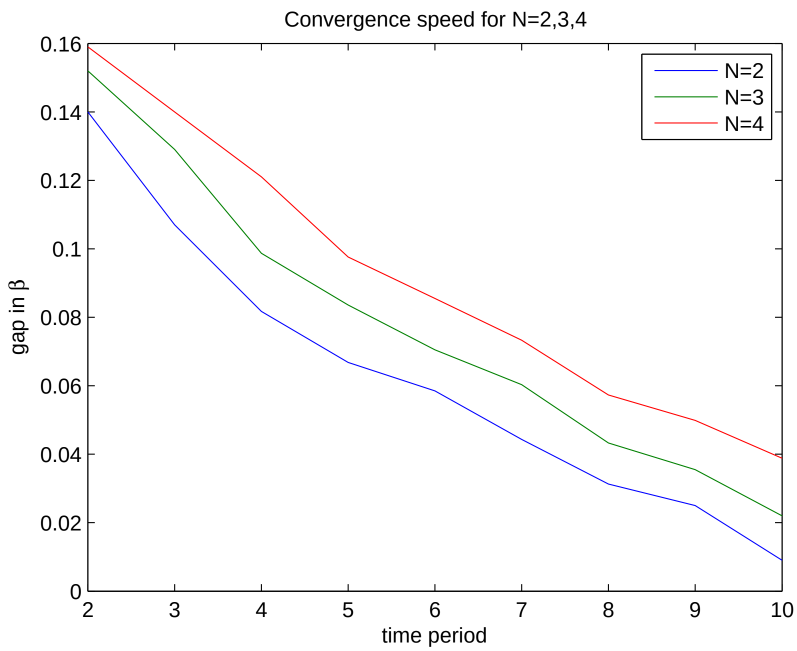
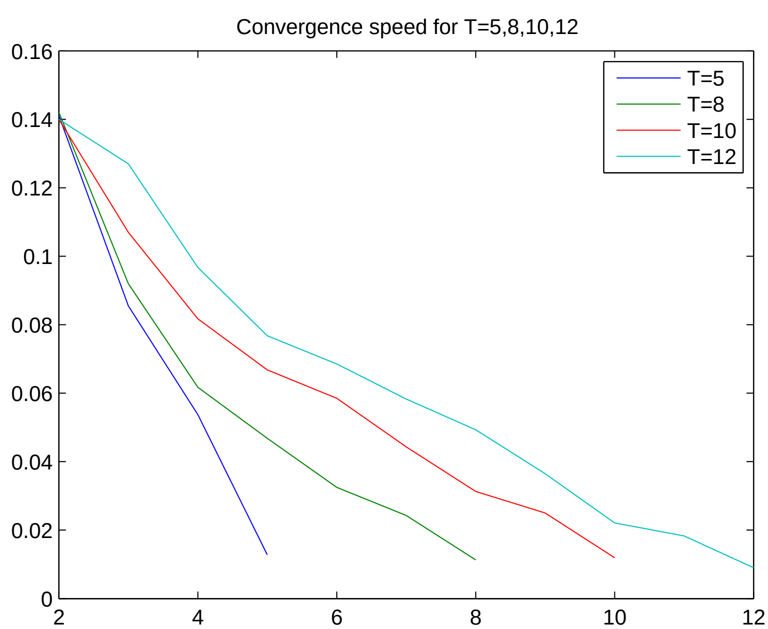
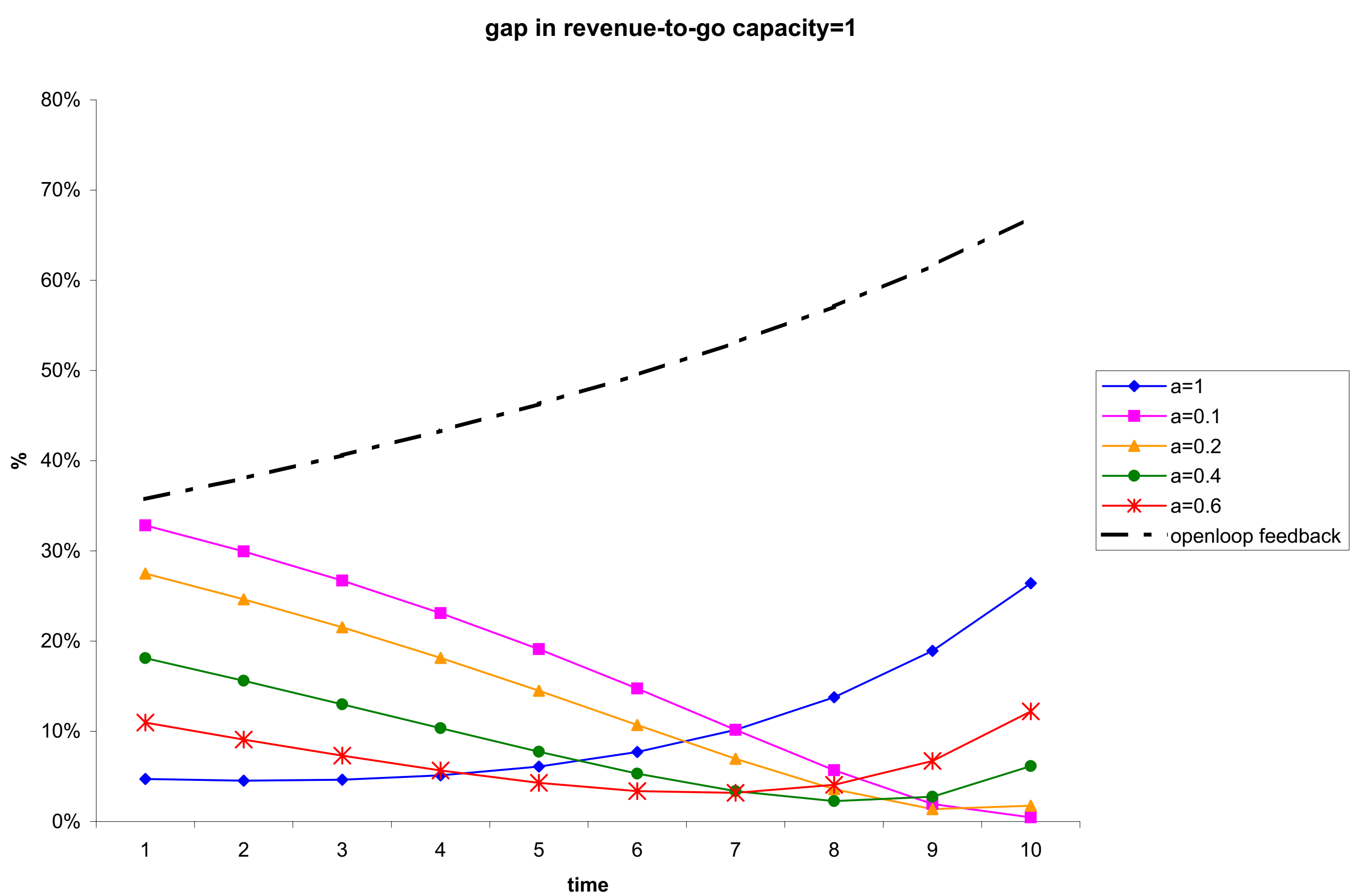
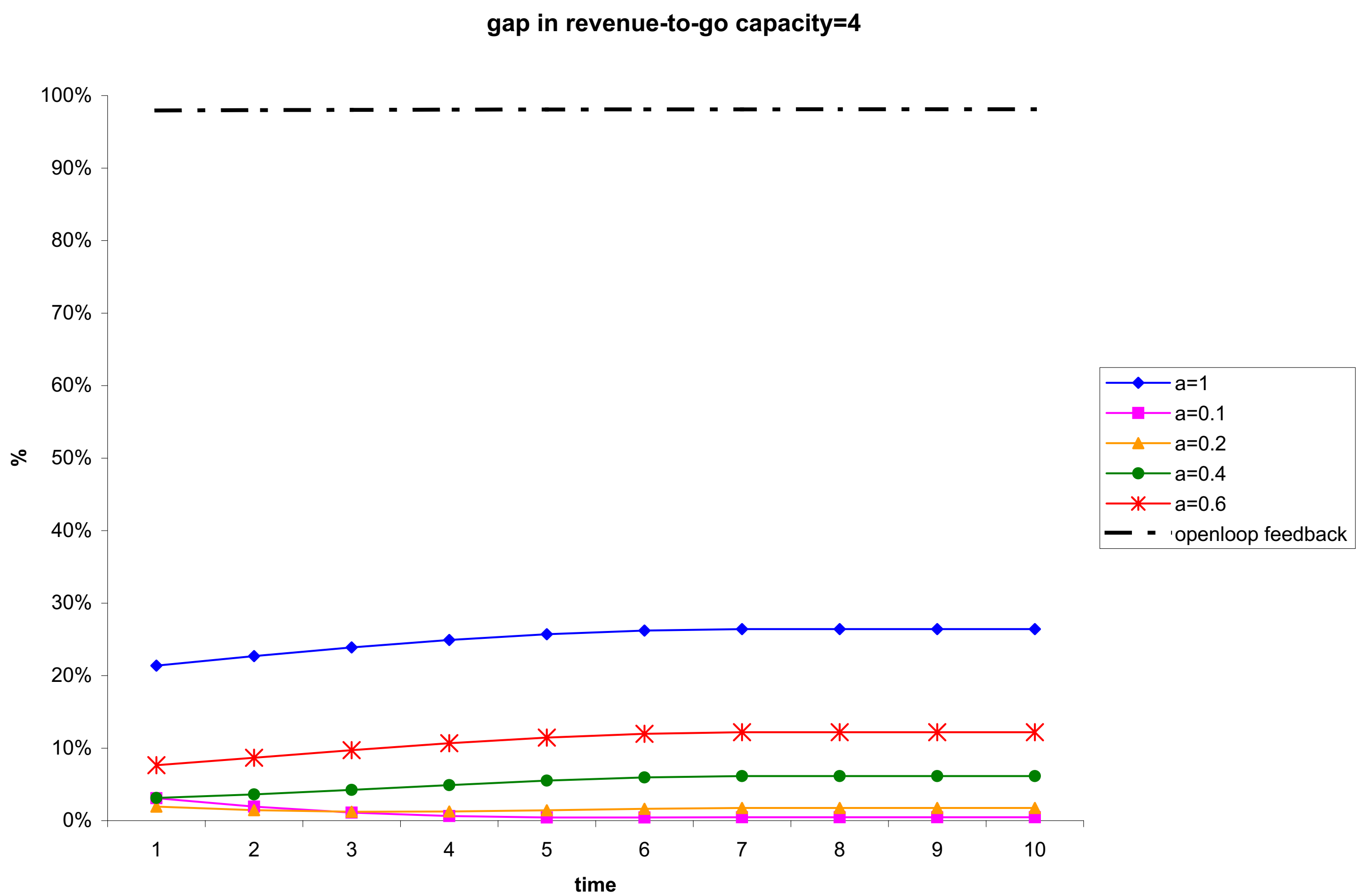


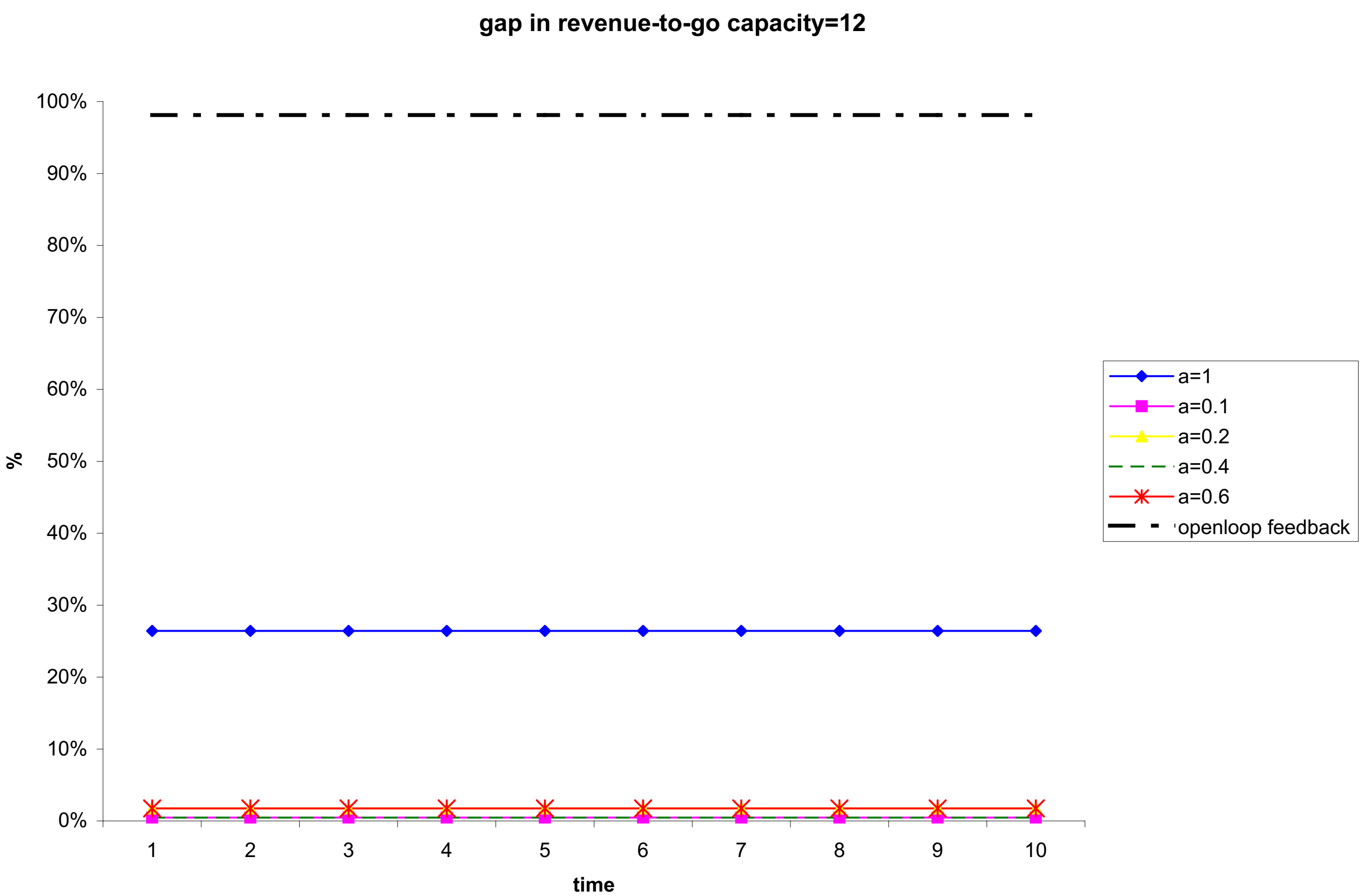
Figure B-16: Comparison of the Learning Speeds as a Function of the Time Horizon

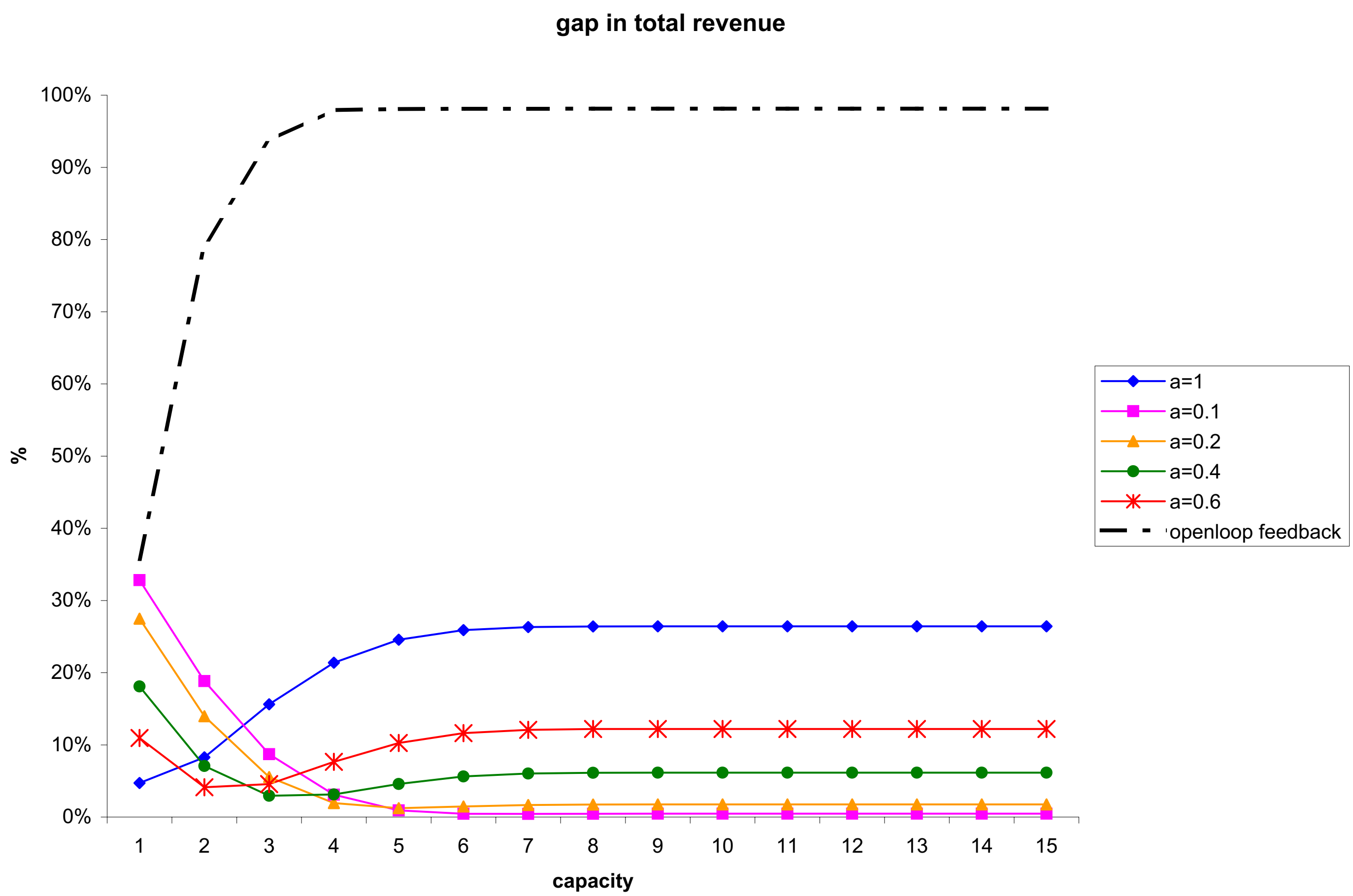


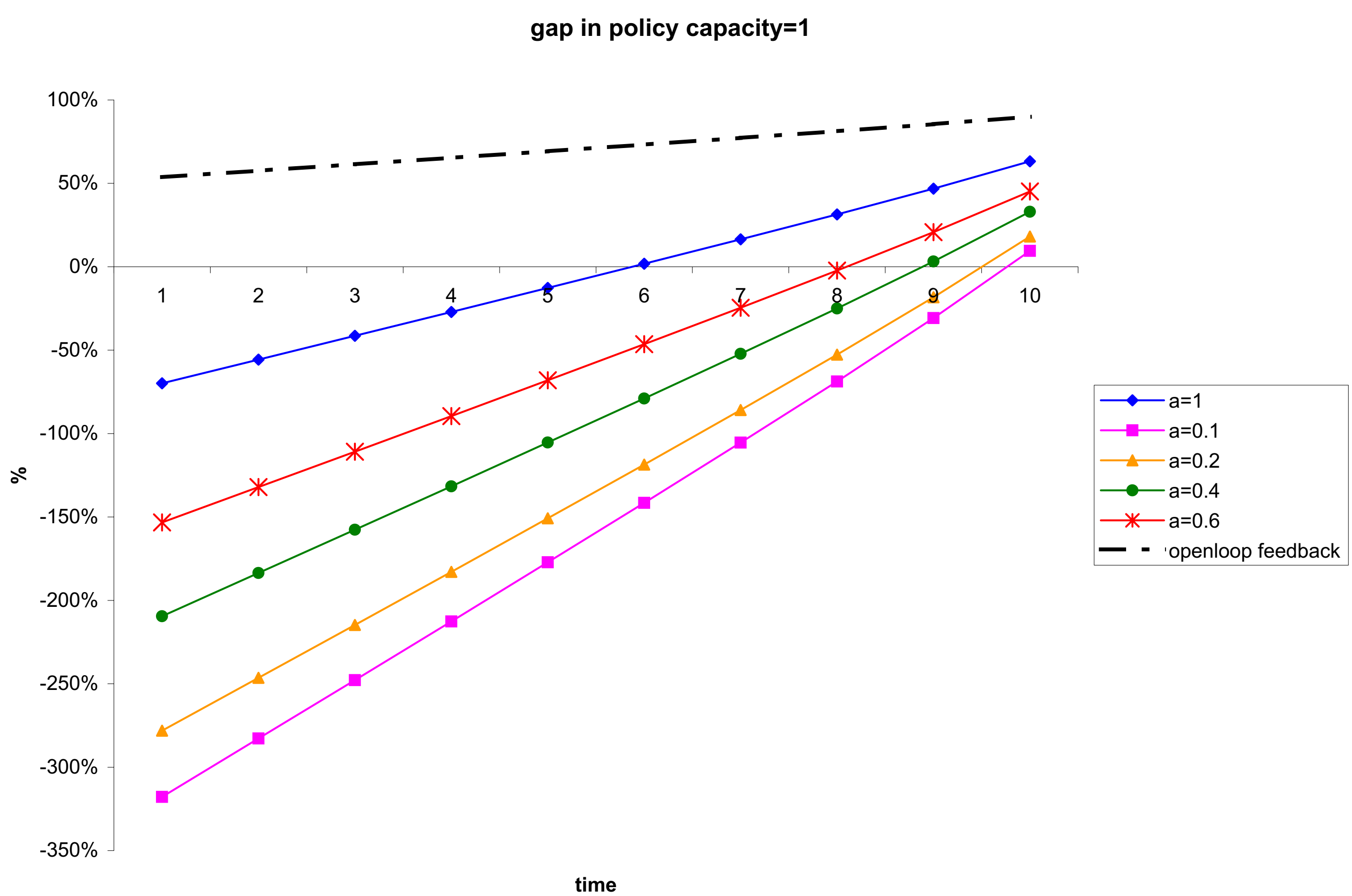

gap in revenue-to-go capacity=1
80%
70%
60%
50%
40%
30%
20%
10%
0%
%
1
2
3
4
5
6
7
8
9
10
time
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback



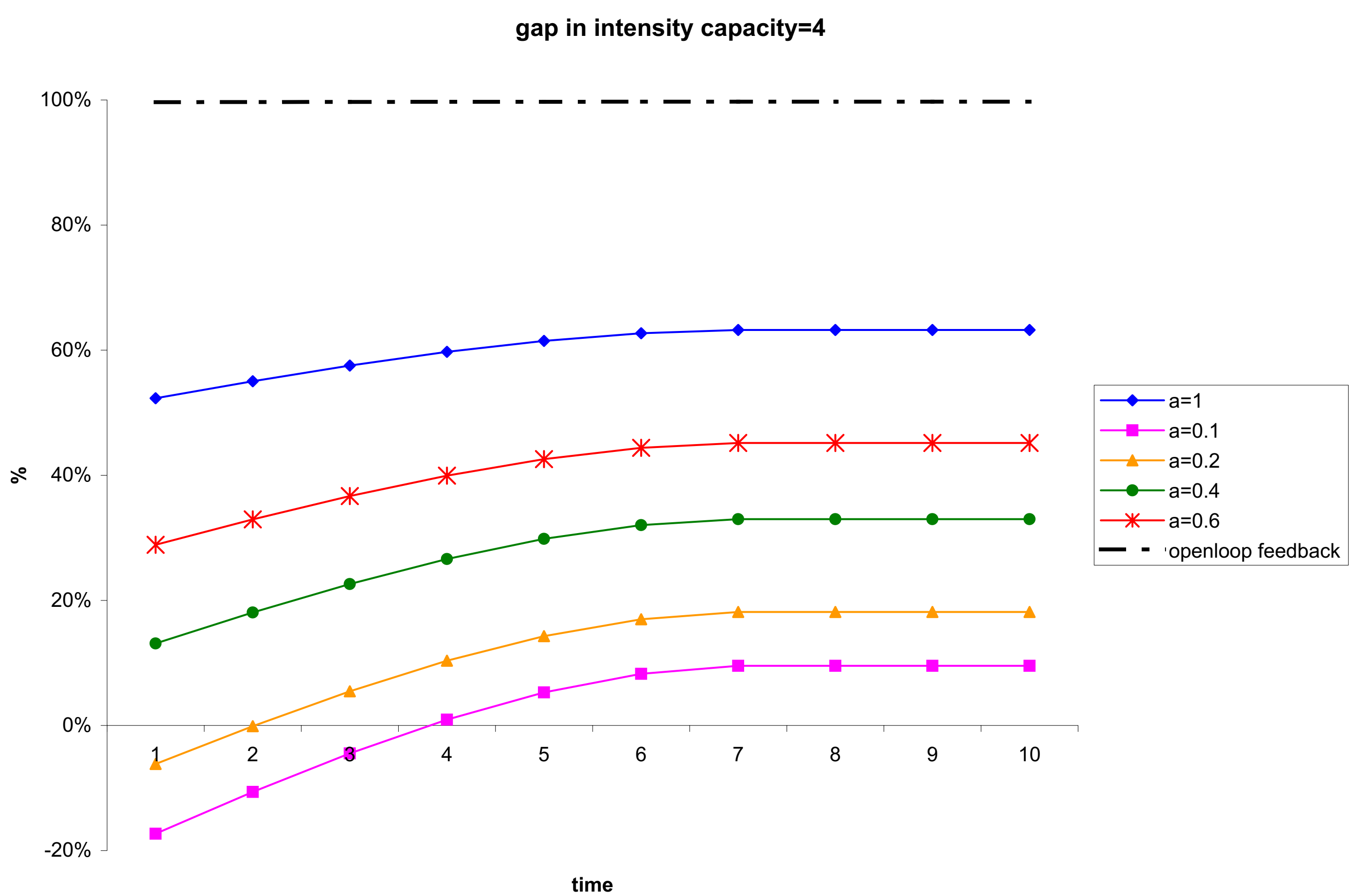
gap in revenue-to-go capacity=4
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
%
1
2
3
4
5
6
7
8
9
10
time
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback


gap in revenue-to-go capacity=12
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
%
1
2
3
4
5
6
7
8
9
10
time
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback

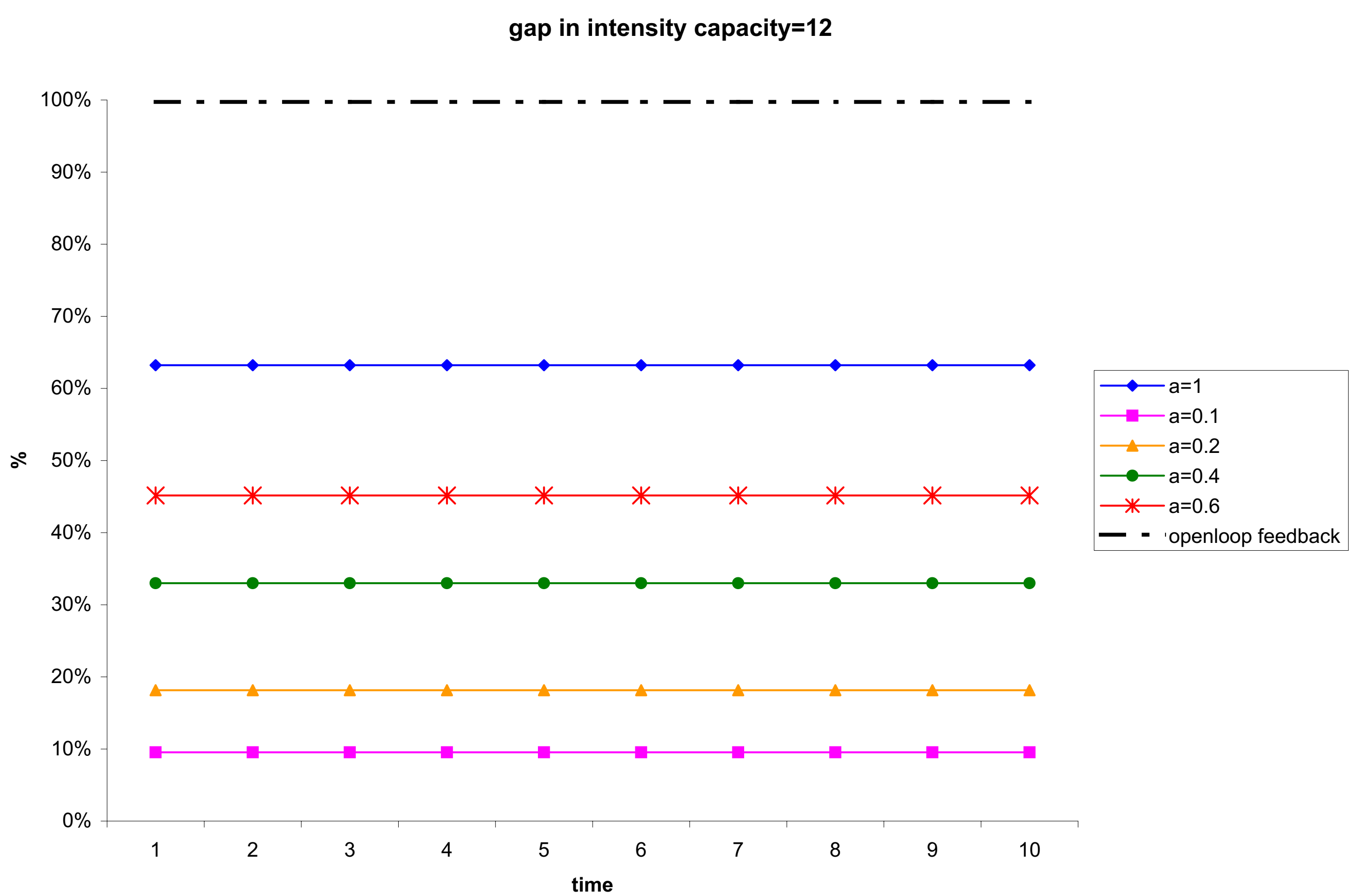

gap in total revenue
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
capacity
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback

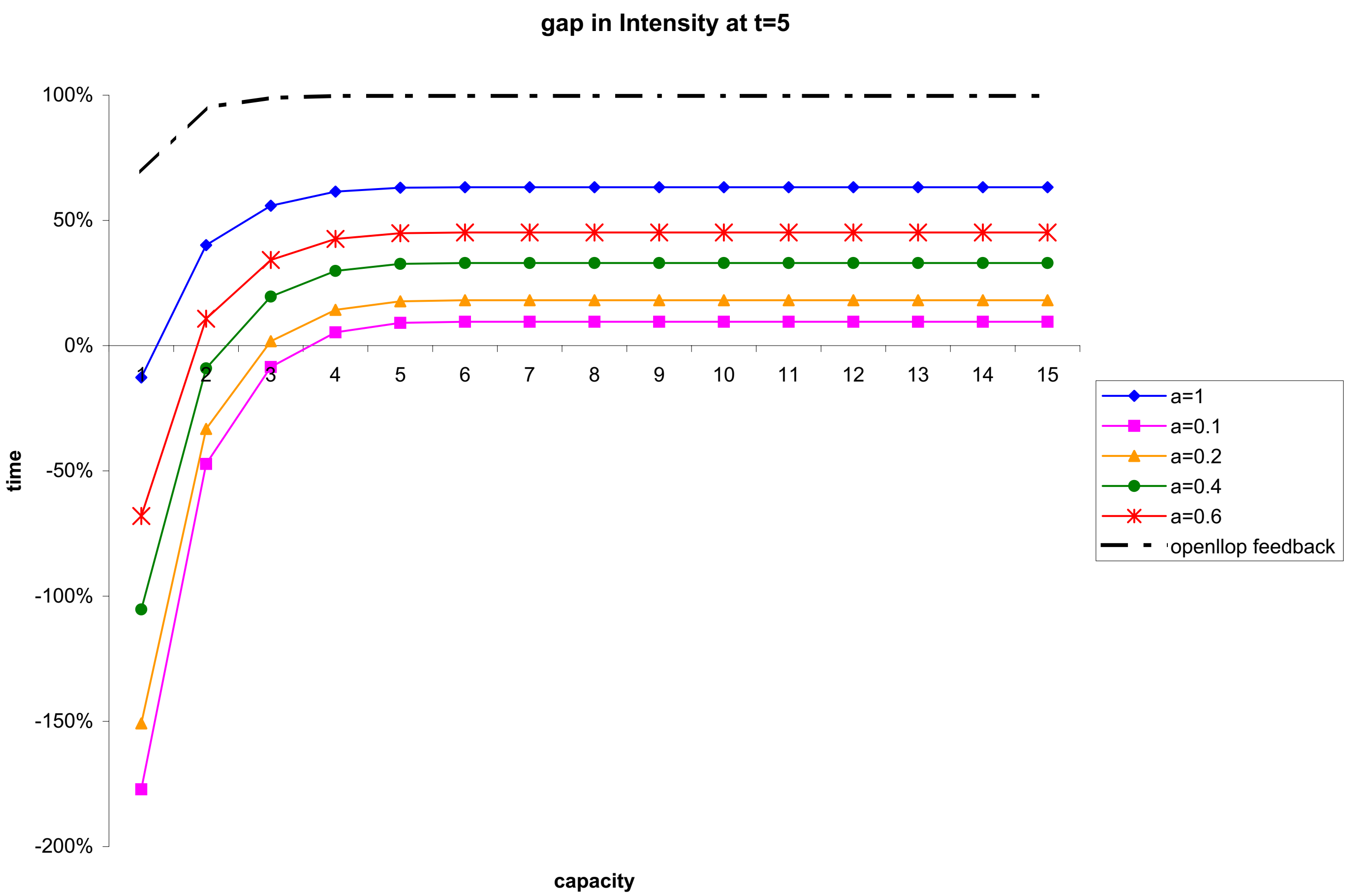


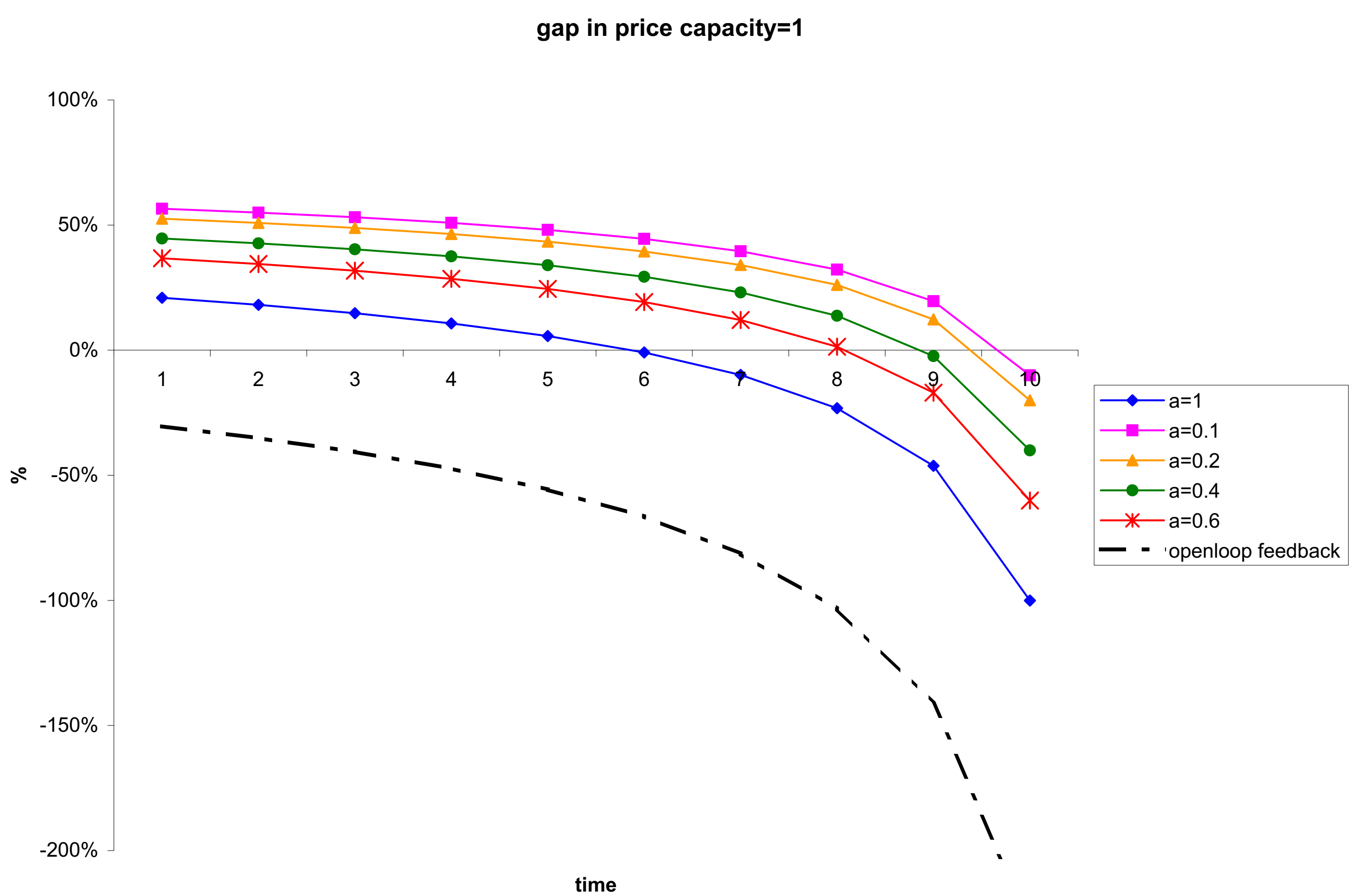
gap in policy capacity=1
100%
50%
0%
-50%
-100%
-150%
-200%
-250%
-300%
-350%
%
1
2
3
4
5
6
7
8
9
10
time
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback

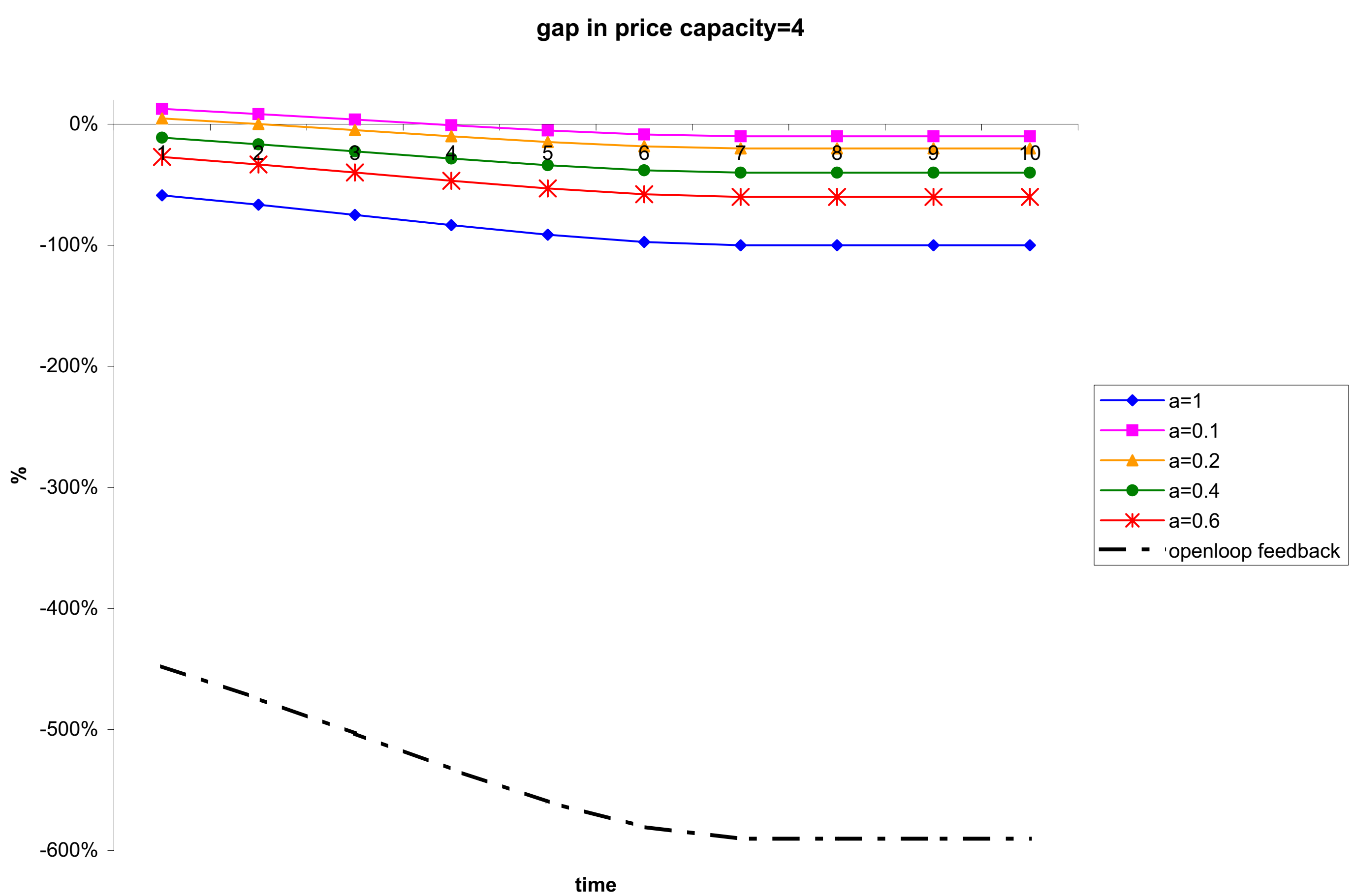


gap in intensity capacity=4
100%
80%
60%
40%
20%
0%
-20%
%
1
2
3
4
5
6
7
8
9
10
time
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback

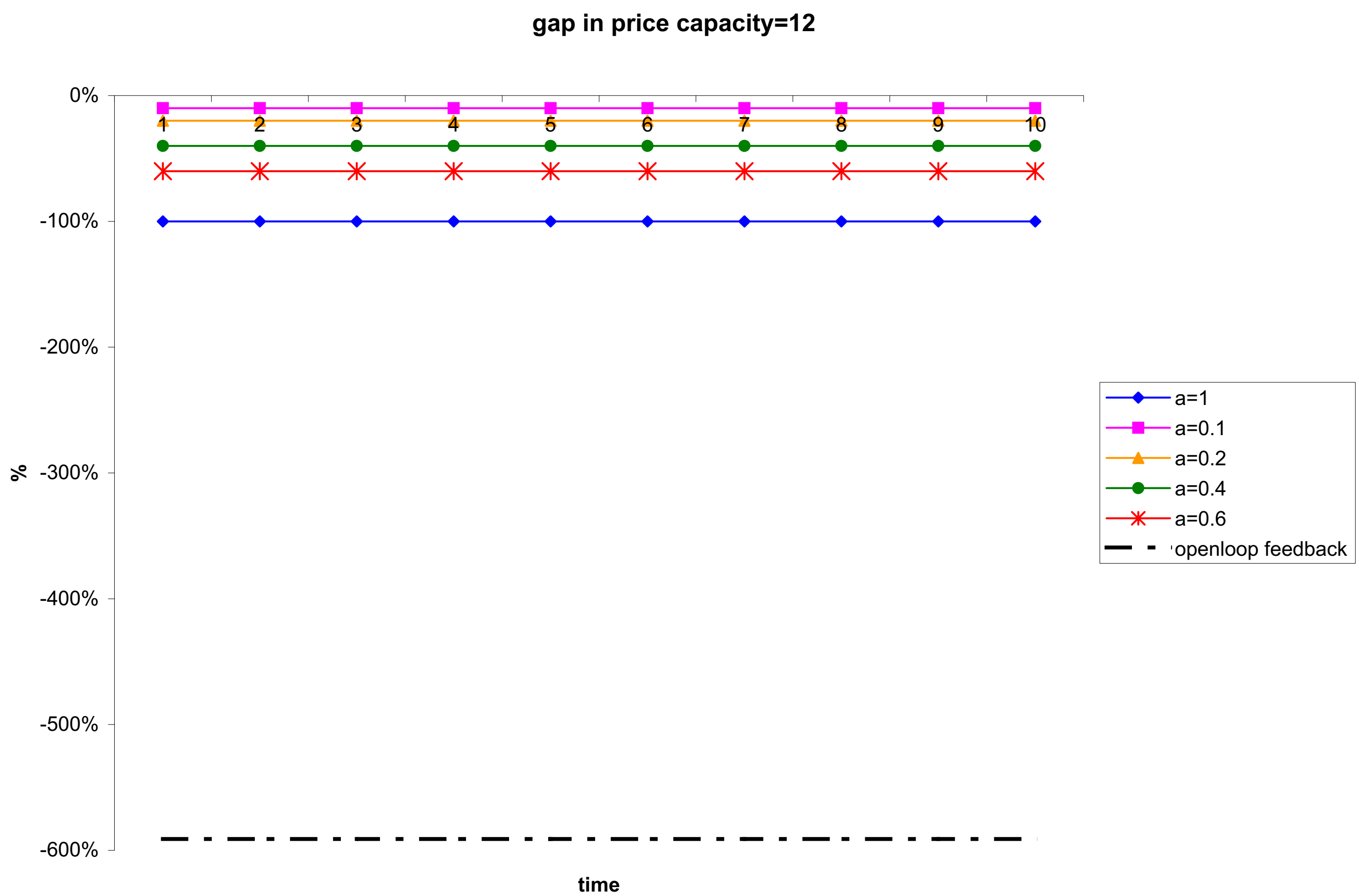

gap in intensity capacity=12
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
%
1
2
3
4
5
6
7
8
9
10
time
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback

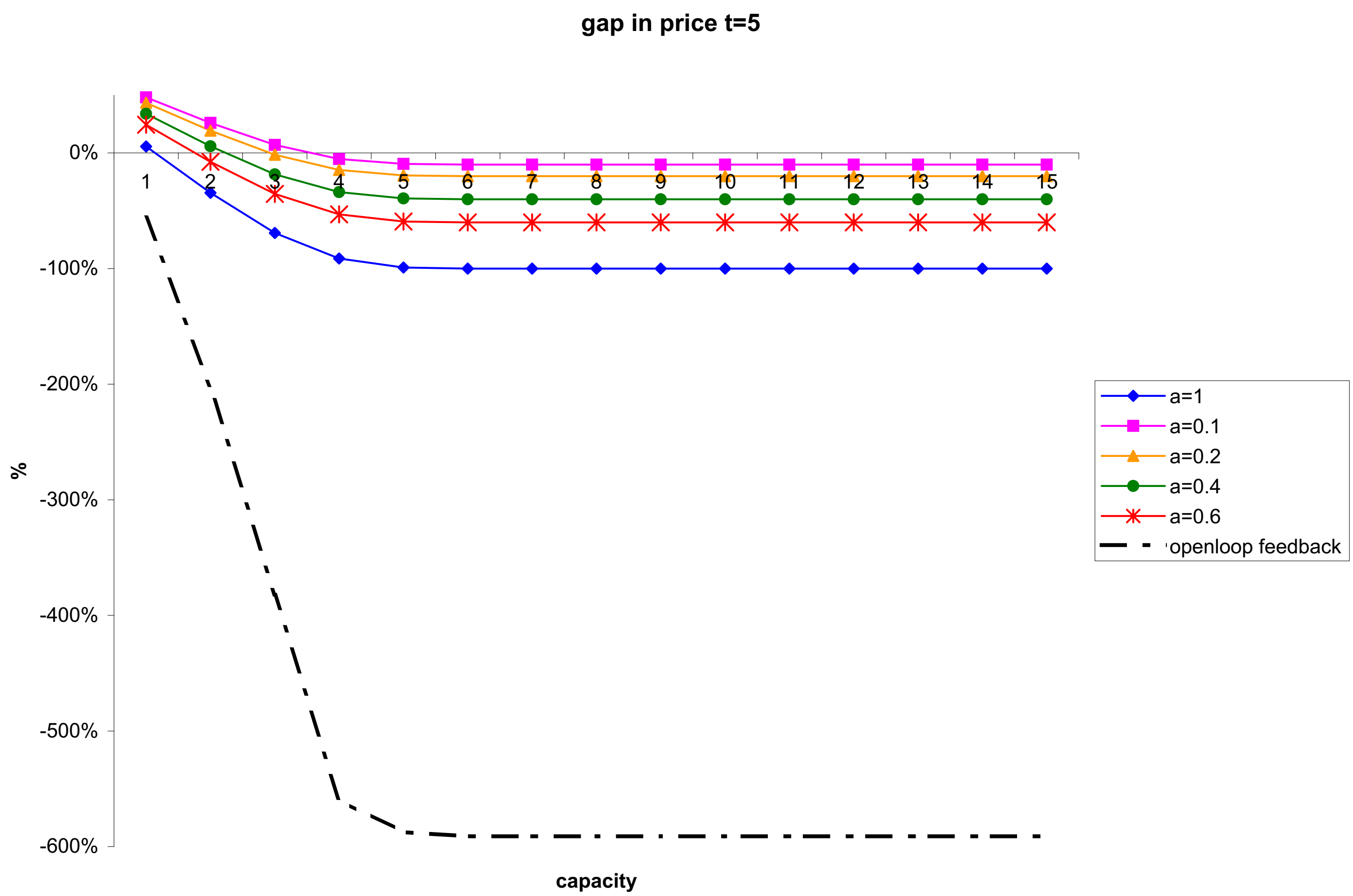

gap in Intensity at t=5
100%
50%
0%
-50%
-100%
-150%
-200%
time
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
capacity
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openllop feedback


gap in price capacity=1
100%
50%
0%
-50%
-100%
-150%
-200%
%
1
2
3
4
5
6
7
8
9
10
time
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback

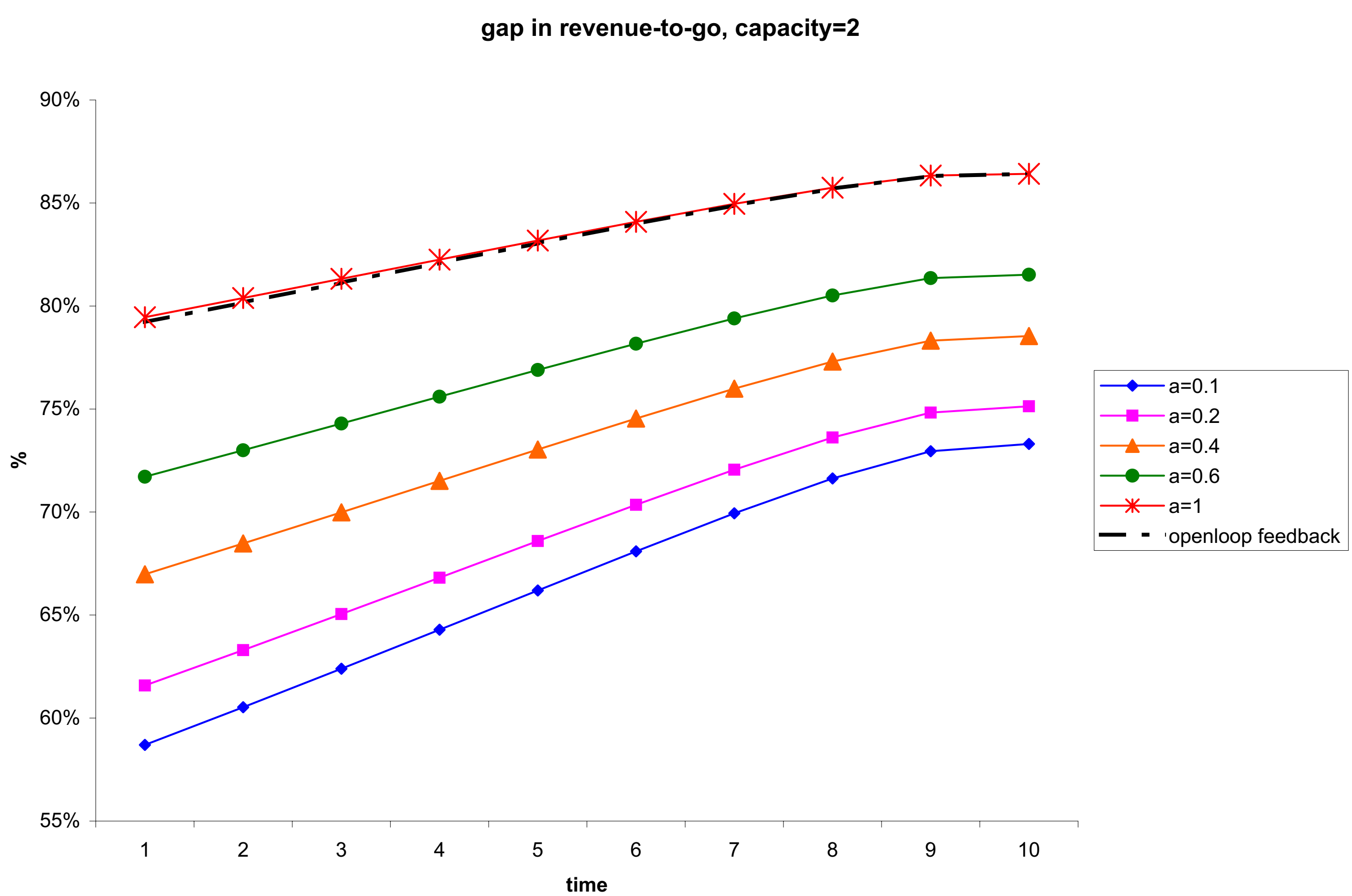


gap in price capacity=4
0%
-100%
-200%
-300%
-400%
-500%
-600%
%
1
2
3
4
5
6
7
8
9
10
time
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback

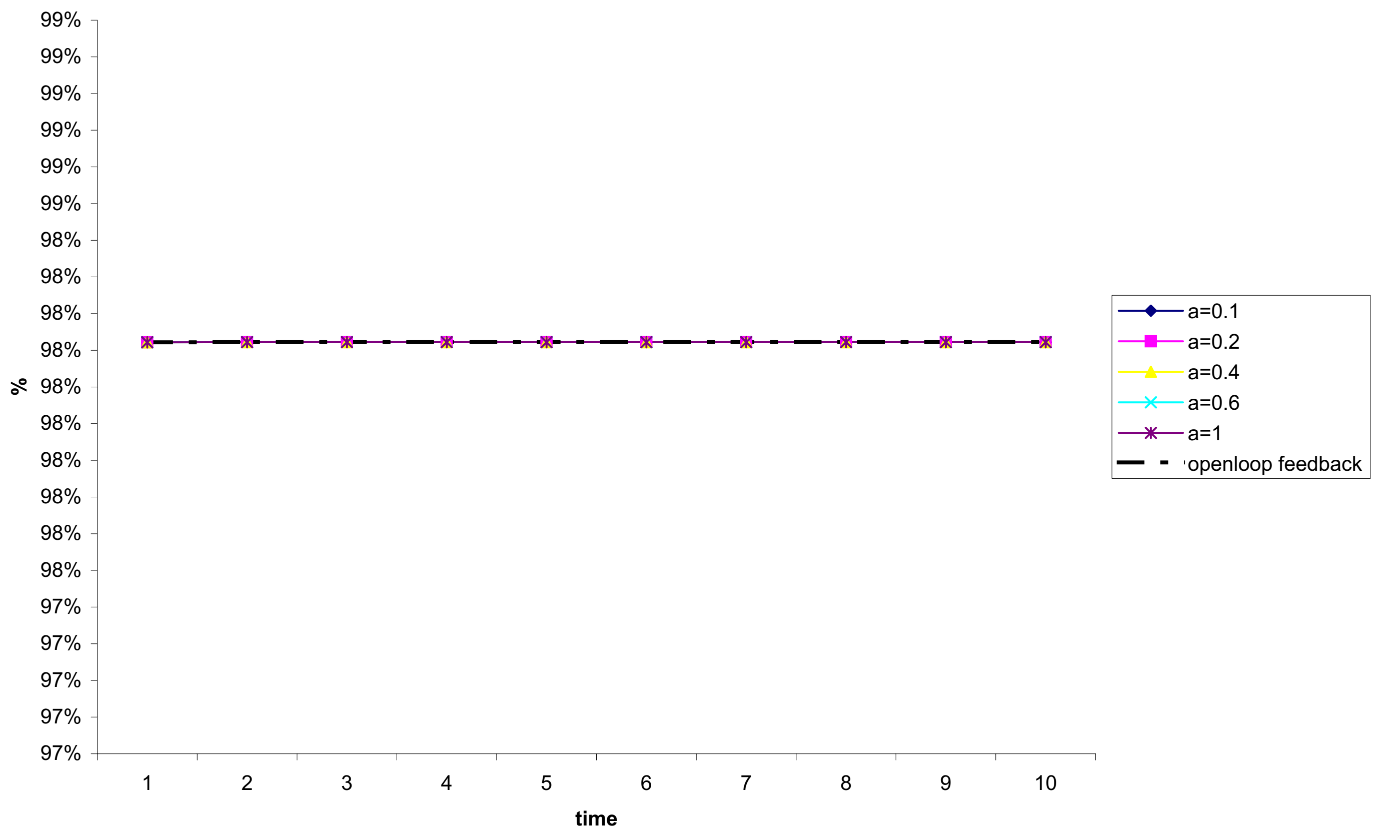

gap in price capacity=12
0%
-100%
-200%
-300%
-400%
-500%
-600%
%
1
2
3
4
5
6
7
8
9
10
time
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback

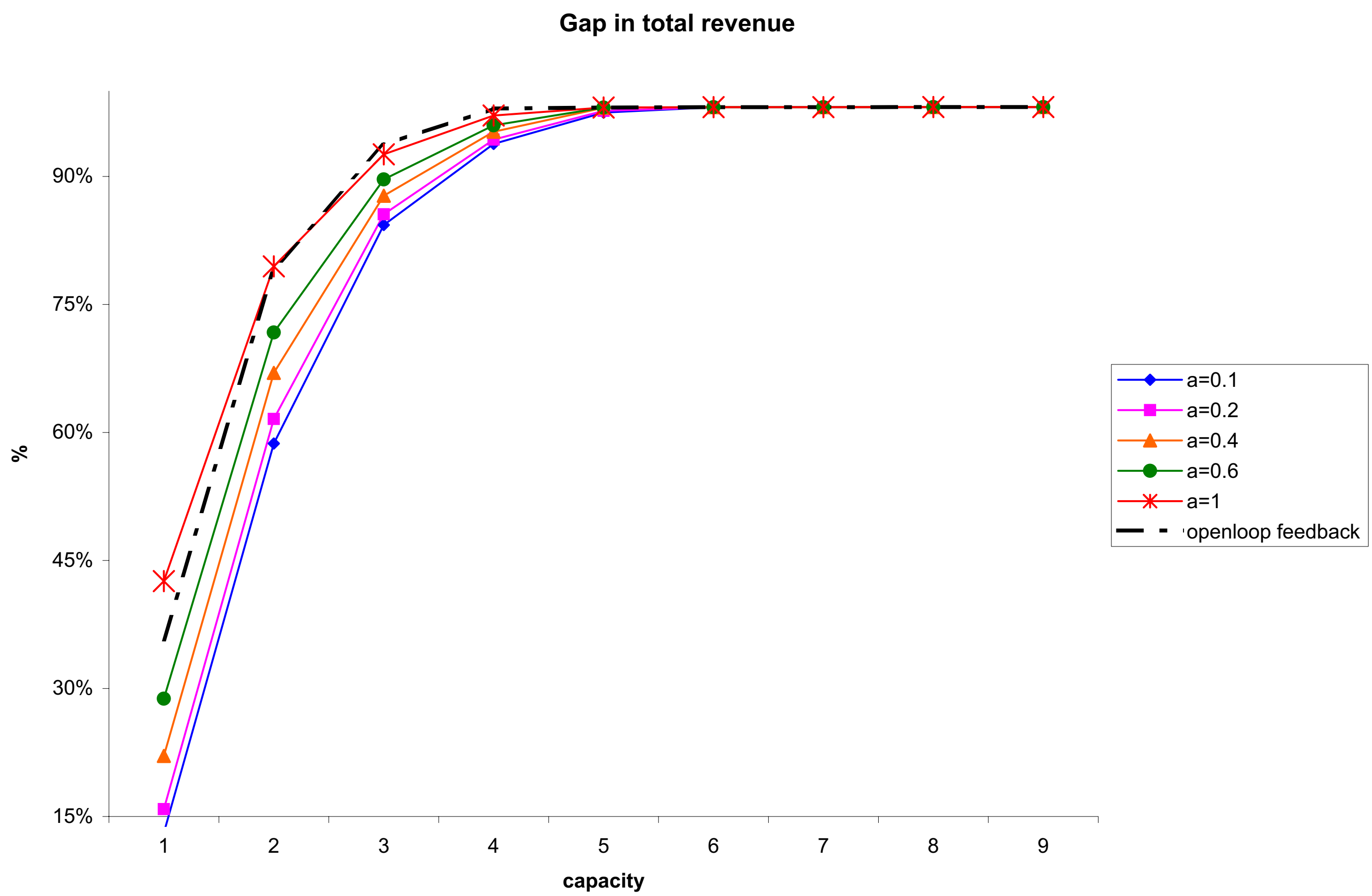


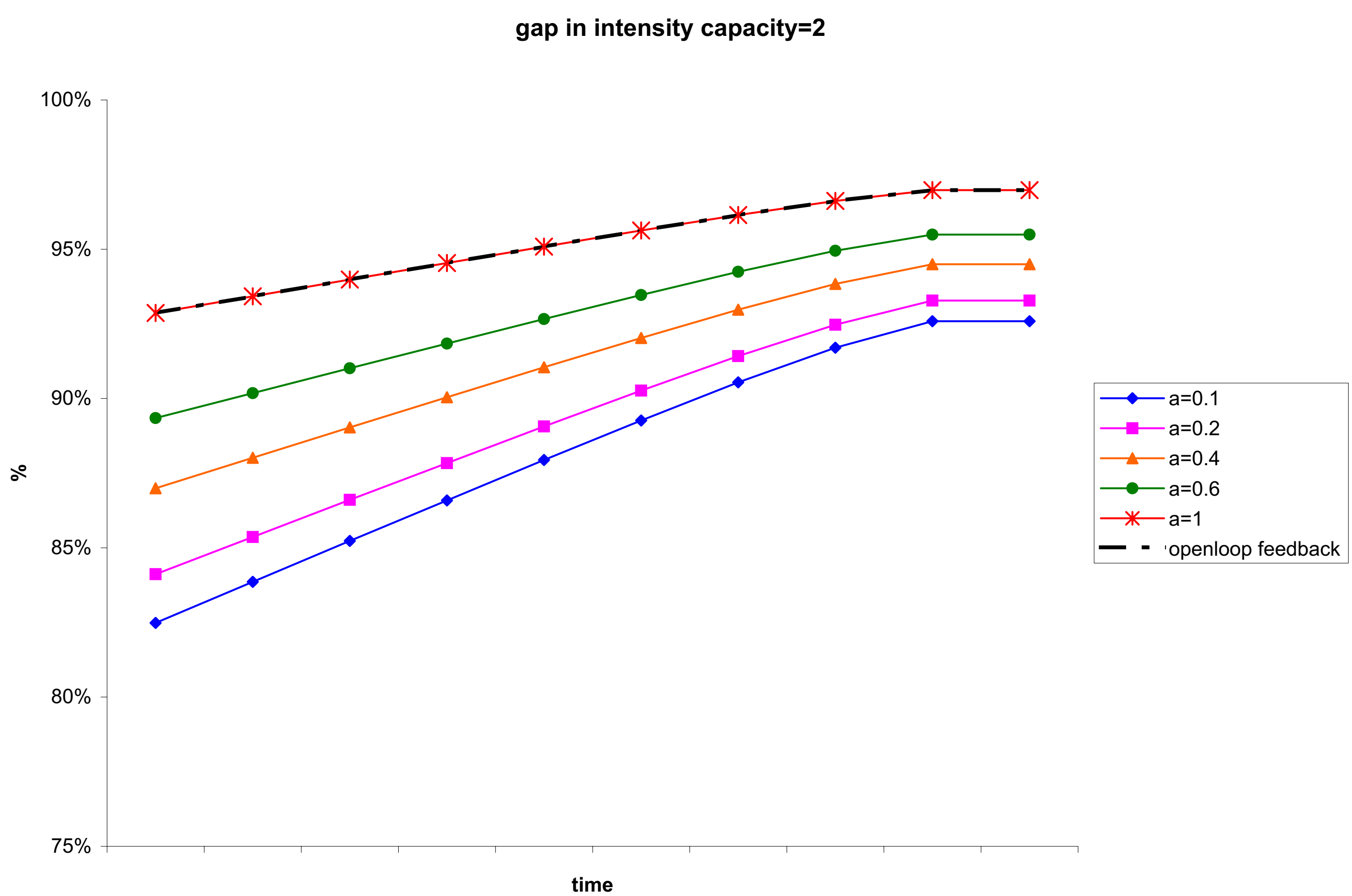
gap in price t=5
0%
-100%
-200%
-300%
-400%
-500%
-600%
%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
capacity
a=1
a=0.1
a=0.2
a=0.4
a=0.6
openloop feedback

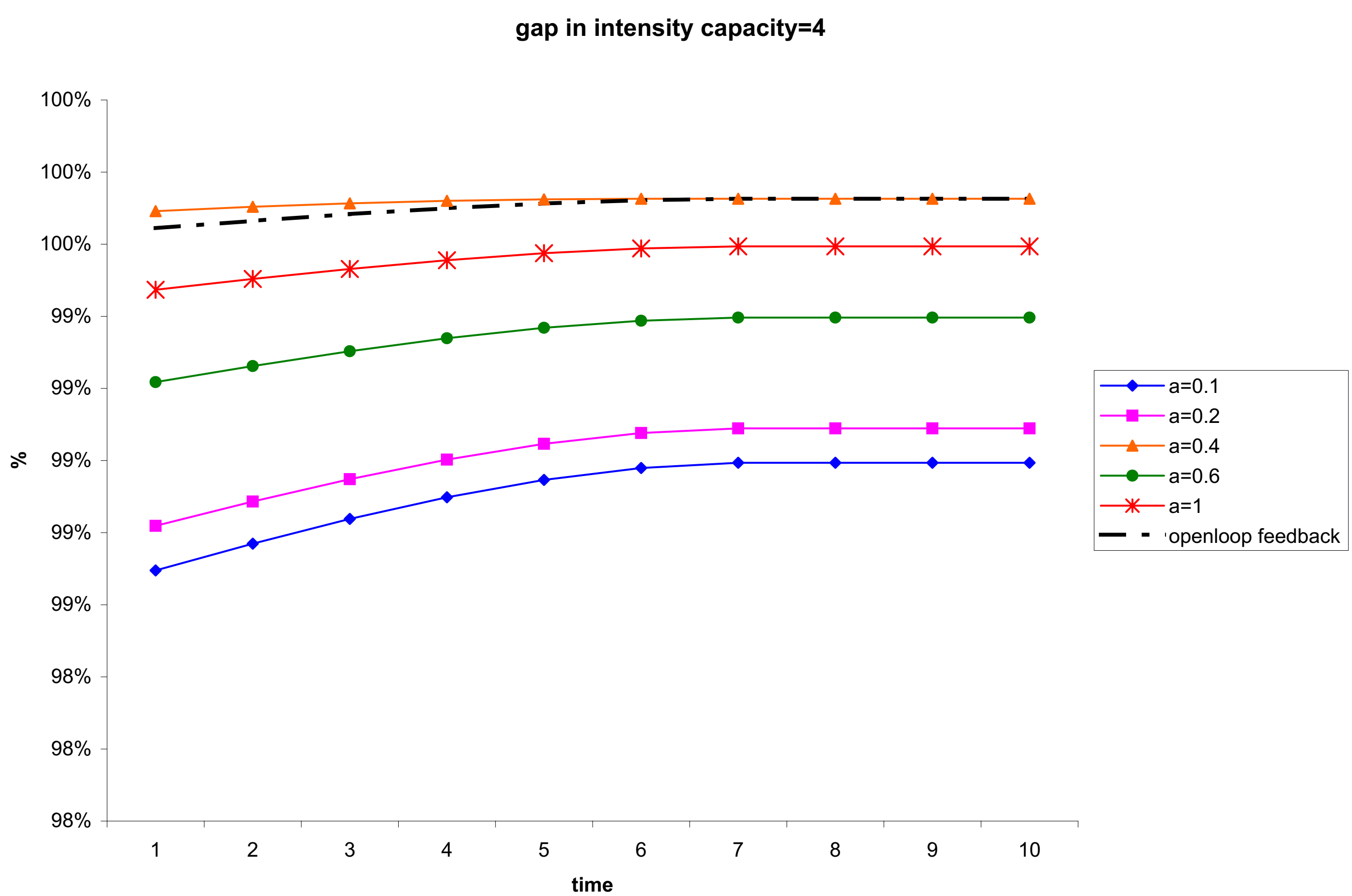

gap in revenue-to-go, capacity=2
90%
85%
80%
75%
70%
65%
60%
55%
%
1
2
3
4
5
6
7
8
9
10
time
a=0.1
a=0.2
a=0.4
a=0.6
a=1
openloop feedback

gap in revenue-to-go capacity=8

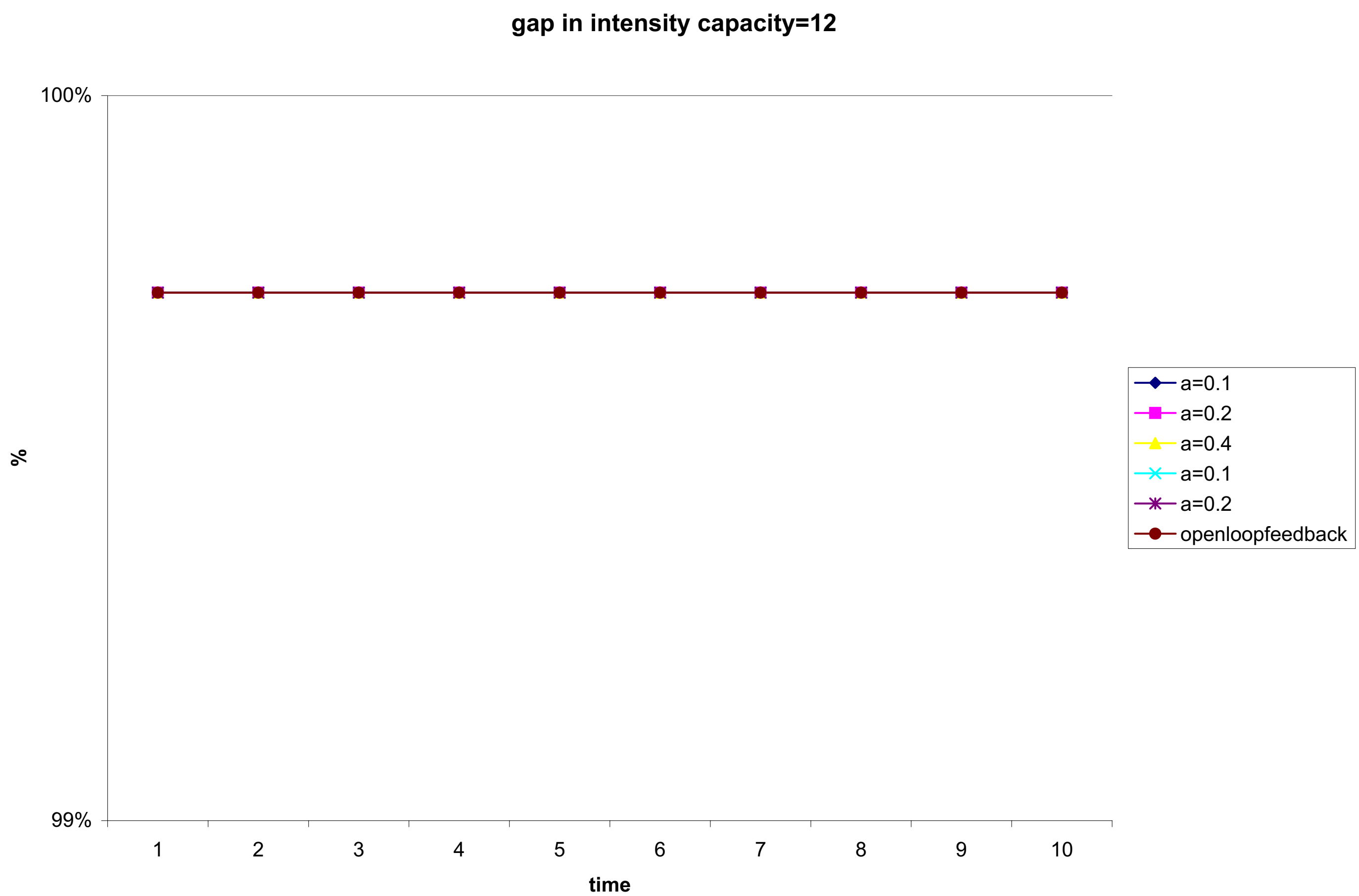



Gap in total revenue
90%
75%
60%
45%
30%
15%
%
1
2
3
4
5
6
7
8
9
capacity
a=0.1
a=0.2
a=0.4
a=0.6
a=1
openloop feedback

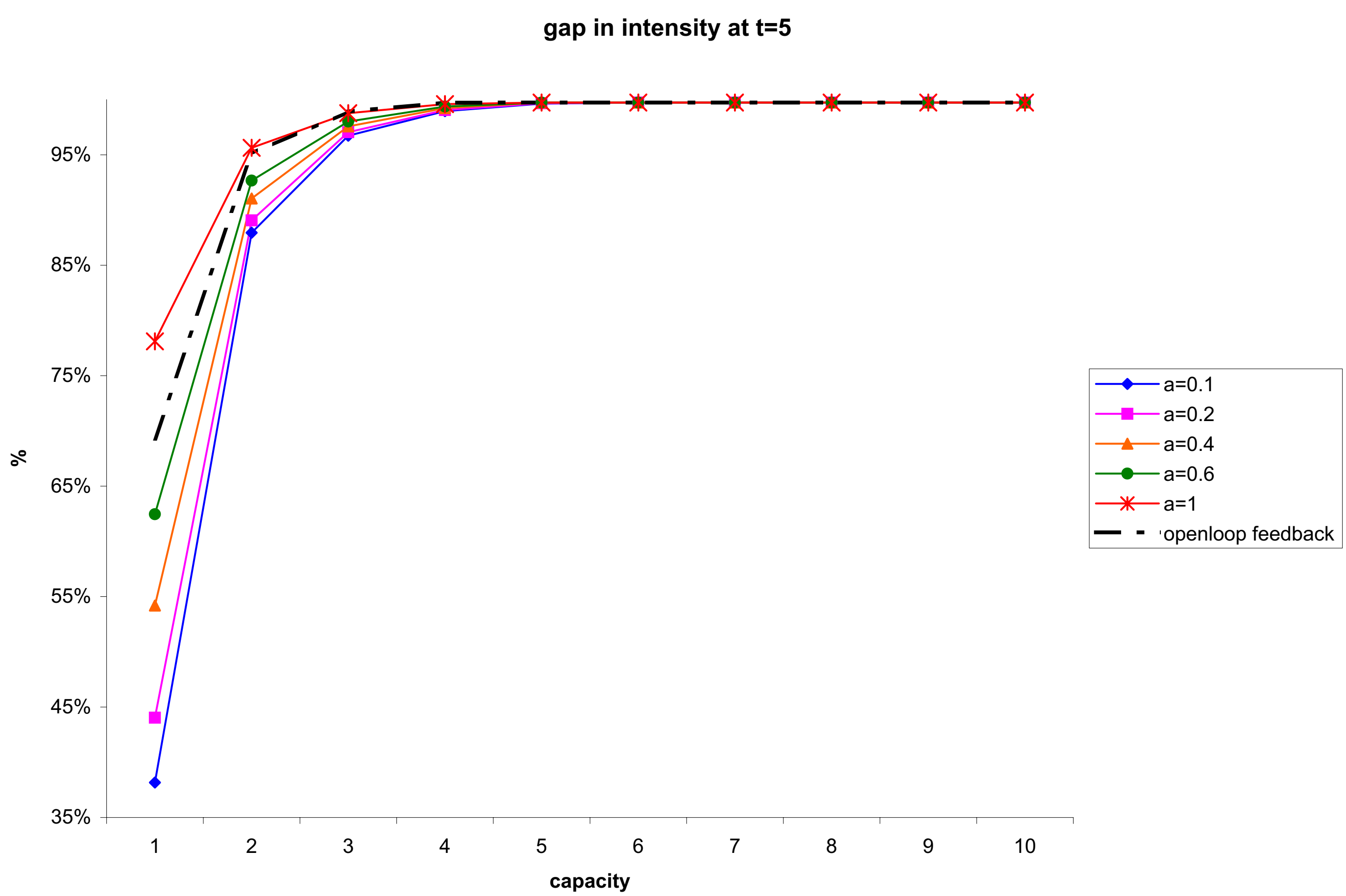

gap in intensity capacity=2
100%
95%
90%
85%
80%
75%
%
time
a=0.1
a=0.2
a=0.4
a=0.6
a=1
openloop feedback



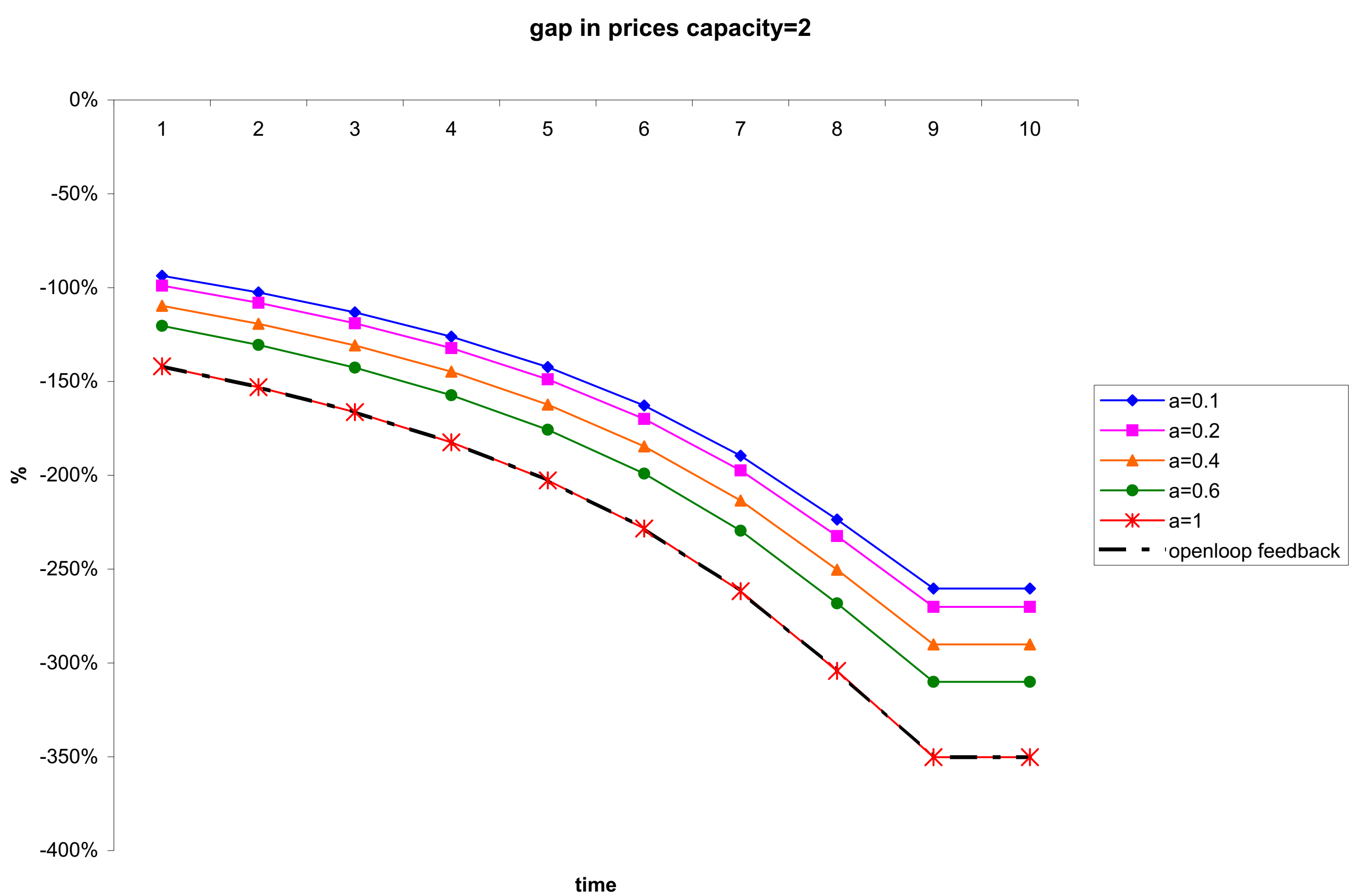
gap in intensity capacity=4
100%
100%
100%
99%
99%
99%
99%
99%
98%
98%
98%
%
1
2
3
4
5
6
7
8
9
10
time
a=0.1
a=0.2
a=0.4
a=0.6
a=1
openloop feedback



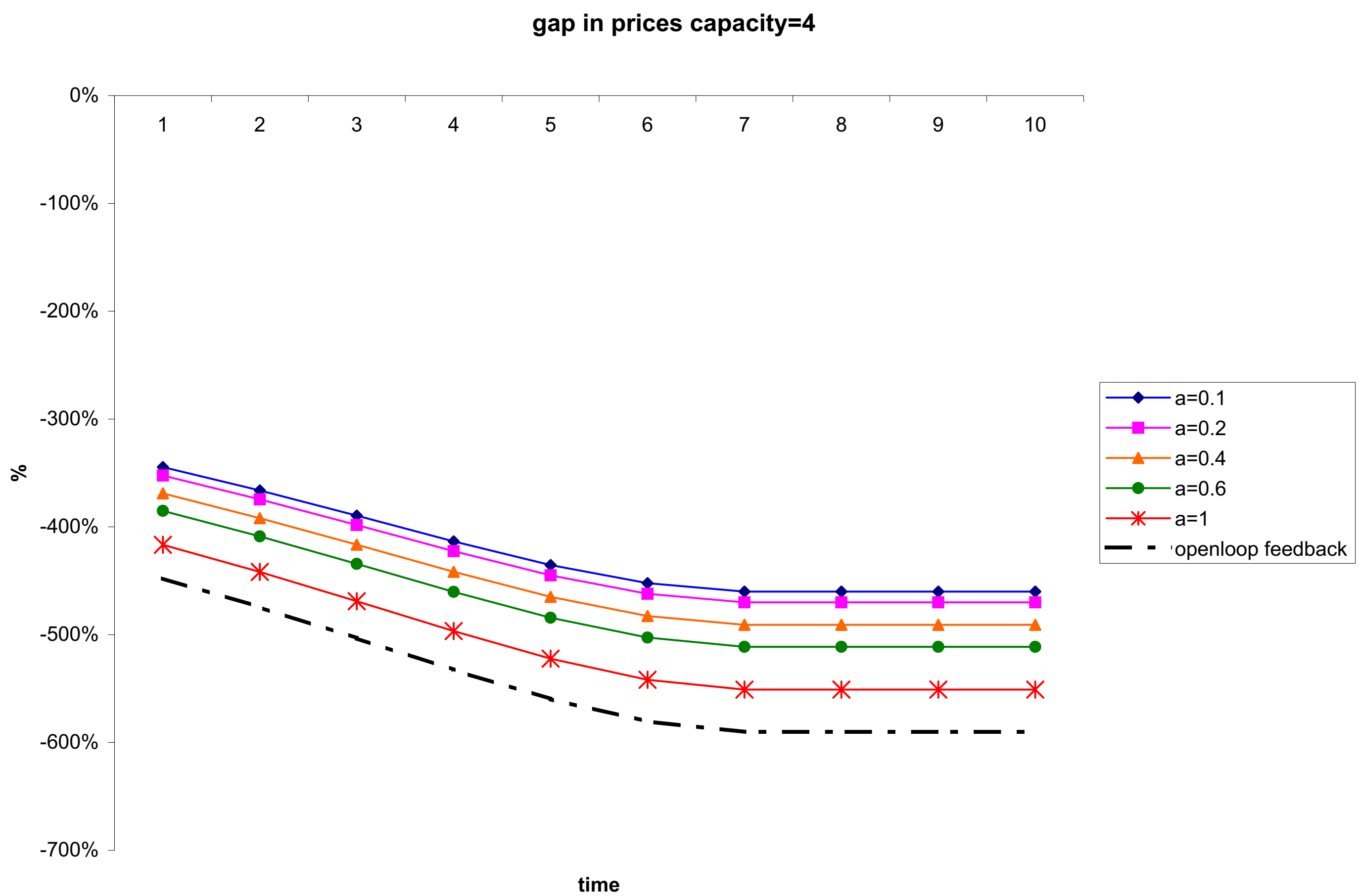
gap in intensity capacity=12
100%
99%
%
1
2
3
4
5
6
7
8
9
10
time
a=0.1
a=0.2
a=0.4
a=0.1
a=0.2
openloopfeedback

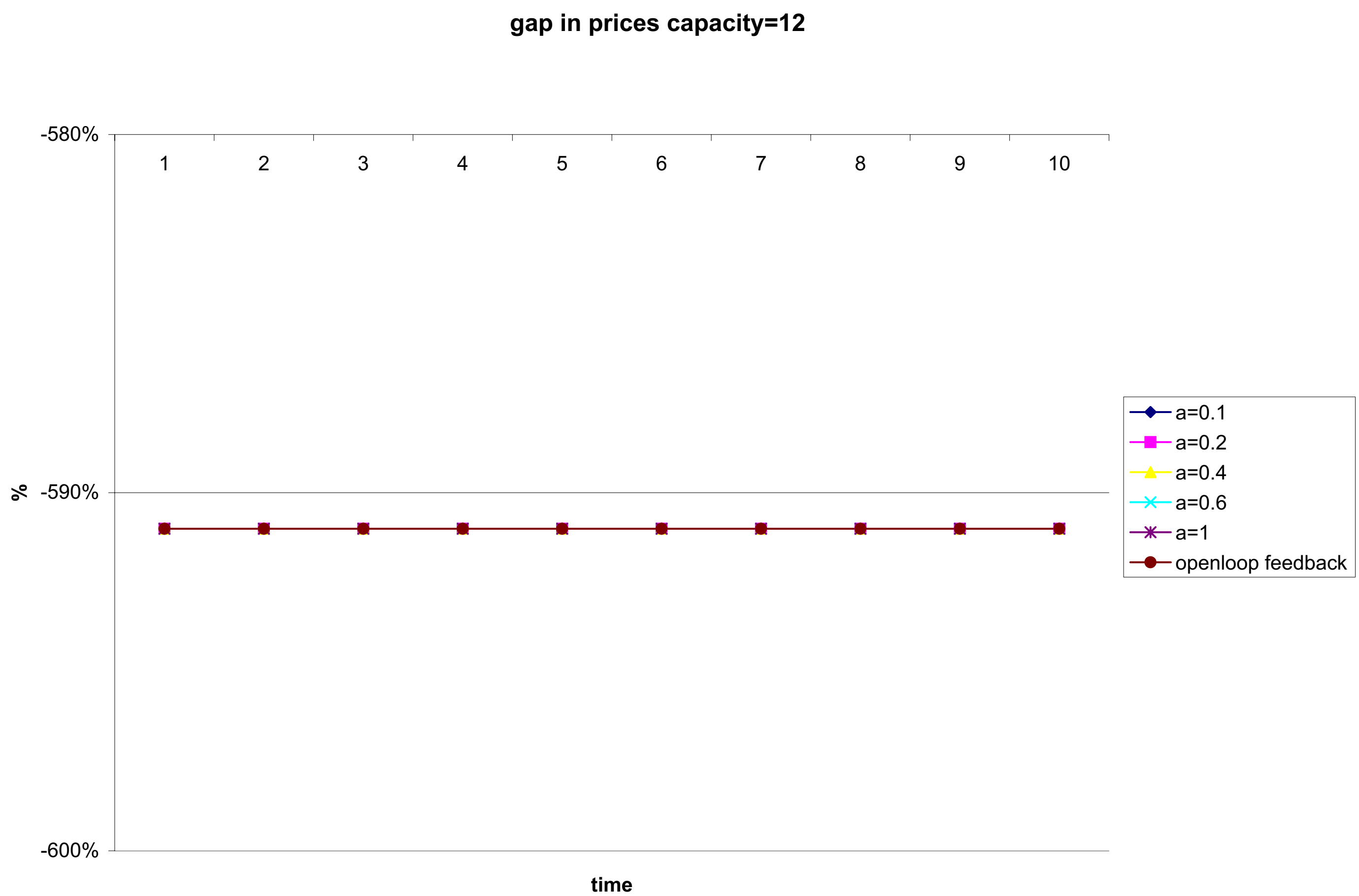


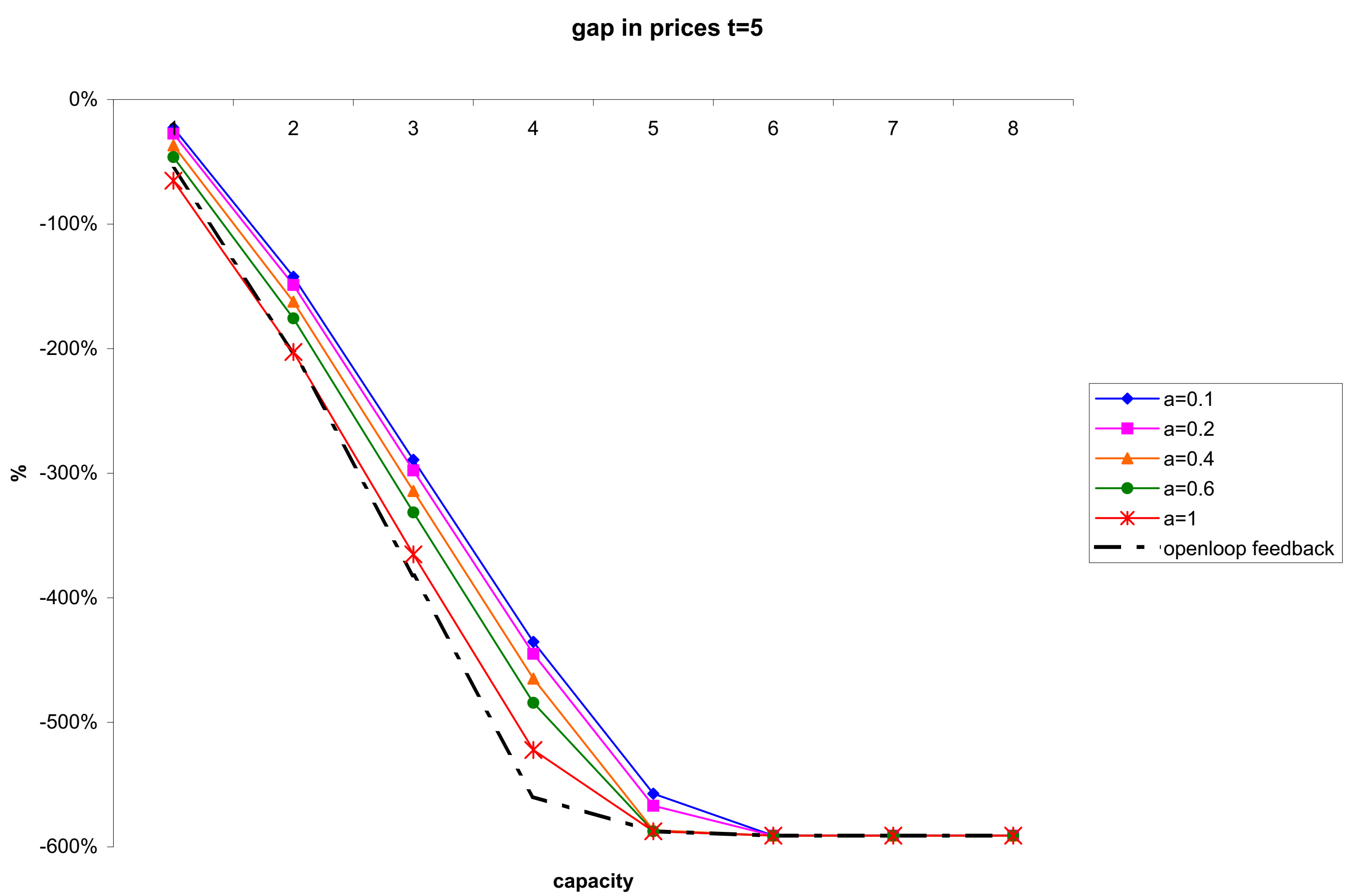
gap in intensity at t=5
%
95%
85%
75%
65%
55%
45%
35%
1
2
3
4
5
6
7
8
9
10
capacity
a=0.1
a=0.2
a=0.4
a=0.6
a=1
openloop feedback



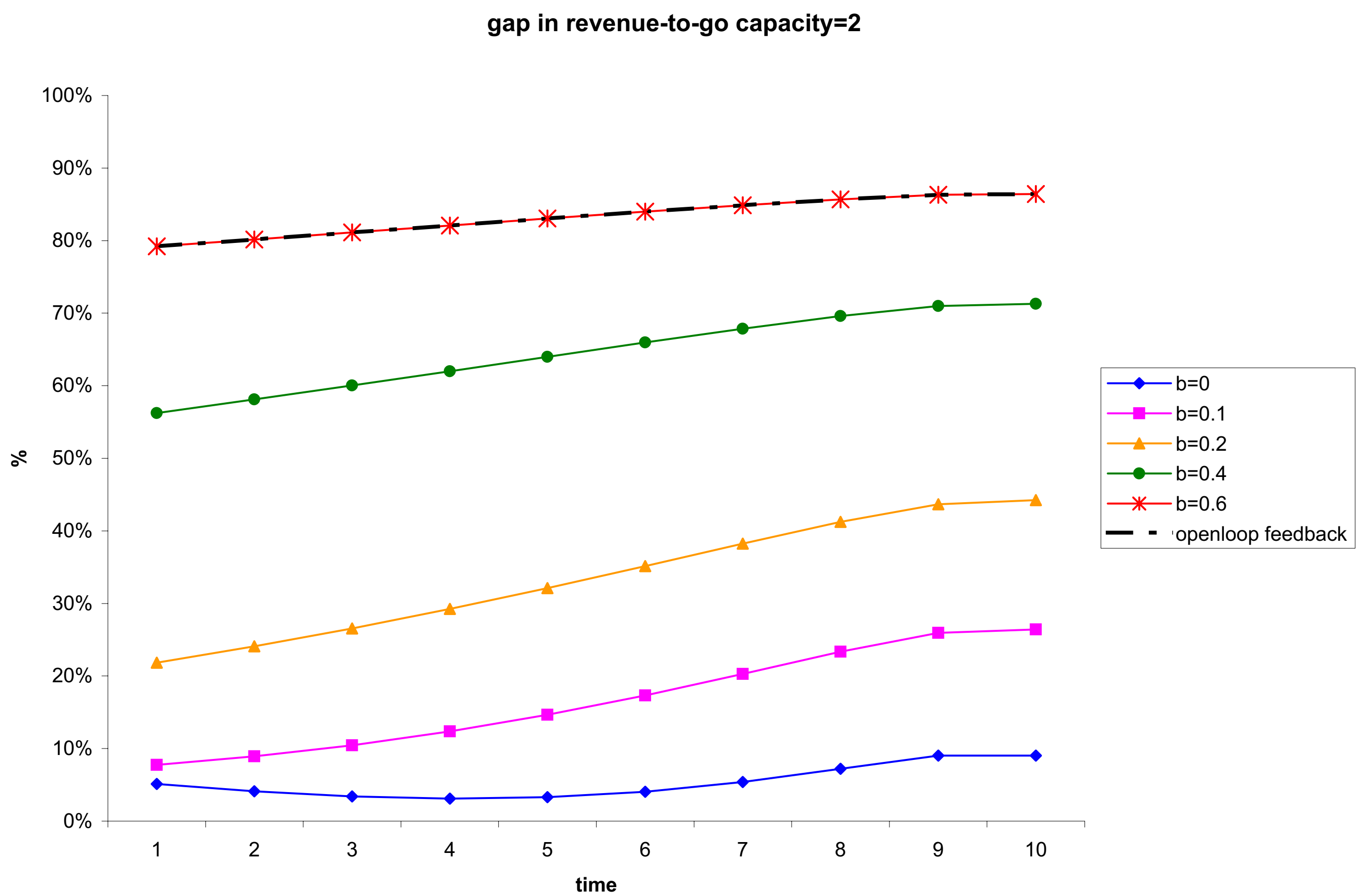
gap in prices capacity=2
0%
-50%
-100%
-150%
-200%
-250%
-300%
-350%
-400%
%
1
2
3
4
5
6
7
8
9
10
time
a=0.1
a=0.2
a=0.4
a=0.6
a=1
openloop feedback

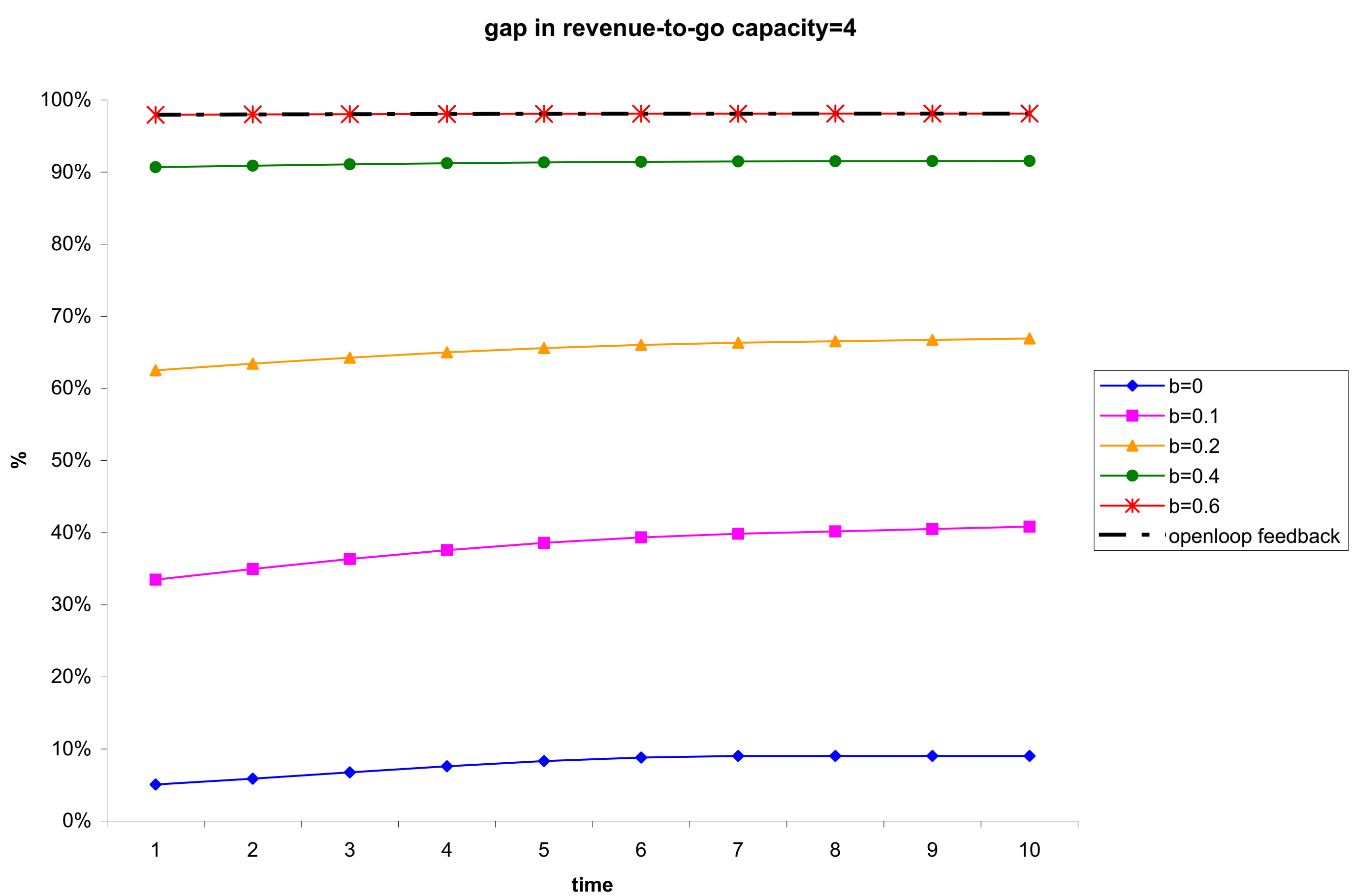

gap in prices capacity=4
0%
-100%
-200%
-300%
-400%
-500%
-600%
-700%
%
1
2
3
4
5
6
7
8
9
10
time
a=0.1
a=0.2
a=0.4
a=0.6
a=1
openloop feedback

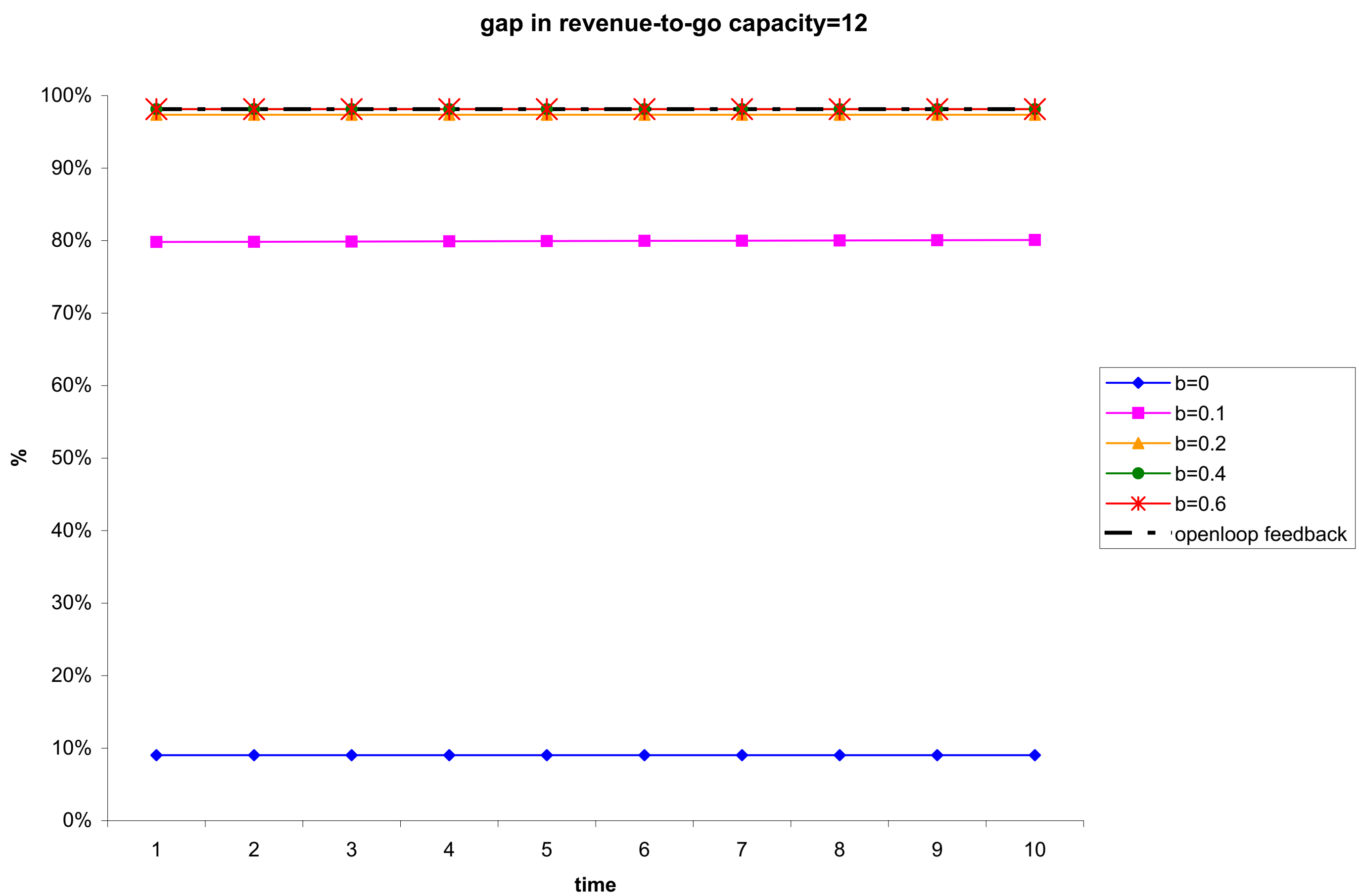

gap in prices capacity=12
-580%
-590%
-600%
%
1
2
3
4
5
6
7
8
9
10
time
a=0.1
a=0.2
a=0.4
a=0.6
a=1
openloop feedback

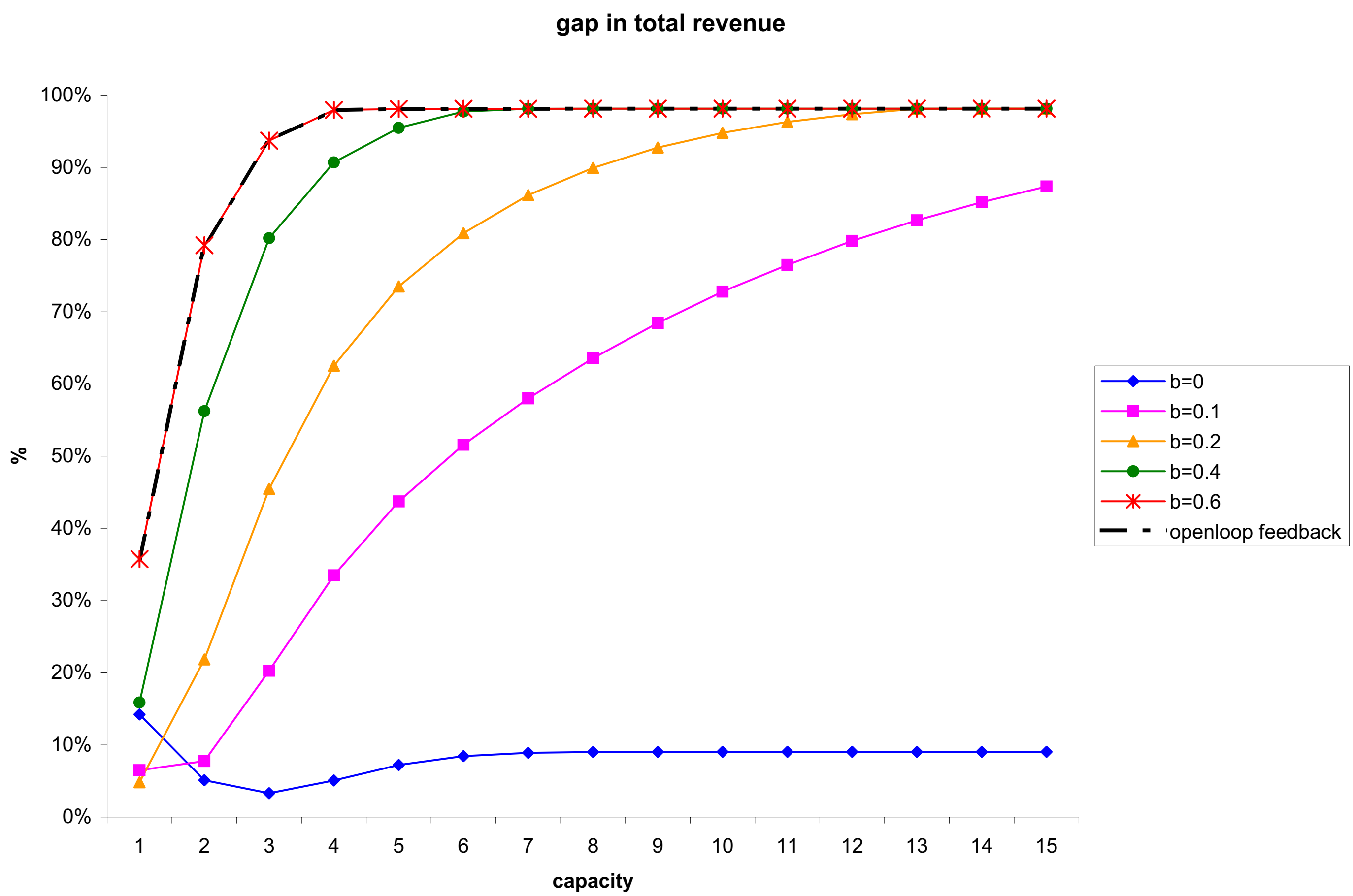


gap in prices t=5
0%
-100%
-200%
-300%
-400%
-500%
-600%
%
2
3
4
5
6
7
8
capacity
a=0.1
a=0.2
a=0.4
a=0.6
a=1
openloop feedback

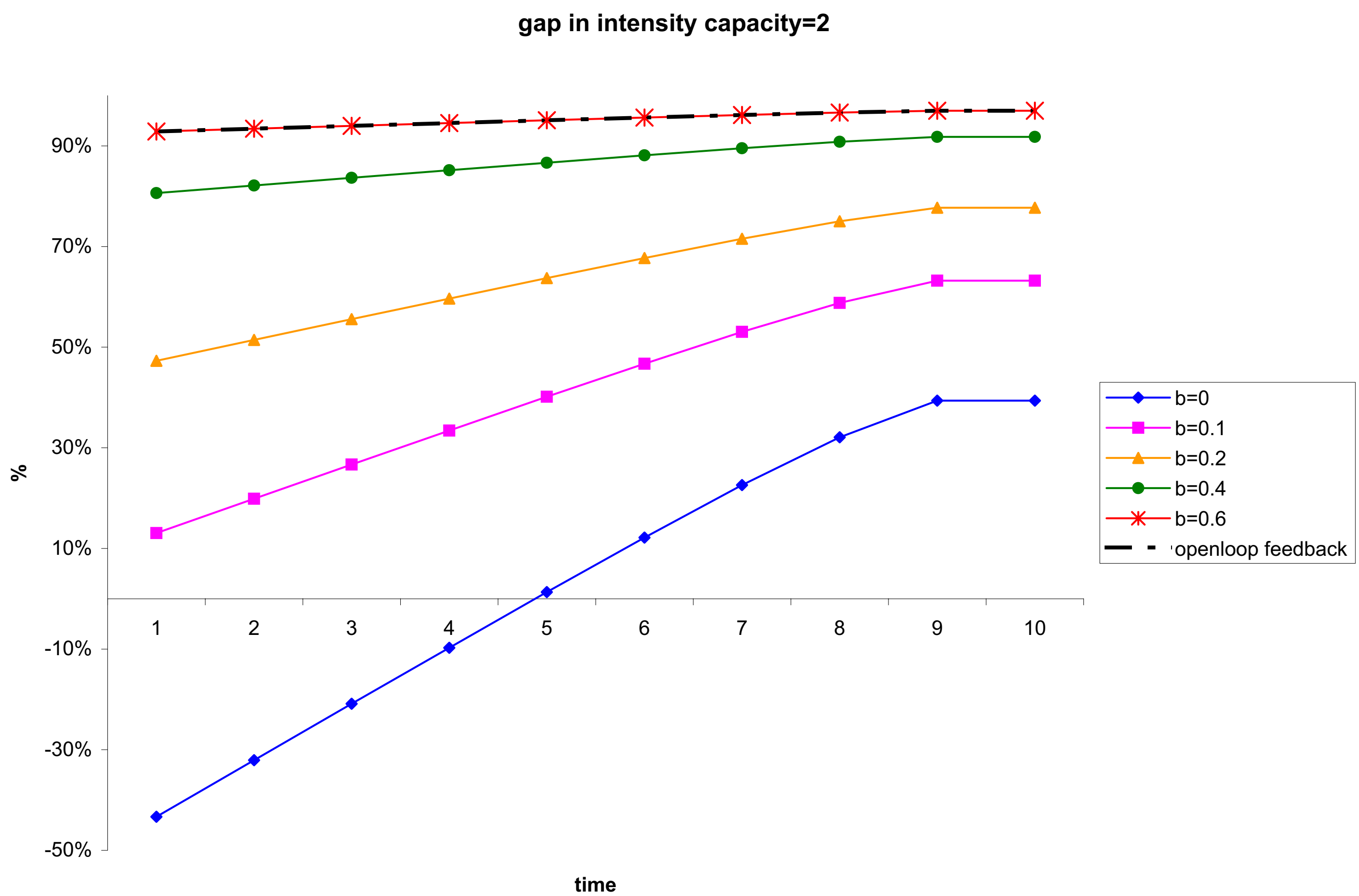

gap in revenue-to-go capacity=2
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
%
1
2
3
4
5
6
7
8
9
10
time
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback


gap in revenue-to-go capacity=4
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
%
1
2
3
4
5
6
7
8
9
10
time
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback

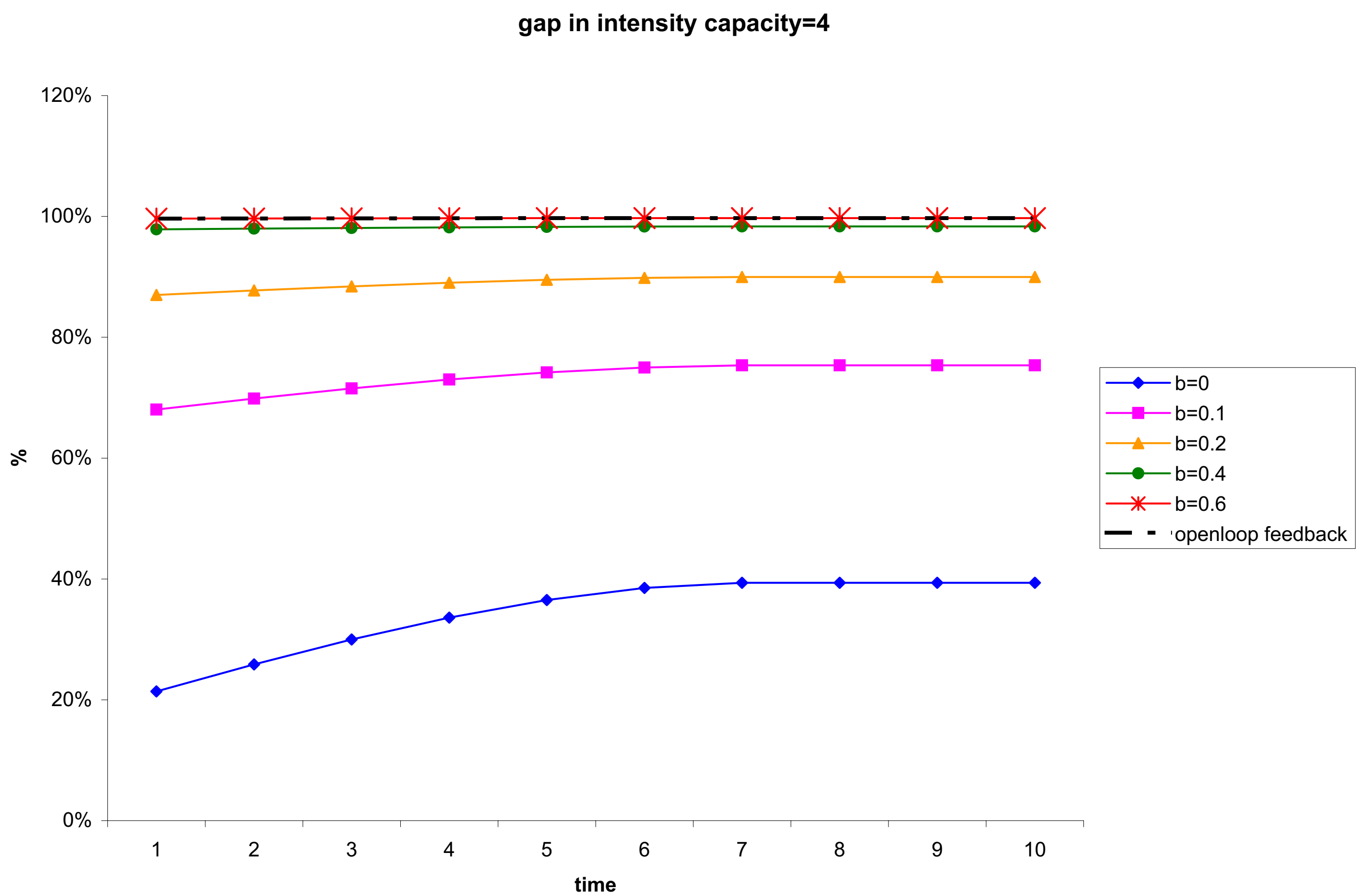

gap in revenue-to-go capacity=12
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
%
1
2
3
4
5
6
7
8
9
10
time
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback



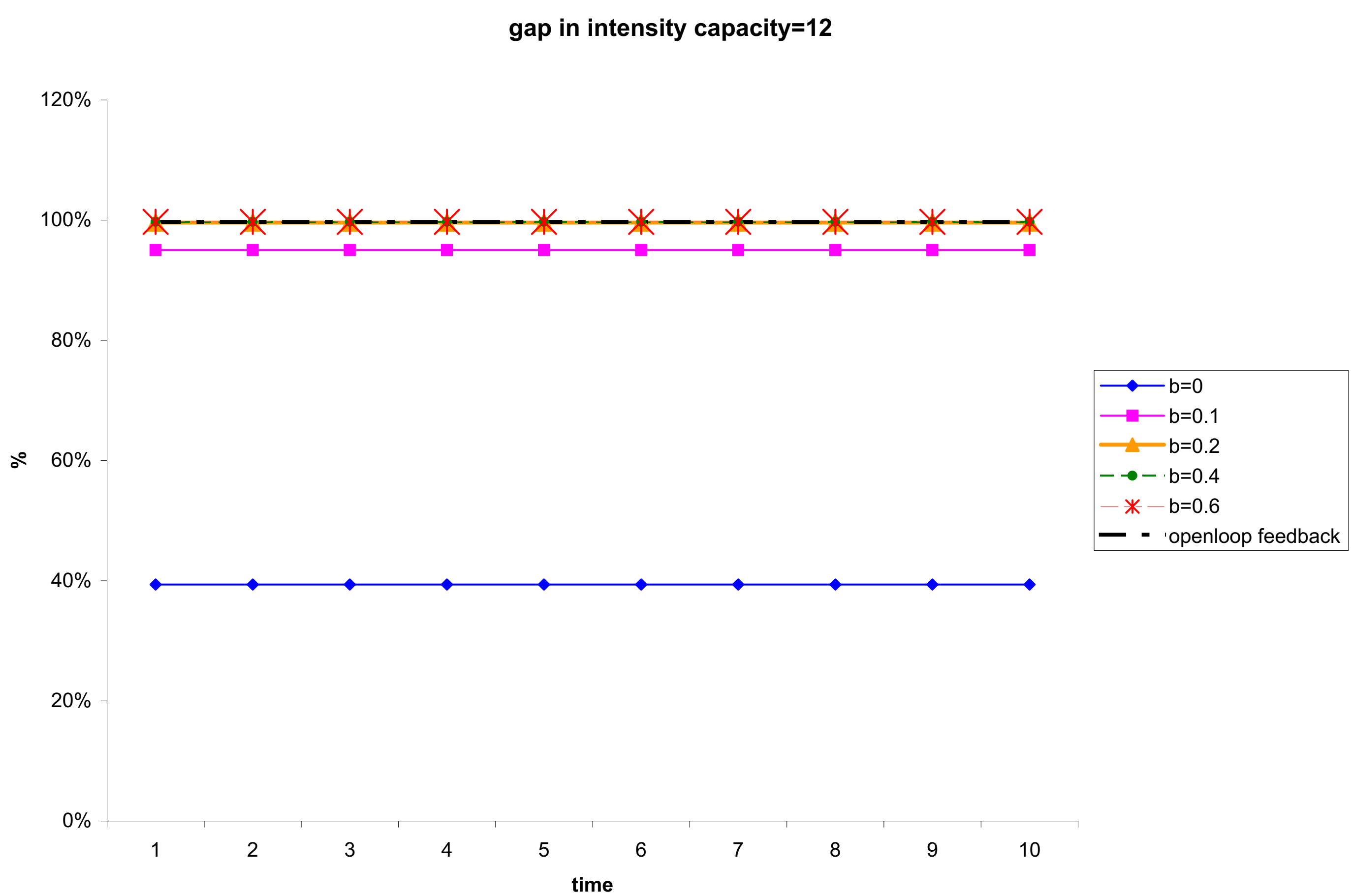
gap in total revenue
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
capacity
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback

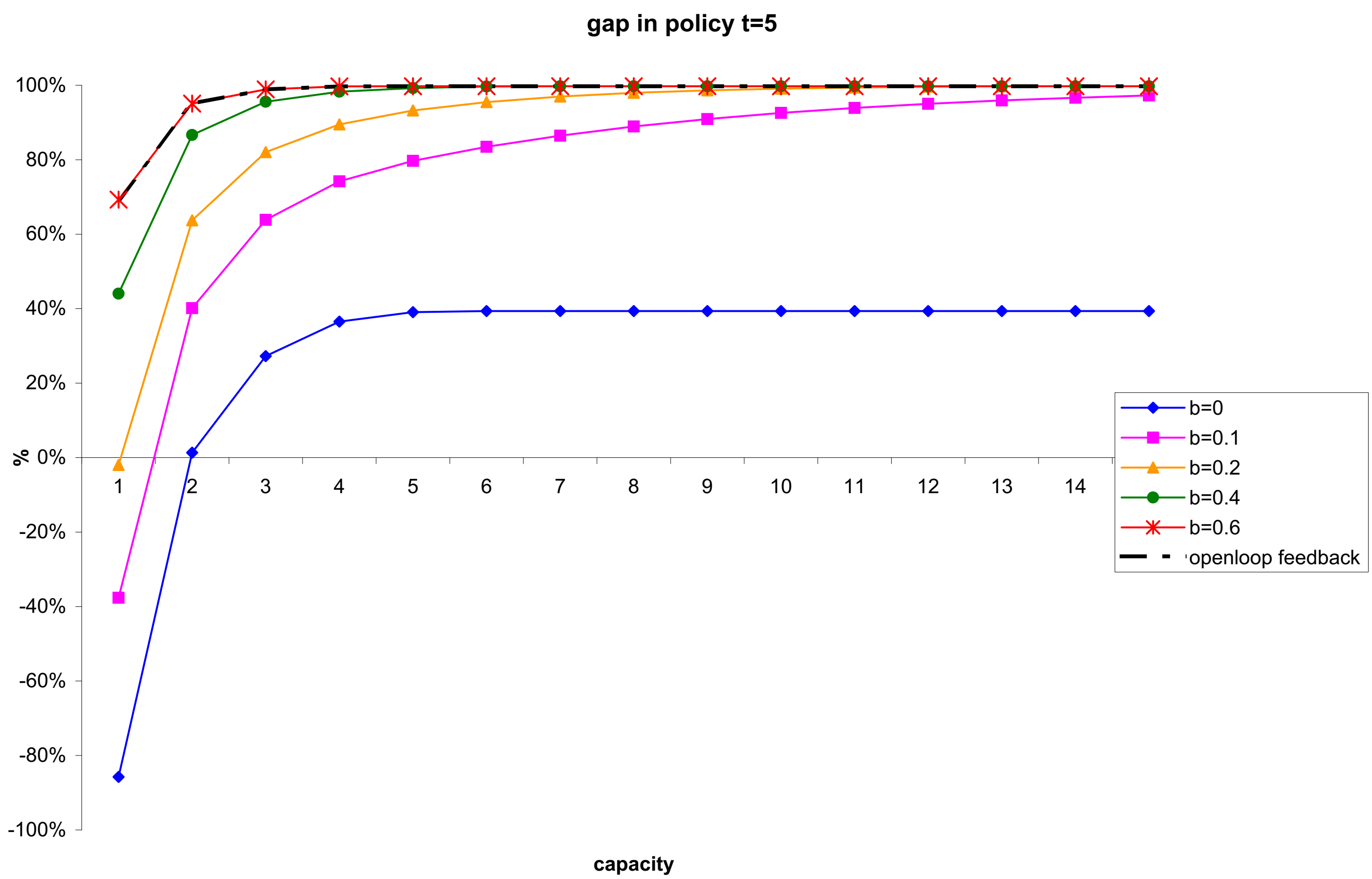


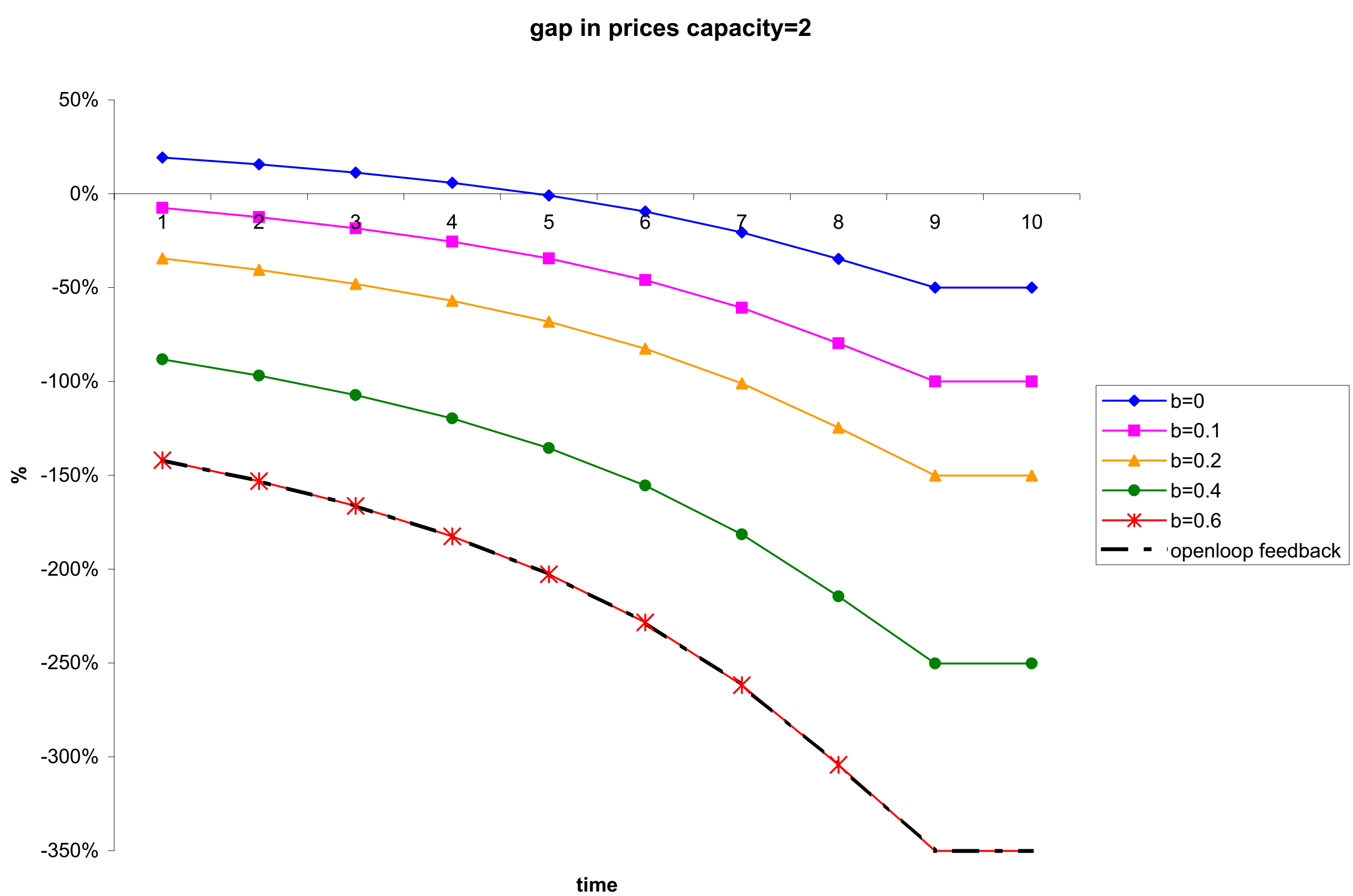
gap in intensity capacity=2
90%
70%
50%
30%
10%
-10%
-30%
-50%
%
1
2
3
4
5
6
7
8
9
10
time
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback



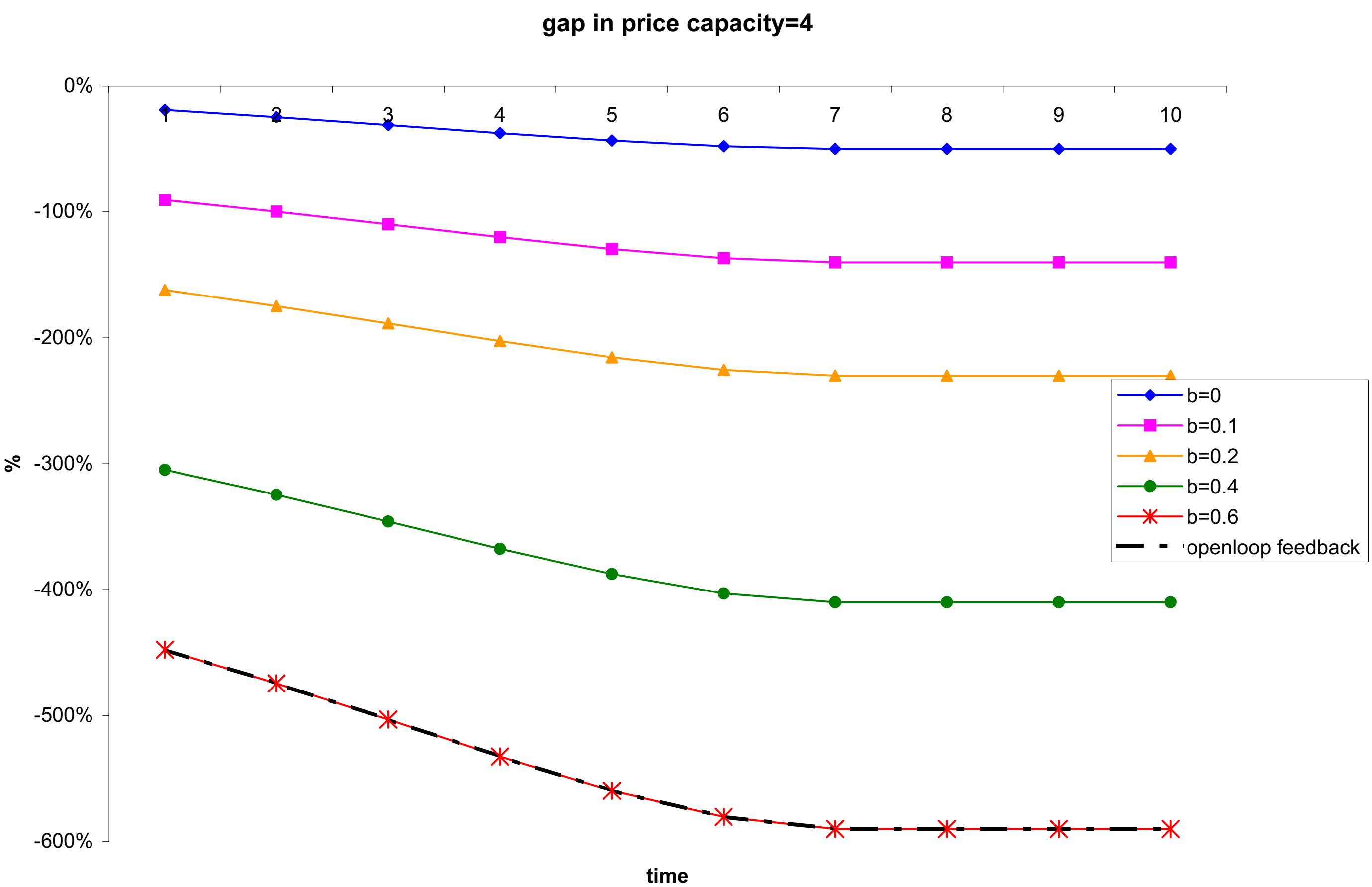
gap in intensity capacity=4
120%
100%
80%
60%
40%
20%
0%
%
1
2
3
4
5
6
7
8
9
10
time
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback

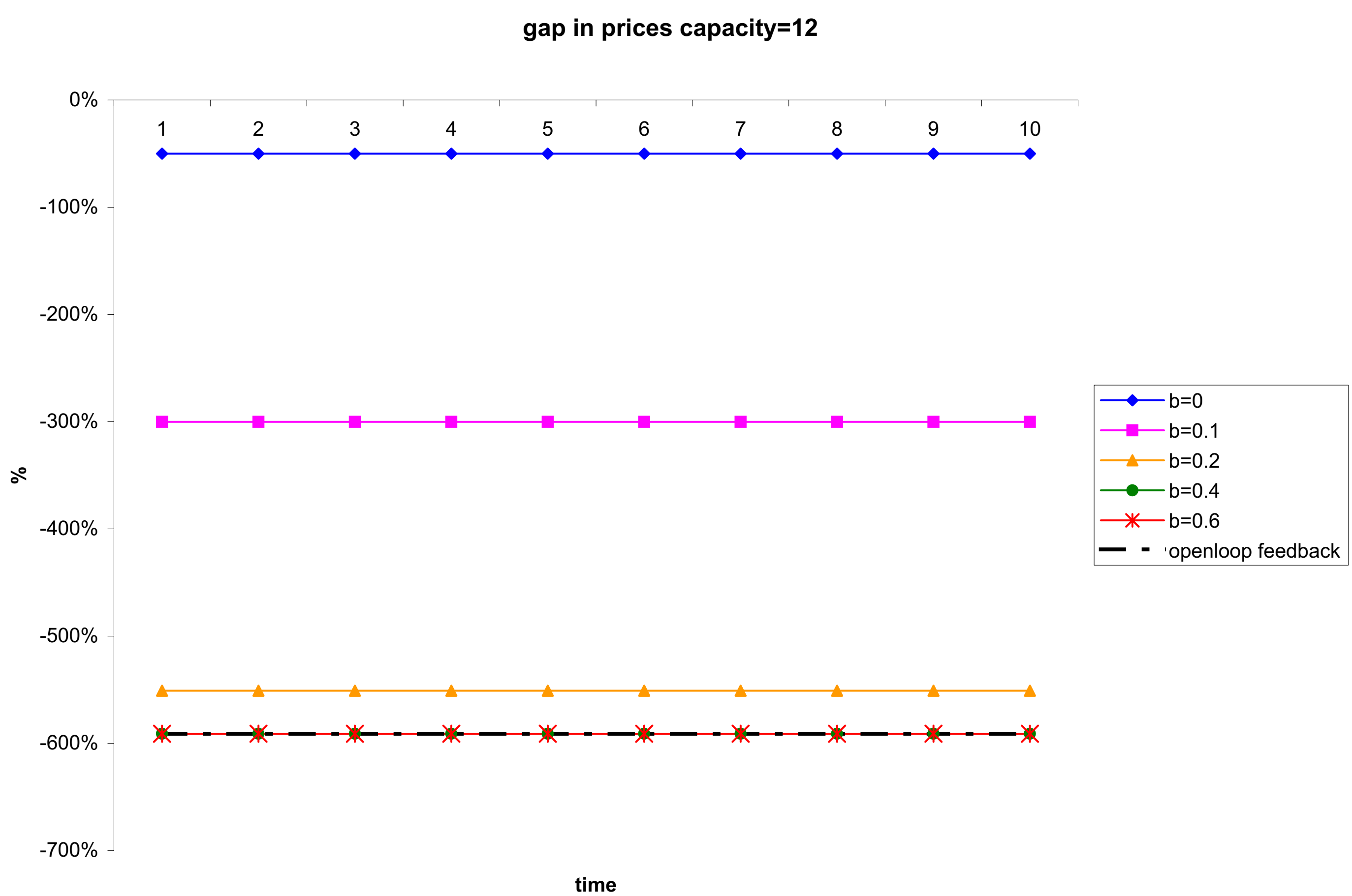

gap in intensity capacity=12
120%
100%
80%
60%
40%
20%
0%
%
1
2
3
4
5
6
7
8
9
10
time
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback

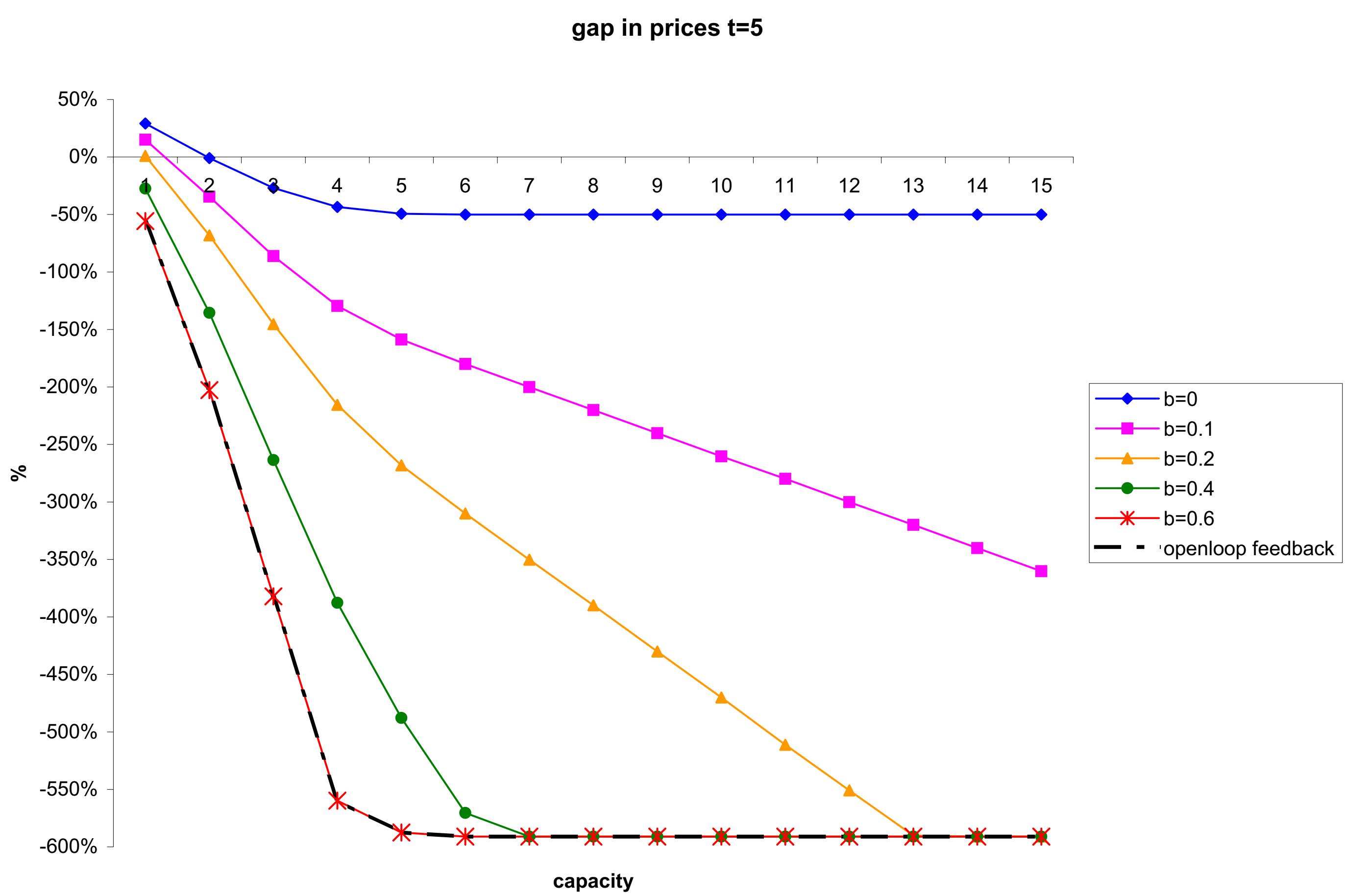

gap in policy t=5
100%
80%
60%
40%
20%
0%
-20%
-40%
-60%
-80%
-100%
%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
capacity
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback



gap in prices capacity=2
50%
0%
-50%
-100%
-150%
-200%
-250%
-300%
-350%
%
1
2
3
4
5
6
7
8
9
10
time
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback


gap in price capacity=4
0%
-100%
-200%
-300%
-400%
-500%
-600%
%
1
2
3
4
5
6
7
8
9
10
time
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback


gap in prices capacity=12
0%
-100%
-200%
-300%
-400%
-500%
-600%
-700%
1
2
3
4
5
6
7
8
9
10
%
time
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback



gap in prices t=5
50%
0%
-50%
-100%
-150%
-200%
-250%
-300%
-350%
-400%
-450%
-500%
-550%
-600%
%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
capacity
b=0
b=0.1
b=0.2
b=0.4
b=0.6
openloop feedback

Figure B-52: Gap as Function of the Iteration Number for Horizon of Length 6

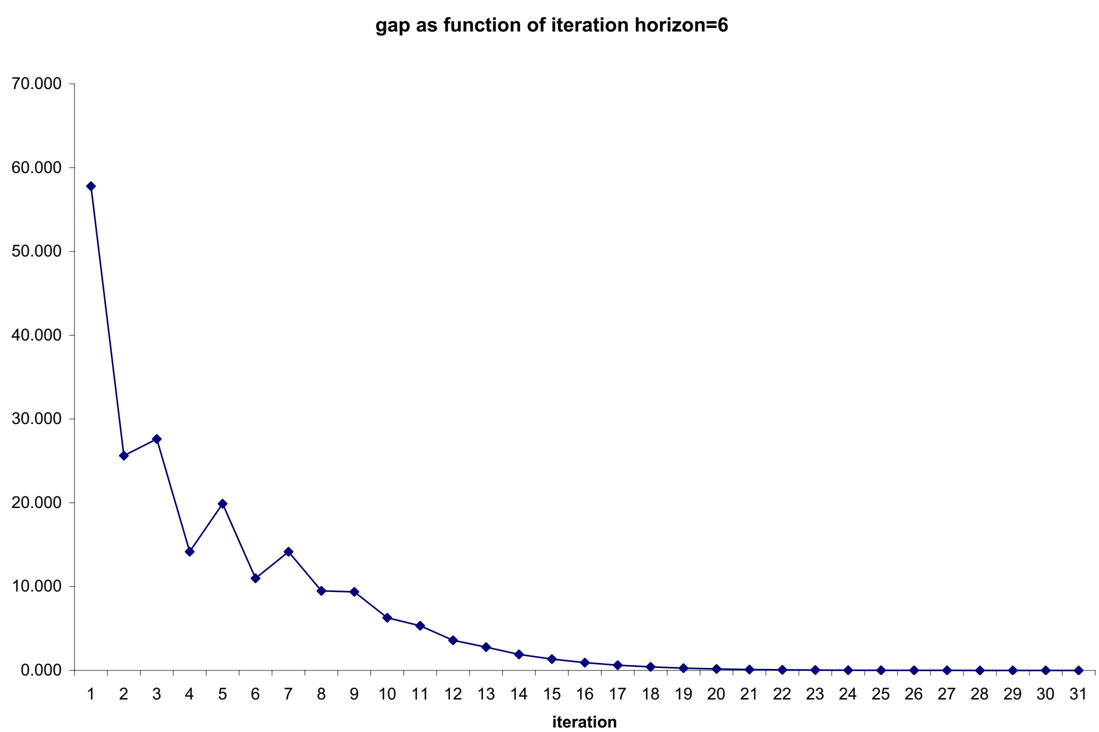


Figure B-53: Gap as Function of the Iteration Number for Horizon of Length 8

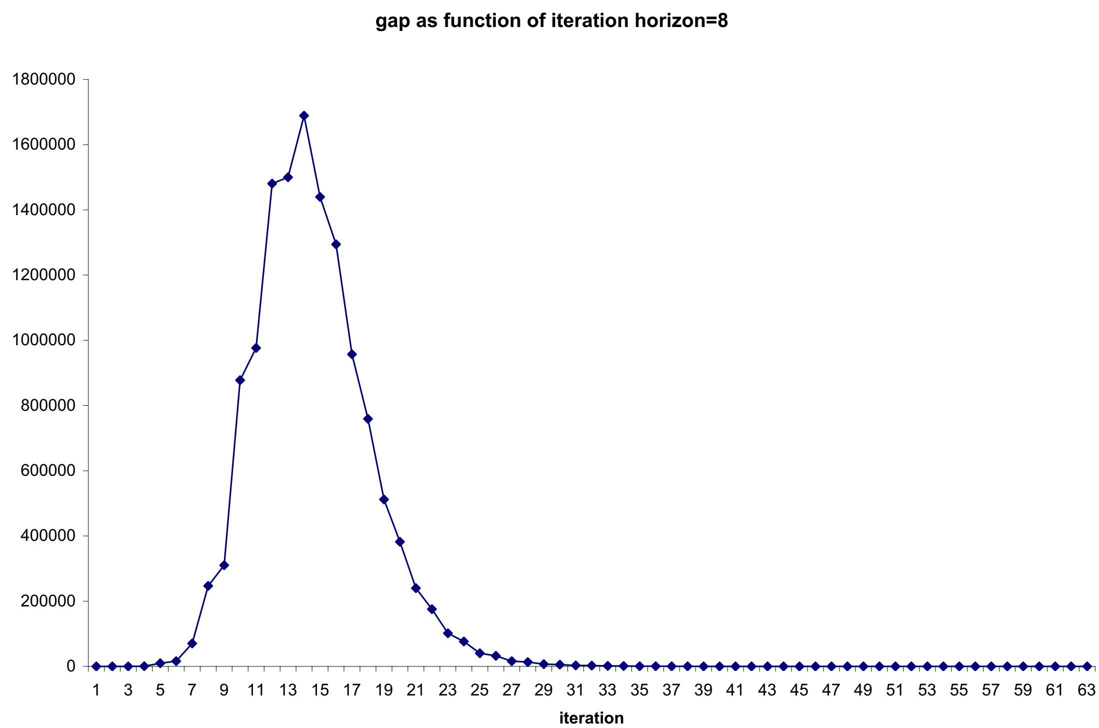


Figure B-54: Number of Iterations as Function of the Length of the Horizon

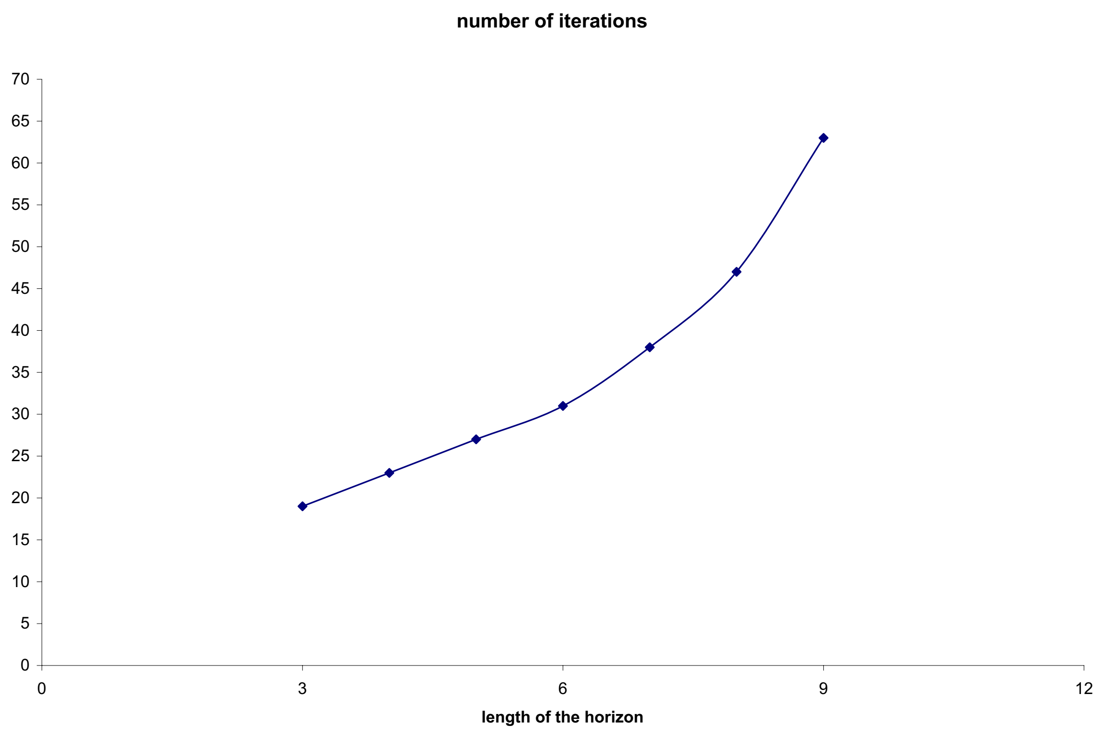
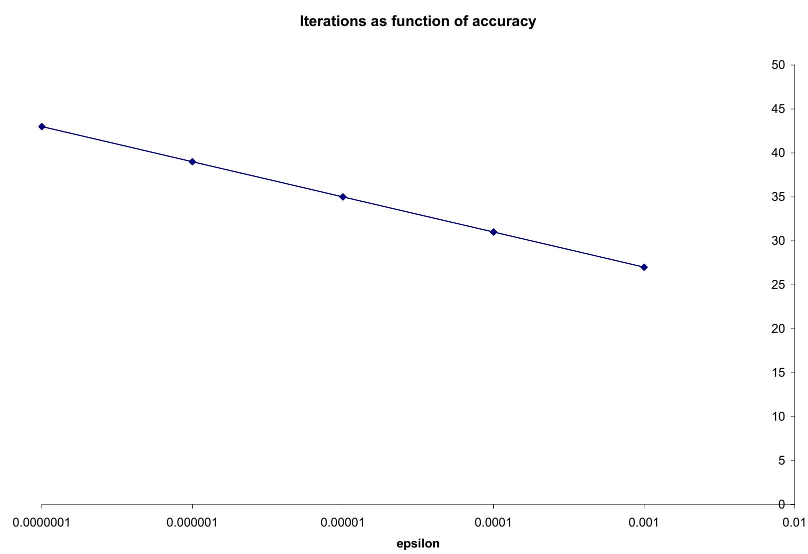


Figure B-55: Number of Iterations as Function of the Accuracy Level



Bibliography

[1] E. Adida and G. Perakis. A robust optimization approach to dynamic pricing and inventory control with no backorders. *Mathematical Programming*, 107(1-2):97–129, 2006. 31.

[2] P. Afèche and B. Ata. Revenue management in queueing systems with unknown demand characteristics. Submitted to: Manufacturing and Service Operations Management, 2005.

[3] V. Araman and R. Caldentey. Dynamic pricing for non-perishable products with demand learning. Under review in Operations Research, Submitted first in 2005.

[4] P. Armstrong. *Quality Control in Services*. PhD thesis, Department of Decision Sciences, Wharton School of Business, University og Pennsylvania, Philadelphia, 1993.

[5] Y. Aviv and A. Pazgal. A partially observed markov decision process for dynamic pricing. *Management Science*, 51(9):1400–1416, 2005.

[6] Y. Aviv and A. Pazgal. Pricing of short life-cycle products through active learning. 2005.

[7] C. Baiocchi and A. Capelo. *Variational and Quasi-Variational Inequalities: Applications to Free Boundary Problems*. John Wiley and Sons, 1984.

[8] R.J. Balvers and T.F Cosimano. Actively learning about demand and the dynamics of price adjustment. *Econom. Journal*, 100:882–898, 1990.

[9] A. Barto and R. Sutton. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, 1998.

[10] A. Bensoussan and J-L. Lions. *Applications of Variational Inequalities to Stochastic Control*. Number 12 in Studies in Mathematics and its Applications. Elsevier, 1982.

[11] F. Bernstein and A. Federgruen. Dynamic inventory and pricing models for competing retailers. *Naval Research Logistics*, 51(2):258–274, 2004. 32.

[12] D. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1995.

[13] O. Besbes and A. Zeevi. Dynamic pricing without knowing the demand function: Risk bounds and near optimal pricing algorithms. 2006.

[14] G. R. Bitran, R. Caldentey, and S. V. Mondschein. Coordinating clearance markdown sales of seasonal products in retail chains. *Operations Research*, 46(5), 1998.

[15] G. R. Bitran and S. V. Mondschein. Periodic pricing of seasonal products in retailing. *Management Science*, 43(1):64–79, 1997. 11.

[16] R. Bitran, G. R. amd Caldentey. An overview of pricing models for revenue management. *Manufacturing and Service Operations Management*, 5:27, 2003.

[17] S. E. Bodily and L. R. Weatherford. Perishable-asset revenue management - generic and multiple-price yield management with diversion. *Omega-International Journal of Management Science*, 23(2):173–185, 1995. 11.

[18] S. Boyd and M. S. Lobo. Pricing and learning under uncertain demand. November 2003.

[19] J. Cardell, C. C. Hitt, and W. W Hogan. Market power and strategic interaction in electricity networks. *Resource and Energy Economics*, 19:109–137, 1997.

[20] A. X. Carvalho and M. L. Puterman. Dynamic pricing and reinforcement learning. *Neural Networks*, 2003.

[21] A. X. Carvalho and M. L. Puterman. Learning and pricing in an internet environment with binomial demands. *Journal of Revenue and Pricing, 3, 4, 320-336, 2005*, 3(4):320–336, 2005.

[22] D. Chan and J-S. Pang. The generalized quasi-variational inequality problem. *Mathematics of Operations Research*, 7:211–222, 1982.

[23] S. C. Choi, W. S. Desarbo, and P.T. Harker. Product positionning under price competition. *Management Science*, 36:175–199, 1990.

[24] collective. *Mathematical and Computational Models for Congestion Charging*, volume 101 of *Applied Optimization*, chapter Dynamic Pricing: A Learning Approach, Bertsimas, D. J. and Perakis, G. Springer, Regionalwissenschaften, 2006., 2006.

[25] B. Colson, P. Marcotte, and G. Savard. Bilevel programming, a survey. *4OR: A Quarterly Journal*, 3(2):87–107, June 2005.

[26] P. Cubiotti. On the discontinuous infinite-dimensional generalized quasivariational inequality problem. *Journal of Optimization Theory and Applications*, 115:97–111, 2002.

[27] Y. Dai, X. L. Chao, S. C. Fang, and H. L. W. Nuttle. Pricing in revenue management for multiple firms competing for customers. *International Journal of Production Economics*, 98(1):1–16, 2005. 29.

[28] S. Dempe and D. Ralph. Directional derivatives of the solutions of a parametric nonlinear program. *Mathematical Programming*, 70:159–172, 1995.

[29] E. Dockner and S. Jørgensen. Optimal pricing strategies for new products in dynamic oligopolies. *Marketing Science*, 7(4):315–334, 1988.

[30] D. Easley and N. Kiefer. Controlling a stochastic process with unknown parameters. *Econometrica*, 56:1045–64, 1988.

[31] W. Elmaghraby, A. Guleu, and P. Keskinocak. Optimal markdown mechanisms in the presence of rational customers with multi-unit demand. 2002.

[32] W. Elmaghraby and P. Keskinocak. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science*, 49(10):1287–1309, 2003. 83.

[33] R. Erikson and A. Pakes. Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62(1):53–82, 1995.

[34] F. Facchinei and J.S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*, volume I and II. Springer Verlag, Berlin, 2003.

[35] A. Farahat and G. Perakis. A comparison of bertrand and cournot profits in oligopolies with gross substitutes. 2006.

[36] V. F. Farias and B. van Roy. Dynamic pricing with a prior on market response. *Management Science*, Forthcoming.

[37] A. Federgruen and A. Heching. Combined pricing and inventory control under uncertainty. *Operations Research*, 47(3):454–475, 1999.

[38] Y. Y. Feng and C. Gallego. Perishable asset revenue management with markovian time dependent demand intensities. *Management Science*, 46(7):941–956, 2000. 39.

[39] Y. Y. Feng and G. Gallego. Optimal starting times for end-of-season sales and optimal stopping-times for promotional fares. *Management Science*, 41(8):1371–1391, 1995. 26.

[40] Y. Y. Feng and B. C. Xiao. Optimal policies of yield management with multiple predetermined prices. *Operations Research*, 48(2):332–343, 2000. 29.

[41] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Technique*. John Wiley and Sons, 1968.

[42] J. Friedman. *Oligopoly Theory*. Cambridge University Press, 1983.

[43] D. Fudenberg and J. Tirole. *Dynamic models of oligopoly*. Harwood Academic Publishers, Chur; New York, 1986.

[44] M Fukushima and J-S. Pang. Quasi-variational inequalities, generalized nash equilibria, and multi-leader-follower games. *Computational Management Science*, 1:21–56, 2005.

[45] G. Gallego and T. Hu. Finite horizon dynamic pricing of perishable assets under competiton. AGIFORS Anna Valicek Award, 2007.

[46] G. Gallego and G. VanRyzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40(8):999–1020, 1994. 55.

[47] G. Gallego and G. VanRyzin. A multiproduct dynamic pricing problem and its applications to network yield management. *Operations Research*, 45(1):24–41, 1997. 36.

[48] P. T. Harker. Generalized nash games and quasi-variational inequalities. *European Journal of Operations Research*, 54:81–94, 1991.

[49] P. T. Harker and J-S. Pang. Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. *Mathematical Programming*, 48(1-3):161–220, March 1990.

[50] B. Jun and X. Vives. Strategic incentives in dynamic duopoly. *Journal of Economic Theory*, 116(2):249–281, 2004. 48.

[51] S. Kachani and G. Perakis. Fluid dynamics models and their applications in transportation and pricing. *European Journal of Operational Research*, 170(2):496–517, 2006. 56.

[52] K. Kalyanam. Pricing decision under uncertainty: A bayesian mixure model approach. *Marketing Science*, 15(3):207–221, 1996.

[53] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and their Applications*. Academic Press, New York, 1980.

[54] M. Kojima. Strongly stable stationary solutions in nonlinear programs. Technical report, Wisconsin University at Madison, Mathematics Research Center, February 1979.

[55] J. B. Krawczyk and M. Tidball. A discrete-time dynamic game of seasonal water allocation. *Journal of Optimization Theory and Applications*, 128(2):411–429, 2006. Krawczyk, J. B. Tidball, M. 13.

[56] C.E. Larson, L.J. Olson, and S. Sharma. Optimal inventory policies when the demand distribution is not known. *Journal of Economic Theory*, 101:281–300, 2001.

[57] E.P. Lazear. Retail pricing and clearance sales. *The American Economic Review*, 76(1):14–32, 1986.

[58] Y. Levin, J. McGill, and M. Nediak. Dynamic pricing in the presence of strategic consumers and oligopolistic competition. Working Paper, School of Business, Queen's University, Kingston, ON, 2006.

[59] Z. Luo, J-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, 1996.

[60] C. Maglaras and J. Meissner. Dynamic pricing strategies for multiproduct revenue management problems. *MSom-Manufacturing and Service Operations Management*, 8(2):136–148, 2006. Maglaras, Constantinos Meissner, Joern 23.

[61] E. Maskin and J. Tirole. A theory of dynamic oligopoly .3. cournot competition. *European Economic Review*, 31(4):947–968, 1987. 8.

[62] E. Maskin and J. Tirole. A theory of dynamic oligopoly .1. overview and quantity competition with large fixed costs. *Econometrica*, 56(3):549–569, 1988. 15.

[63] E. Maskin and J. Tirole. A theory of dynamic oligopoly .2. price-competition, kinked demand curves, and edgeworth cycles. *Econometrica*, 56(3):571–599, 1988. 26.

[64] E. Maskin and J. Tirole. Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory*, 100:191–219, 2001.

[65] J. I. McGill and G. J. Van Ryzin. Revenue management: Research overview and prospects. *Transportation Science*, 33(2):233–256, 1999. 199.

[66] L. Mirman, L. Samuelson, and A. Urbano. Monopoly experimentation. *International Economic Review*, 34(3):549–63, August 1995.

[67] F.H Murphy, H.D. Sherali, and A.L Soyster. A mathematical programming approach for determining oligopolistic market equilibrium. *Mathematical Programming*, 24:92–106, 1982.

[68] A. Nagurney. *Network Economics: A Variational Inequality Approach*. Kluwer Academic Publishers, Boston, 2000.

[69] T. D. Nguyen and G. Perakis. Robust competitive multi-period pricing for multiple products with fixed capacity. June 2005.

[70] J.V Outrata. Solution behaviour for parameter-dependent quasi variational inequalities. *RAIRO - Recherche Operationnelle*, 30(4):399–415, 1996.

[71] J.S. Pang. Solution differentiability and continuation of newton's method for variational inequality problems over polyhedral sets. *Journal of Optimization Theory and Applications*, 1, July 1990.

[72] G. Perakis and A. Sood. Competitive multi-period pricing for perishable products: A robust optimization approach. *Mathematical Programming*, 107(1-2):295–335, 2006. 34.

[73] N. C. Petruzzi and M. Dada. Dynamic pricing and inventory control with learning. *Naval Research Logistics*, 49(3):303–325, 2002. 37.

[74] Y. Popescu and Y. Wu. Dynamic pricing strategies with reference effects. *Operations Research*, 55(3):413–429, May-June 2007.

[75] A. Rajan, Rakesh, and R. Steinberg. Dynamic pricing and ordering decisions by a monopolist. *Management Science*, 38(2):240–262, 1992. 39.

[76] J. Robinson. Generalized equations and their applications: Part i: Basic theory. *Mathematical Programming Study*, 10:128–141, 1979.

[77] J. Robinson. Generalized equations and their applications: Part ii: Applications to nonlinear programming. *Mathematical Programming Study*, 19:200–221, 1982.

[78] J.B. Rosen. Existence and uniqueness of equilibrium points for convave n-person games. *Econometrica*, 33(3):520–534, June 1965.

[79] A. Rustichini and A. Wolinsky. Learning about variable demand in the long run. *Journal of Economic Dynamics and Control*, 19:1283–92, 1995.

[80] S. A. Smith and D. D. Achabal. Clearance pricing and inventory policies for retail chains. *Management Science*, 44(3), 1998.

[81] A. Sood. *Competitive Multi-period Pricing for Perishable Products*. PhD thesis, MIT Operations Research Center, June 2004.

[82] G.J. Stigler. The kinky oligopoly demand curve and rigid prices. *The Journal of Political Economy*, 55(5):432–449, 1947.

[83] P.M. Sweezy. Demand under conditions of oligopoly. *The Journal of Political Economy*, 47(4):568–573, 1939.

[84] K. Talluri and G. van Ryzin. *The Theory and Practice of Revenue Management*. Kluwers Academic Publishers, 2004.

[85] G. Thowsen. A dynamic nonstationary inventory problem for a price/quantity setting firm. *Naval Research Logistics Quaterly*, 47:461–476, 1975.

[86] G. Tian and J. Zhou. Quasi variational inequalities with non compact sets. *Journal of Mathematical Analysis and Applications*, 169:583–595, 1991.

[87] X. Vives. *Oligopoly Pricing - Old Ideas and New Tools*. MIT Press, Cambridge, 1999.

[88] E. Zabel. Monopoly and uncertainty. *The Review of Economic Studies*, 37(2):205–219, 1970.