

# Pricing for Retail, Social Networks and Green Technologies

by

Maxime C. Cohen

B.S., Technion - Israel Institute of Technology (2006)

M.S., Technion - Israel Institute of Technology (2009)

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 2015

© Massachusetts Institute of Technology 2015. All rights reserved.

Author .....  
Sloan School of Management  
August 7, 2015

Certified by .....  
Georgia Perakis  
William F. Pounds Professor of Management  
Sloan School of Management  
Thesis Supervisor

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



# Pricing for Retail, Social Networks and Green Technologies

by

Maxime C. Cohen

Submitted to the Sloan School of Management  
on August 7, 2015, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy in Operations Research

## Abstract

What is the right price to charge your customers? Many retailers and researchers are facing this question. In the last three decades, tremendous progress was made, both in the academic and business worlds. In this thesis, we investigate four novel pricing applications.

In the first part, we study the promotion optimization problem for supermarket retailers. One needs to decide which products to promote, the depth of price discounts and when to schedule the promotions. To capture the stockpiling behavior of consumers, we propose two general classes of demand functions that can be easily estimated from data. We then develop an approximation that allows us to solve the problem efficiently and derive analytical results on its accuracy.

The second part is motivated by the ubiquity of social networking platforms. We consider a setting where a monopolist sells an indivisible good to consumers embedded in a social network. First, the firm designs prices to maximize its profits. Subsequently, consumers choose whether to purchase the item or not. Assuming positive externalities, we show the existence of a pure Nash equilibrium. Using duality theory and integer programming techniques, we reformulate the problem into a linear mixed-integer program. We then derive efficient ways of optimally solving the problem for discriminative and uniform pricing strategies.

The third part considers the problem of pricing a product for which demand information is very limited. We impose minimal assumptions on the problem: that is, only the constant marginal cost and the maximal price the firm can set are known. We propose a simple way of pricing the product by approximating the true inverse demand by a linear function. Surprisingly, we show that this approximation yields a good profit performance for a wide range of demand curves.

In the final part, we consider green technology products such as electric vehicles. We propose a Stackelberg model where the government offers consumer subsidies in order to encourage the technology adoption, whereas the supplier decides price and production to maximize profits. We address the question: How does demand uncertainty affect the government, the industry and the consumers, when designing

policies?

Thesis Supervisor: Georgia Perakis  
Title: William F. Pounds Professor of Management  
Sloan School of Management

## Acknowledgments

To begin, I would like to thank my advisor Georgia Perakis for her outstanding mentorship throughout my PhD. Georgia's dedication to her students is unique and limitless. Actually, Georgia is far from being only a research advisor to me. Georgia constantly motivates her students to aspire for excellence and to get the most out of MIT. From my early days as a graduate student, Georgia provided me with opportunities to attend conferences, intern in top companies and encourage me to TA several MBA classes. I had (and hope to continue having) the privilege to learn from Georgia skills including creative thinking, state of the art research methodologies and time management. I could feel that Georgia was caring at least as much about my personal growth and happiness as she was about my research progress. Georgia was always here to support and assist me by spending countless hours discussing our research. Inspired by Georgia, I hope one day to become such a great mentor to others.

I would also like to thank the members of my thesis committee Aharon Ben-Tal and Retsef Levi for their feedback and valuable comments on my work. On the academic side still, I learnt from a number of wonderful professors at MIT. I would especially like to thank Itai Ashlagi, Dimitris Bertsimas, Vivek Farias, Patrick Jaillet, Retsef Levi, Asu Ozdaglar, Robert Pindyck, Zeynep Ton, John Tsitsiklis, Juan Pablo Vielma, Tauhid Zaman and Karen Zheng for all they taught me. I will always be indebted to Nahum Shimkin, Allen Tannenbaum and Aharon Ben-Tal from the Technion for introducing me to Operations Research and directing me to MIT.

I would like to express a special thank you to the brilliant research collaborators I worked with: Ruben Lobel, Charles Thraves, Pavithra Harsha, Zachary Leung, Kiran Panchangam, Andrew Vakhutinsky, Swati Gupta, Ludovica Rizzo, Jonathan Chemama, Jeremy Kalas, Xubo Sun, Marjan Baghaie, Anthony Smith, Danny Segev and Lennart Baardman.

The Operations Research Center has been my home at MIT for the last five years. A great thanks to Dimitris Bertsimas and Patrick Jaillet for their dedication at the

head of the Center. My sincere appreciation also goes to Laura Rose and Andrew Carvalho for making sure the center runs smoothly and the students have everything they need. In addition, I would like to thank Ariel Brandner for the logistical support at Sloan.

Life in the United States would not have been the same without the presence of my friends. I interacted with many wonderful colleagues and friends at the Operations Research Center. In particular, I enjoyed working and spending time with Yaron Shaposhnik, Leon Valdes, Kris Johnson, Nathan Kallus, Will Ma, Angie King, Fernanda Bravo, Chaitanya Bandi, Vishal Gupta, Matthieu Monsch, Andre Calmon, Gonzalo Romero, Florin Ciocan, Phil Keller, Zhe Zhu, Ben Letham, Velibor Misic, Nikita Korolko, Jason Acimovic, Nikos Trichakis, Gareth Williams, Alex Rikun, Virgile Galle, Ali Aouad and Iain Dunning. A special thanks to my Sloan officemates Adam Elmachtoub and Paul Grigas for spending long days and nights discussing research and motivating each other. I would also like to thank my roommates Gal Shulkind, Arthur Flajolet, Carl Boand and Tarun Jain who had the courage to live with me over the years.

Spending five years in Boston allowed me to meet great people that became very close friends. A special thanks to Audren Cloitre, Lauren Ammar, Jonathan Klumberg, Gal Shulkind, David Puyraimond Zemmour, Tzachi Hazan, Alexandre Bouaziz, Michael Chemama, Alexis Turjman, Adriana Schultz, Luis Voloch, Guillaume Saint-Jacques, Leonid Rapoport, Yaniv Altshuler, Rami Abi Akl, Alexandre Jacquillat, Geraldine Gudefin, Trevor Frankel, Elisha Rejovitzky, Carine Margez, Daniel Rock, Rachid Naciri, Nathaniel Cohen and Oren Berkowitz for always being there for me to have fun and support me.

I was fortunate to witness that geographical distance is definitely not an obstacle to strong friendship. I would like to specially thank four of my undergraduate classmates from the Technion: Nathan Bettach, Alex Lellouche, Michael Fiszer and Gaddiel Ouaknin. I would also like to thank Olivier Rubel, Alexandre Benchimol, Jonathan Sidi, Miryam Dreyfuss, Daniel Guetta, Rabbi Eliahou Zini, Enath Fhal,

David Adjashvili, Ken Bugayong, Michael Benborhoum and Elie Toubiana.

Another key aspect of my life in Cambridge has been my involvement in the local jewish community. I cannot thank enough the Harvard Chabad, the BU Chabad, Rabbi Posner, Rabbi Ganz and his wife and the Chai Center in Brookline for the community they help build, for the fabulous meals I have shared with them and for their insightful teachings.

Last but not least, I am extremely grateful to my family for encouraging me to come to MIT and to always pursue my dreams. I could not imagine getting through this journey without their love and support. In particular, I am very grateful to my mother for her unconditional support, love and dedication. I would also like to express my gratitude to my father that always believed in me. Everything I am today and all I have achieved, I owe it to my family. To them I dedicate this thesis.



# Contents

<b>1</b>	<b>Introduction</b>	<b>17</b>
1.1	Summary . . . . .	17
1.2	Technical Contributions . . . . .	19
1.3	Broader Impact . . . . .	23
<b>2</b>	<b>Promotion Optimization for Supermarket Retailers</b>	<b>25</b>
2.1	Introduction . . . . .	25
2.2	Literature review . . . . .	31
2.3	Model and Assumptions . . . . .	33
2.3.1	Business Rules . . . . .	35
2.3.2	Assumptions . . . . .	37
2.4	Problem Formulation . . . . .	38
2.5	IP Approximation . . . . .	40
2.6	Demand Models . . . . .	43
2.6.1	Multiplicative Demand . . . . .	45
2.6.2	Additive Demand . . . . .	49
2.7	Multiple Items . . . . .	49
2.7.1	Problem Formulation . . . . .	50
2.7.2	Solution Approach . . . . .	53
2.7.3	Insights . . . . .	61
2.8	Computational Results . . . . .	64
2.8.1	Estimation Method . . . . .	65

2.8.2	Estimation Results and Discussion . . . . .	68
2.8.3	Optimization Results and Discussion . . . . .	70
2.9	Conclusions . . . . .	74
<b>3</b>	<b>Designing Price Incentives in a Network with Social Interactions</b>	<b>79</b>
3.1	Introduction . . . . .	79
3.2	Model . . . . .	87
3.3	Purchasing equilibria . . . . .	93
3.4	Optimal pricing: MIP formulation . . . . .	97
3.5	Efficient algorithms . . . . .	101
3.5.1	Discriminative Prices . . . . .	101
3.5.2	Uniform Price . . . . .	104
3.6	Extension to uncertain valuations . . . . .	106
3.7	Price-incentives to guarantee influence . . . . .	110
3.8	Computational experiments . . . . .	118
3.9	Conclusions . . . . .	124
<b>4</b>	<b>Pricing without knowledge of demand</b>	<b>129</b>
4.1	Introduction . . . . .	129
4.2	Common Demand Functions . . . . .	134
4.2.1	Quadratic Demand . . . . .	135
4.2.2	Monomial Demand . . . . .	137
4.2.3	Semi-Log Demand . . . . .	138
4.2.4	Log-Log Demand . . . . .	139
4.3	Uncertain Maximum Price . . . . .	141
4.4	General Demand Functions . . . . .	143
4.4.1	Concave Demand Curves . . . . .	144
4.4.2	General Random Demand Curves . . . . .	145
4.4.3	Uncertain Maximum Price . . . . .	147
4.5	Conclusions . . . . .	148

<b>5</b>	<b>The Impact of Demand Uncertainty on Consumer Subsidies for Green Technology Adoption</b>	<b>153</b>
5.1	Introduction . . . . .	153
5.2	Literature Review . . . . .	158
5.3	Model . . . . .	163
5.4	The Model . . . . .	169
5.4.1	Additive Noise . . . . .	170
5.4.2	Multiplicative Noise . . . . .	179
5.4.3	Consumer Surplus . . . . .	181
5.5	Supply-chain Coordination . . . . .	186
5.6	Subsidizing the manufacturer's cost . . . . .	188
5.7	Computational Results . . . . .	191
5.8	Conclusions . . . . .	198
<b>6</b>	<b>Conclusion</b>	<b>207</b>
<b>A</b>	<b>Proofs for Chapter 2</b>	<b>209</b>
<b>B</b>	<b>Proofs for Chapter 3</b>	<b>233</b>
<b>C</b>	<b>Proofs for Chapter 4</b>	<b>247</b>
<b>D</b>	<b>Proofs for Chapter 5</b>	<b>253</b>



# List of Figures

2-1	Prices and sales for a particular brand of coffee during 35 weeks . . .	26
2-2	Example of the promotion fatigue effect . . . . .	36
2-3	Profit function for demand with promotion fatigue effect . . . . .	39
2-4	Illustration of the loss leader effect . . . . .	62
2-5	Summary of POP insights regarding cross-item effects . . . . .	63
2-6	Actual versus forecasted sales over the 35 test weeks for Brand1 . . .	69
2-7	Profits for different scenarios using a log-log demand model . . . . .	72
2-8	Running times of the LP formulation and the exhaustive search method for the POP . . . . .	74
3-1	Value of incorporating network externalities for the discriminative and uniform pricing strategies . . . . .	118
3-2	Centrality effect: losing money on an influential agent . . . . .	119
3-3	Value of incorporating incentives that guarantee influence . . . . .	121
3-4	Symmetric Graph with asymmetric incentives: with and without in- centives . . . . .	122
3-5	Optimal prices for various network topologies . . . . .	123
4-1	Illustration for a representative demand curve . . . . .	132
4-2	Profit and price ratios for the semi-log inverse demand curve as a func- tion of $c/P_m$ . . . . .	139
4-3	Profit and price ratios for the log-log inverse demand curve as a func- tion of $P_m/c$ . . . . .	141

4-4	Profit ratios as a function of $\epsilon$ for different demand curves . . . . .	144
4-5	Randomly generated inverse demand curve with $S = 5$ pieces . . . . .	146
4-6	Histogram of profit ratios when $S = 5$ for $c = 0$ and $c = 0.5P_m$ . . . . .	148
5-1	Order of events: 1. Subsidies; 2. Price and Quantity; 3. Sales . . . . .	163
5-2	Consumer surplus for stochastic demand, before demand rationing . . . . .	182
5-3	Optimal values as a function of the target level . . . . .	193
5-4	Relative normalized differences in subsidies (left) and supplier's profit (right) . . . . .	195
5-5	Expected sales relative to target . . . . .	196

# List of Tables

2.1	Forecast accuracy of two regression models for a brand of coffee . . .	44
2.2	A subset of the estimation results for two coffee brands . . . . .	68
2.3	Summary of the estimation results for four categories . . . . .	70
4.1	Price and profit ratios for several “true” demand curves . . . . .	141
4.2	Expected profit ratios when $\hat{P}_m = P_m(1 + \epsilon)$ and $\epsilon \sim U[-0.2, 0.2]$ . .	142
4.3	Profit ratios for randomly generated demands . . . . .	147
4.4	Expected profit ratios when the maximum price is unknown . . . . .	149



# Chapter 1

## Introduction

### 1.1 Summary

Finding the right pricing strategy has always been an intriguing and interesting topic for retailers and researchers. In the past three decades, tremendous progress was made, both in the academic and business worlds. It is not surprising anymore to observe prices that are dynamic in time, tailored to the individual and/or to the situation. On one hand, the goal is still the same and most of the time, retailers seek to maximize profits. On the other hand, the tools and models have been evolving very rapidly. It became very common to use data analytics and optimization methods in order to capture the psychology of consumers, their heterogeneity and to balance the trade-off between short term profits and long term strategies. In this spirit, this thesis investigates four novel applications of pricing.

In Chapter 2, we partnered with Oracle Retail in order to study the promotion optimization problem for supermarket retailers. One needs to decide which products to promote, what is the depth of price discounts and finally, when to schedule the promotions. Optimizing promotions while capturing consumer behavior (e.g., stockpiling behavior of consumers and cross-item effects such as substitution and complementarity) and incorporating practical business rules has not received much attention in the literature, yet it is a very important problem in practice. To capture the stockpiling

behavior of consumers, we propose two general classes of demand functions that can be easily estimated from data. Since the exact formulation is “hard”, we propose a linear approximation that allows us to solve the problem efficiently. We then show computationally that the formulation solves fast using actual data and present analytical results on the accuracy of the approximation relative to the optimal solution. Together with our collaborators, our framework allows to develop a tool that can help supermarket managers to schedule promotions by testing various strategies and business constraints. Finally, we calibrate our models using actual data and determine that they can improve profits by 3-5%.

The recent ubiquity of social networks has revolutionized the way people interact and influence each other. The overwhelming success of social networking platforms such as Facebook and Twitter allows firms to collect unprecedented volumes of data about their customers and their social interactions. The challenge that confronts every firm is how to process this data and turn it into actionable policies. In Chapter 3, we consider a setting where a monopolist sells an indivisible good to consumers embedded in a social network. We model the optimal pricing problem of the firm as a two-stage game. First, the firm designs prices to maximize its profits. Subsequently, the consumers choose whether to purchase the item or not. Assuming positive network externalities, we show the existence of a pure strategy Nash equilibrium. Using duality theory and integer programming techniques, we reformulate the problem into a linear mixed-integer program (MIP). We then derive efficient ways of optimally solving the MIP using its linear programming relaxation for two different pricing strategies: discriminative and uniform pricing. Finally, we present computational insights by comparing the pricing strategies and highlighting the benefits of incorporating social interactions in pricing models. In particular, we provide instances where it is beneficial for the seller to earn a negative profit on an influential agent in order to extract significant positive profits on others.

One may be interested in pricing a product for which demand information is very costly or even not available (the recent iWatch or new drugs are such examples).

This is the main motivation behind Chapter 4. We impose minimal assumptions on the information of the problem: that is, only the constant marginal cost and the maximal price the firm can set are known. We propose a simple way of pricing by approximating the true unknown inverse demand curve by a linear function for which one can easily compute the optimal price. We derive analytical bounds on the performance guarantee for different demand models and test computationally for a wide range of cases. This work presents an operational and very simple method to price items for which no demand information is available along with the guarantee that the profit is not far from optimal.

In Chapter 5, we introduce and study a model where the government sets subsidies for green technology adoption while considering the industry's response. In particular, we address the question: How does demand uncertainty affect the government, the industry and the consumers, when designing policies? Our methodology relies on a Stackelberg game where the supplier solves a price setting newsvendor problem. We show that demand uncertainty benefits the government and hurts the supplier. Interestingly, consumers do not always benefit from demand uncertainty. We also conclude that when policy makers ignore demand uncertainty, they can significantly miss the desired adoption target level. This research indicates that governments need to account for demand uncertainty while designing consumer subsidies.

## 1.2 Technical Contributions

In this thesis, we have been using methodologies and tools such as optimization, mathematical modeling, game theory and data analytics. We next describe in more details the technical contributions of each part of the thesis.

In Chapter 2, the promotion optimization problem was formulated as a non-linear integer program. The difficulty comes from the generality of the demand function. In particular, we consider time dependent and potentially highly non-linear demand models that are commonly used in practice. The presence of business rules as con-

straints in the optimization formulation makes the problem more realistic but also harder to solve in practice. In a recent paper, entitled “An Efficient Algorithm for Dynamic Pricing using Graph Theory” (Cohen, Gupta, Kalas, Perakis and Panchangam), the authors show that even the single item problem is NP-hard by reducing it from the Travel Salesman Problem. As a result, it motivates us to develop efficient ways of solving the problem with provably guarantees. In this thesis, starting with the single item problem, we present a method based on approximating the non-linear objective by the sum of the marginal contributions (i.e., having a single promotion at a time). This approximation leads to a linear integer program. We then show that the feasible region is totally unimodular so that one can solve the continuous relaxation of the linear problem. Next, we derive tight analytical parametric bounds on the quality of the approximation for two general classes of demand functions. The first class considers the effect of past prices on current demand to be multiplicative. Under a mild technical assumption, we show that for this class of models, the objective (i.e., the total profits) is submodular in the number of promotions. The second class assumes that the effect of past prices is additive and this leads to a supermodular objective. We then develop for each class of models a parametric bound in closed form to assess the performance of our approximation. We show that this bound is tight and that it performs very well in all the practical instances we tested. We then extend the bound for a unified demand model that can capture different segments of customers. Subsequently, we address the promotion problem for multiple items. Unfortunately, one cannot easily extend the previous results. Instead, we propose a class of approximation methods for which we develop two main analytical contributions. First, we propose an algorithm based on approximating the non-linear objective by the sum of marginal contributions (i.e., having a single promotion at a time) but also the sum of pairwise contributions (i.e., having two items on promotion simultaneously). We then prove that this approximation captures accurately the cross-item effects for any non-linear demand function with additive cross-item effects. Our proof is based on observing that the pairwise contributions are sufficient to characterize exactly all the

cross-effects. Second, for the case of multiple items, we note that the matrix of the constraint set is not totally unimodular anymore. In particular, incorporating the necessary consistency constraints (due to promoting two items simultaneously) ruins the integrality of the feasible region. Nevertheless, we show that under substitution effects (that is the most common case within any category of products), one can still solve the problem through a linear program. Geometrically, this means that some of the extreme points are fractional, but the structure of the objective implies that one can only converge to an integer vertex. Our proof is based on reducing our formulation to a quadratic binary problem and show that when the cross coefficients are non-negative, the problem can be solved by a linear programming reformulation. This result allows us to solve the promotion optimization problem for a large number of items very efficiently. Finally, we extend the analytical bounds for the case of multiple items.

In Chapter 3, we model the problem of selling an indivisible good to consumers embedded in a social network as a two-stage game. Using duality theory, we derive equilibrium constraints and reformulate the two-stage problem faced by the seller into a non-convex integer program. We then transform it into an equivalent mixed-integer program (MIP) using reformulation techniques from integer programming. This resulting MIP can be viewed as an operational pricing tool as any firm can easily incorporate business rules on prices and constraints on network segmentation. We then develop efficient and scalable methods to optimally solve the MIP for two pricing strategies using the linear programming relaxation. We consider the discriminative and uniform pricing strategies and present a solution method that is efficient (polynomial in the number of agents) and scalable to large networks. Finally, we show that the price of an agent that buys in the optimal discriminative pricing solution is the sum of its own value and a markup term that corresponds to the influence by the network of agents that buy. The seller offers this price to each agent depending on its own valuation for the item and the influence the agent exerts on the network. The seller's profit from network externalities in essence comes from two types of customers:

high valued customers who influence their neighbors and low valued customers who are highly influential. In addition, when comparing submodular, linear and supermodular influence models, we show that as we move from the submodular to the linear and then to the supermodular influence models, additional agents will buy. In addition, the buyers will pay a higher price so that it induces a larger profit for the seller.

In Chapter 4, we propose a simple pricing rule when demand information is very costly or even not available. Our method is based on approximating the true unknown inverse demand curve by a linear function. We derive analytical bounds on the performance guarantee for several demand models. For each demand form, we identify important structural properties (e.g., convexity versus concavity) that will drive the performance accuracy of our pricing rule. We then derive tight bounds in closed form for four commonly used demand models. We further extend our results for the case where the maximal price is not exactly known. In this case, our bounds depend explicitly on the confidence level of an available estimate of the maximal price. We then illustrate that with an error of  $\pm 20\%$  on the true maximal price, the performance of our pricing rule still yields a good performance on average. In all the cases we tested, our bounds guarantee a near optimal performance in terms of profits.

In Chapter 5, we study a model where the government sets subsidies for green technology adoption while considering the industry's response. We propose a Stackleberg game where each player solves sequentially its own optimization problem. We then proceed to solve the problem by backward induction. By showing some key monotonicity properties, we observe that the government problem can be solved to optimality by the tightness of the target adoption constraint. This allows us to solve the non-linear two-stage game and draw several interesting insights on the impact of demand uncertainty. For example, we determine that demand uncertainty does not always benefit consumers and that nonlinearity in demand plays a key role. We then show that the cost of demand uncertainty is shared between the supplier and the government. In particular, we analyze who bears the cost of demand uncertainty

between government and supplier, which we show depends on the profit margin of the product. In general, the government expenditure increases with the added inventory risk. For linear demand models, the cost of demand uncertainty shifts from the government to the supplier as the adoption target increases or the production cost decreases. Finally, we prove that consumer subsidies are a sufficient mechanism to coordinate the government and the supplier. More precisely, we compare the optimal policies to the case where a central planner manages jointly the supplier and the government. We determine that the price paid by the consumers and the production levels coincide for both the decentralized and the centralized models. In other words, consumer subsidies coordinate the supply-chain in terms of price and quantities.

### 1.3 Broader Impact

It is also important to indicate that this thesis investigates important practical problems motivated from real-world applications. In particular, two out of the four parts are in collaboration with industry practitioners: Chapter 2 is joint with Oracle Retail and Chapter 3 with IBM Research. The work in Chapter 2 is part of a broader collaboration between MIT and Oracle that ultimately allowed to develop a decision support tool to help retailers to schedule promotions. Using the models and optimization algorithms developed in this thesis as well as other tools, a supermarket retailer can solve the promotion optimization problem for a large number of items. The problem can be tailored to the client by calibrating our demand models with historical data. Then, one can solve the optimization problem very efficiently (running time in milliseconds) to obtain a recommendation on scheduling the promotions for the selling season. Given the tractability of our methods, one can solve several instances in order to test the robustness of the recommended solution. As we previously mentioned, extensive preliminary testing shows that using our model can improve profits by 3-5%. Note that for an industry such as groceries for which profit margins are very small, this is a very significant impact. The work in Chapter 3 has also an underlying

practical application. Very often, online retailers have access to a large volume of data about their customers and in particular, social networks data. The question is how to use this information and turn it into profitable business decisions (in our case, designing price incentives). One can definitely use social network data that are available to retailers in order to calibrate our models. We hope one day to implement and test the research conducted in Chapter 3 for a client of IBM Research. The tractability of our approach will allow the retailer to solve very large instances in a few seconds and to compute the optimal price incentives tailored to the influence of each customer in the social network. Finally, we also see the works in Chapters 4 and 5 as very timely problems motivated from important real-world applications. Our hope would be to have an impact in practice by providing key insights to decision makers. The work in Chapter 4 can be of great help to any firm that launches a new item (or service) for which demand information is very limited. The work in Chapter 5 may hopefully assist governments for policy decisions while designing consumer subsidies for green technology adoption.

# Chapter 2

## Promotion Optimization for Supermarket Retailers

### 2.1 Introduction

Sales promotions have become ubiquitous in various settings that include the grocery industry. During a sales promotion, the retail price of an item is temporarily lowered from the regular price, often leading to a large increase in sales volume. To illustrate how important promotions are in the grocery industry, we consider a study by A.C. Nielsen, which estimated that during January–June 2004, 12–25% of supermarket sales in five European countries were made during promotions.

Our own analysis also supports that promotions can be a key driver for increasing profits. We were able to obtain sales data from a large supermarket retailer for different categories of items. In Figure 2-1, we show the (normalized) prices and resulting sales for a particular brand of coffee in a single grocery store during a period of 35 weeks. One can see that this brand was promoted 8 out of 35 weeks (i.e., 23% of the time considered). In addition, the sales during promotions accounted for 41% of the total sales volume. Using a demand model estimated from real data (see Section 2.8.3 for details), we observe that the promotion prices of the retailer achieved a profit gain of 3% compared to using only the regular price (i.e., no promotions). A paper published by the Community Development Financial Institutions (CDFI)

Fund reports that the average profit margin for the supermarket industry was 1.9% in 2010. According to analysis of Yahoo! Finance data, the average net profit margin for publicly traded US-based grocery stores for 2012 is close to 2010’s 1.9% average. As a result, our finding suggests that promotions might make a significant difference in the retailer’s profits. Furthermore, it motivates us to build a model that answers the following question: How much money does the retailer leave on the table by using the implemented prices relative to “optimal” promotional prices?

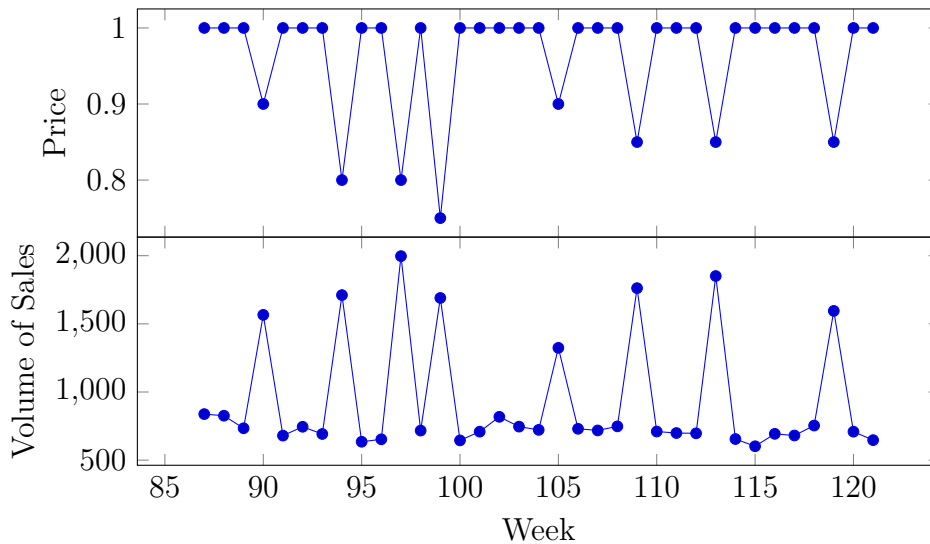


Figure 2-1: Prices and sales for a particular brand of coffee during 35 weeks

Given the importance of promotions in the grocery industry, it is not surprising that supermarkets pay great attention to how to design promotion schedules. The promotion planning process is complex and challenging for multiple reasons. First, demand exhibits a promotion fatigue effect, i.e., for certain categories of products, customers stockpile products during promotions, leading to reduced demand following the promotion. Second, promotions are constrained by a set of business rules specified by the supermarket and/or product manufacturers. Example of business rules include prices chosen from a discrete set, limited number of promotions and separating successive promotions (more details are provided in Section 2.3.1). Finally, the problem is difficult even for a single store because of its large scale - an average supermarket

has of the order of 40,000 SKUs, and the number of items on promotion at any time is about 2,000 leading to a large number of decisions that has to be made.

Despite the complexity of the promotion planning process, it is still to this day performed manually in most supermarket chains. This motivates us to design and study promotion optimization models that can make promotion planning more efficient (reducing man-hours) and at the same time more profitable (increasing profits and revenues) for supermarkets. To accomplish this, we introduce a Promotion Optimization Problem (POP) formulation and propose how to solve it efficiently. We introduce and study classes of demand functions that incorporate the features we discussed above as well as constraints that model important business rules. The output will provide optimized prices together with performance guarantees. In addition, thanks to the the short running times of our formulation, the manager can test various what-if scenarios to examine the robustness of the solution.

The POP formulation we introduce is a nonlinear IP and as a result, not computationally tractable. In practice, prices take values from a discrete price ladder (set of allowed prices at each time period) dictated by business rules. Even if we relax this requirement, the objective is in general neither concave nor convex due to the promotion fatigue effect.

We begin our analysis with the POP for a single item. This allows us to draw some insights on scheduling promotions and on how effects such as stockpiling and business rules play a role in promotions. In addition, this model can be applied to several categories of products for which cross item effects are not significant. Then, we extend our analysis to the POP with multiple items by including cross item effects in demand as well as cross item constraints, dictated by business rules.

We propose a linear IP approximation and show that the problem can be solved efficiently as an LP. This new formulation approximates the POP problem for any general demand and hence, any desired objective function. We also establish analytical lower and upper bounds relative to the optimal objective depending on the way past prices affect the demand. These results allow us to derive guarantees on

the performance of the LP approximation relative to the optimal POP solution. We show using actual data that the models run fast in practice and yield increased profits for the retailer by maintaining the same business rules. We start with the single item problem and then extend to multiple items and demonstrate good performance guarantees.

The impact of our models can be significant for supermarkets in practice. One of the goals of this research has been to develop data driven optimization models that can guide the promotion planning process for grocery retailers, including the clients of Oracle Retail. They span the range of Mid-market (annual revenue below \$1 billion) as well as Tier 1 (annual revenue exceeding \$5 billion and/or 250+ stores) retailers. One key challenge for implementing our models into software that can be used by grocery retailers is the large-scale nature of this industry. For example, a typical Tier 1 retailer has roughly 1000 stores, with 200 categories each containing 50-600 items. An important criterion for our models to be adopted by grocery retailers, is that the software solution needs to run in the order of a few seconds up to a minute. This is what has prompted us to reformulate our model as we discussed above as an LP. We apply our models (and tools) to two different clients of Oracle: one large supermarket chain and one US Farm & Home store chain. Preliminary tests using actual supermarket data, suggest that our model can increase profits by 3% just by optimizing the promotion schedule and up to 5% by slightly increasing the number of promotions allowed. If we assume that implementing the promotions recommended by our models does not require additional fixed costs (this seems to be reasonable as we only vary prices), then a 3% increase in profits for a retailer with annual profits of \$100 million translates into a \$3 million increase. As we previously discussed, profit margins in this industry are thin and therefore 3% profit improvement is significant. Finally, we are currently running a pilot with the second client.

## Contributions

This research was conducted in collaboration with our co-authors and industry practitioners from the Oracle Retail Science group, which is a business unit of Oracle Corporation. One of the end outcomes of this work is the development of sales promotion analytics that will be integrated into enterprise resource planning software for supermarket retailers.

- *We propose a POP formulation motivated by real-world retail environments.*

We introduce a nonlinear IP formulation for the POP. Unfortunately, this model is in general not computationally tractable. An important requirement from our industry collaborators is that an executive of a medium-sized supermarket (100 stores,  $\sim$ 200 categories,  $\sim$ 100 items per category) can run the tool (whose backbone is the model and algorithms we are developing in this work) and obtain a high quality solution in a few seconds. This motivates us to propose an LP approximation.

- *We propose an LP reformulation that allows us to solve the problem efficiently.*

We first introduce a linear IP approximation of the POP. We then show that the constraint matrix is totally unimodular and therefore, our formulation is tractable. Consequently, one can use the LP approximation to obtain a provably near-optimal solution to the original nonlinear IP formulation.

- *We introduce general classes of demand functions that capture promotion fatigue effects.* An important feature of the application domain is the promotion fatigue effect observed. We propose general classes of demand functions in which past prices have a multiplicative or an additive effect on current demand. These classes are generalizations of several models currently found in the literature, provide some extra modeling flexibility and can be easily estimated from data. We also propose a unified demand model that combines the multiplicative and additive models and as a result, can capture several consumer segments.

- *We develop bounds on performance guarantees for multiplicative and additive demand functions.* We derive upper and lower guarantees on the quality of the LP

approximation relative to the optimal (but intractable) POP solution and characterize the bounds as a function of the problem parameters. We show that for multiplicative demand, promotions have a submodular effect (for some relevant subsets of promotions). This leads to the LP approximation being an upper bound of the POP objective. For additive demand, promotions have a supermodular effect so that the LP approximation leads to a lower bound of the POP objective. Finally, we show the tightness of these bounds.

- *We consider the POP for multiple items, develop an efficient method to solve the problem and derive some useful insights.* We consider cross item effects in demand and formulate the POP for multiple items incorporating several cross item constraints dictated by business rules. We propose a linear IP approximation based on approximating the objective function by the unilateral and bilateral deviations. We show that this approximation yields a good solution relative to the optimal and extend the bounds on the performance guarantee. Finally, we study the interplay of cross item and stockpiling effects to draw useful insights on promotion planning.

- *We implement our models using actual data from two large Oracle clients and demonstrate the added value.* Our industry partners provided us with a collection of sales data from multiple stores and various categories from two of their clients. We apply our analysis to a few selected categories. In particular, we looked into coffee, tea, chocolate and yogurt. We first estimate the various demand parameters and then quantify the value of our LP approximation relative to the optimal POP solution. After extensive numerical testing with the clients' data, we show that the approximation error is in practice even smaller than the analytical bounds we developed. Our model provides supermarket managers recommendations for promotion planning with running times in the order of seconds. As the model runs fast and can be implemented on a platform like Excel, it allows managers to test and compare various strategies easily. We conclude the chapter by briefly discussing the ongoing pilot we have started.

## 2.2 Literature review

Our work is related to at least three streams of literature: optimization, marketing and dynamic pricing. We formulate the promotion optimization problem as a nonlinear mixed integer program (NMIP). In order to give users flexibility in the choice of demand functions, our POP formulation imposes very mild assumptions on the demand functions. Due to the general classes of demand functions we consider, the objective function is typically non-concave. In general, NMIPs are difficult from a computational complexity standpoint. Under certain special structural conditions (see, e.g., [17] and references therein), there exist polynomial time algorithms for solving NMIPs. However, many NMIPs do not satisfy these conditions and are solved using techniques such as Branch and Bound, Outer-Approximation, Generalized Benders and Extended Cutting Plane methods [14].

In a special instance of the POP when demand is a linear function of current and past prices and when discrete prices are relaxed to be continuous, one can formulate the POP as a Cardinality-Constrained Quadratic Optimization (CCQO) problem. It has been shown in [3] that a quadratic optimization problem with a similar feasible region as the CCQO is NP-hard. Thus, tailored heuristics have been developed in order to solve the problem (see for example, [2] and [3]). Our solution approach is based on linearizing the objective function by exploiting the discrete nature of the problem and then solving the POP as an LP. We note that due to the general nature of demand functions we consider, it is not possible to use linearization approaches such as in [24].

As we show later in this chapter, the POP for the two classes of demand functions we introduce is related to submodular and supermodular maximization. Maximizing an unconstrained supermodular function was shown to be a strongly polynomial time problem (see, e.g., [23]). However, in our case, we have several constraints on the promotions and as a result, it is not guaranteed that one can solve the problem efficiently to optimality. In addition, most of the proposed methods to maximize

supermodular functions are not easy to implement and are often not very practical in terms of running time. Indeed, our industry collaborators request solving the POP in at most few seconds and using an available platform like Excel. Unlike supermodular, maximization of submodular functions is generally NP-hard (see for example [20]). Several common problems, such as max cut and the maximum coverage problem, can be cast as special cases of this general submodular maximization problem under suitable constraints. Typically, the approximation algorithms are based on either greedy methods or local search algorithms. The problem of maximizing an arbitrary non-monotone submodular function subject to no constraints admits a  $1/2$  approximation algorithm. In addition, the problem of maximizing a monotone submodular function subject to a cardinality constraint admits a  $1 - 1/e$  approximation algorithm (e.g., [21]). In our case, we propose an LP approximation that does not request any monotonicity or other structure on the objective function. This LP approximation also provides guarantees relative to the optimal profits for two general classes of demand. Nevertheless, these bounds are parametric and not uniform. To compare them to the existing methods, we compute in Section 2.8 the values of these bounds on different demand functions estimated with actual data.

Sales promotions are an important area of research in the field of marketing (see [4] and the references therein). However, the focus in the marketing community is on modeling and estimating dynamic sales models (typically econometric or choice models) that can be used to derive managerial insights [9, 13]. For example, [13] studies parametric econometrics models based on scanner data to examine the dynamic effects of sales promotions. It is widely recognized in the marketing community that for certain products, promotions may have a *pantry-loading* or a *promotion fatigue* effect, i.e., consumers may buy additional units of a product during promotions for future consumption (stock piling behavior). This leads to a decrease in sales in the short term. In order to capture the promotion fatigue effect, many of the dynamic sales models that are used in the marketing literature have demand as a function of not just the current price, but also affected by past prices (see, e.g., [1, 19]). The

demand models used in our chapter can be seen as a generalization of the models used in these papers.

Our work is also related to the field of dynamic pricing (see for example, [25] and the references therein). An alternative method to model the promotion fatigue effect is a reference price demand model, which posits that consumers have a reference price for the product based on their memory of the past prices (see, e.g, [22, 18]). When consumers purchase the product, they compare the posted price to their internal reference price and interpret a discount or surcharge as a gain or a loss. The demand models considered in this work can be seen as a generalization of the reference price demand models as it includes several parameters to model the dependence of current demand in past prices.

The remainder of the chapter is structured as follows. In Section 5.3, we describe the model and assumptions as well as the business rules required for our problem. In Section 2.4, we formulate the Promotion Optimization Problem for a single item. In Section 2.5, we present an approximate formulation based on a linearization of the objective function, which gives rise to a linear IP. In Section 2.6, we consider multiplicative and additive demand models and show tight bounds on the LP approximation performance relative to the optimal POP solution. Section 2.7 extends the treatment for the POP and the bounds for multiple items. Finally, Section 2.8 presents the implementation results using actual data. The proofs of the different propositions, lemmas and theorems are relegated to the Appendix. However, we refer the reader to [8] and [7] for more details.

## 2.3 Model and Assumptions

In what follows, we consider the Promotion Optimization Problem for a single item. Note that solving this problem is important as one can use the results for the single item model for addressing the multiple product case. We extend our analysis and consider the multiple item scenario in Section 2.7. The manager's objective is to

maximize the total profits during some finite time horizon, whereas the decision variables are for each time period, whether to promote a product and what price to set (i.e., the promotion depth). In our formulation, we also incorporate various important real-world business requirements that should be satisfied (a complete description is presented in Section 2.3.1). We first introduce some notation:

- $T$  - Number of weeks in the horizon (e.g., one quarter composed of 13 weeks).
- $L$  - Limitation on the number of times we are allowed to promote.
- $S$  - Number of separating periods (restricting the separation between two successive promotions).
- $\mathcal{Q} = \{q^0 > q^1 > \dots > q^k > \dots > q^K\}$  - Price ladder, i.e., the discrete set of admissible prices.
- $q^0$  - Regular (non-promoted) price, which is the maximum price in the price ladder.
- $q^K$  - Minimum price in the price ladder.
- $c_t$  - Unit cost of the item at time  $t$ .

The decision variables are the prices set at each time period denoted by  $p_t \in \mathcal{Q}$ . Since we are considering a set of discrete prices only (motivated by the business requirement of a finite price ladder, see Section 2.3.1), one can rewrite the price  $p_t$  at time  $t$  as follows:

$$p_t = \sum_{k=0}^K q^k \gamma_t^k, \quad (2.3.1)$$

where  $\gamma_t^k$  is a binary variable that is equal to 1 if the price  $q^k$  is selected from the price ladder at time  $t$  and 0 otherwise. This way, the decision variables are now the set of binary variables  $\gamma_t^k$ ;  $\forall t = 1, \dots, T$  and  $\forall k = 0, \dots, K$ , for a total of  $(K + 1)T$

variables. In addition, we require the following constraint to ensure that exactly a single price is selected at each time  $t$ :

$$\sum_{k=0}^K \gamma_t^k = 1; \quad \forall t. \quad (2.3.2)$$

Finally, we consider a general time-dependent demand function denoted by  $d_t(\mathbf{p}_t)$  that explicitly depends on the current price and up to  $M$  past prices  $p_t, p_{t-1}, \dots, p_{t-M}$  as well as on demand seasonality and trend. We will consider specific demand forms later in this chapter.  $M \in \mathbb{N}_0$  denotes the memory parameter that represents the number of past prices that affect the demand at time  $t$ :

$$d_t(\mathbf{p}_t) = h_t(p_t, p_{t-1}, \dots, p_{t-M}). \quad (2.3.3)$$

We next describe the various business rules we incorporate in our formulation.

### 2.3.1 Business Rules

1. *Promotion fatigue effect.* It is well known that when the price is reduced, consumers tend to purchase larger quantities. This can lead to a larger consumption for particular products but also can imply a stockpiling effect (see, e.g., [1]). In other words, for particular items, customers will purchase larger quantities for future consumption (e.g., toiletries or non-perishable goods). Therefore, due to the consumer stockpiling behavior, a sales promotion for a product increases the demand at the current period but also reduces the demand in subsequent periods, with the demand slowly recovering over time to the nominal level, that is no promotion. This effect is illustrated in Figure 2-2, where the promotion in week 3 yields a boost in current demand but also decreases demand in the following weeks. Finally, demand gradually recovers up to the nominal level (no promotion). We propose to capture this effect by a demand model that explicitly depends on the current price  $p_t$  and on the past prices  $p_{t-1}, p_{t-2}, \dots, p_{t-M}$ . In addition, our models allow to have the flexibility of assigning

different weights to reflect how strongly a past price affects the current demand. The parameter  $M$  represents the memory of consumers with respect to past prices and varies depending on several features of the item. In practice, the parameter  $M$  can be estimated from data (see Section 2.8).

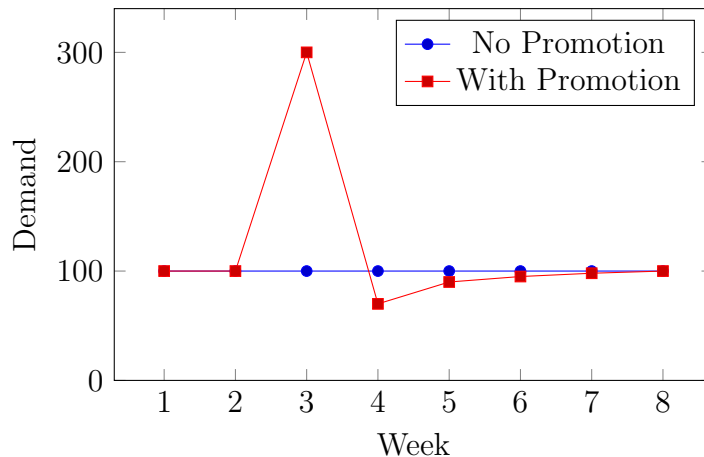


Figure 2-2: Example of the promotion fatigue effect

2. *Prices are chosen from a discrete price ladder.* For each product, there is a finite set of permissible prices. For example, prices may have to end with a ‘9’. In addition, the price ladder for an item can be time-dependent. This requirement is captured explicitly by equation (2.3.1), where the price ladder is given by:  $q^0 > q^1 > \dots > q^K$ . In other words, the regular price  $q^0$  is the maximal price and the price ladder has  $K + 1$  elements. For simplicity, we assume that the elements of the price ladder are time independent but note that this assumption can be relaxed.

3. *Limited number of promotions.* The supermarket may want to limit the frequency of the promotions for a product. This requirement applies because retailers wish to preserve the image of the store/brand. For example, it may be required to promote a particular product at most  $L = 3$  times during the quarter. Mathematically, one can impose the following constraint:

$$\sum_{t=1}^T \sum_{k=1}^K \gamma_t^k \leq L. \quad (2.3.4)$$

4. *Separating periods between successive promotions.* A common additional requirement is to space out promotions by a minimal number of separating periods, denoted by  $S$ . Indeed, if successive promotions are too close to one another, this may hurt the store image and incentivize consumers to behave more as deal-seekers. Mathematically, one can impose the following constraint:

$$\sum_{\tau=t}^{t+S} \sum_{k=1}^K \gamma_{\tau}^k \leq 1 \quad \forall t. \quad (2.3.5)$$

### 2.3.2 Assumptions

We assume that at each period  $t$ , the retailer orders the item from the supplier at a linear ordering cost that can vary over time, i.e., each unit sold in period  $t$  costs  $c_t$ . This assumption holds under the conventional wholesale price contract which is frequently used in practice as well as in the academic literature (see for example, [6]). We also consider the demand to be specified by a deterministic function of current and past prices. This assumption is justified because we capture the most important factors that affect demand (current and past prices), therefore the estimated demand models are accurate in the sense of having low forecast error (see estimation results in Section 2.8 and Figure 2-6). Since the estimated deterministic demand functions seem to accurately model actual demand, for this application, we can use them as input into the optimization model without taking into account demand uncertainty. Indeed, the typical process in practice is to estimate a demand model from data and then to compute the optimal prices based on the estimated demand model. In Section 2.8, we start with actual sales data from a supermarket, estimate a demand model and finally compute the optimal prices using our model. The demand models we consider are commonly used both by practitioners and the academic literature (see e.g., [19] and [12]).

Finally, we assume that the retailer always carries enough inventory to meet demand, so that in each period, sales are equal to demand. The above assumption is

reasonable in our setting because grocery retailers are aware of the negative effects of stocking out of promoted products (see e.g., [10]) and use accurate demand estimation models (e.g., [9] and [26]) in order to forecast demand and plan inventory accordingly. We hence use the terms demand and sales interchangeably in this chapter.

To the best of our knowledge, this work is perhaps the first to develop a model that incorporates the aforementioned features for the POP and propose an efficient solution. These features not only introduce challenges from a theoretical perspective, but also are important in practice.

## 2.4 Problem Formulation

In what follows, we formulate the single-item Promotion Optimization Problem (POP) incorporating the business rules we discussed above:

$$\begin{aligned}
& \max_{\gamma_t^k} && \sum_{t=1}^T (p_t - c_t) d_t(\mathbf{p}_t) \\
& \text{s.t.} && p_t = \sum_{k=0}^K q^k \gamma_t^k \\
& && \sum_{t=1}^T \sum_{k=1}^K \gamma_t^k \leq L \\
& && \sum_{\tau=t}^T \sum_{k=1}^K \gamma_\tau^k \leq 1 \quad \forall t \\
& && \sum_{k=0}^K \gamma_t^k = 1 \quad \forall t \\
& && \gamma_t^k \in \{0, 1\} \quad \forall t, k
\end{aligned} \tag{POP}$$

Note that the only decisions are which price to choose from the discrete price ladder at each time period (i.e., the binary variables  $\gamma_t^k$ ). We denote by  $POP(\mathbf{p})$  (or equivalently  $POP(\boldsymbol{\gamma})$ ) the objective function of (POP) evaluated at the vector  $\mathbf{p}$  (or equivalently  $\boldsymbol{\gamma}$ ). This formulation can be applied to a general time-dependent demand function  $d_t(\mathbf{p}_t)$  that explicitly depends on the current price  $p_t$ , and on the  $M$

past prices  $p_{t-1}, \dots, p_{t-M}$  as well as on demand seasonality and trend (see equation (2.3.3)). Specific examples are presented in Section 2.6.

In Figure 2-3, we plot the profit function for a small example with two time periods, where the demand functions at times 1 and 2 follow the following relations:  $\log d_1(p_1) = \log a_1 + \beta_1 \log p_1 + \beta_2 \log \frac{q^0 - p_1}{q^0}$  and  $\log d_2(p_2, p_1) = \log a_2 + \beta_1 \log p_2 + \beta_2 \log \frac{p_1 - p_2}{2q^0}$ . Here, we used  $a_1 = 100, a_2 = 200$  and  $\beta_2 = -\beta_1 = 4$ . The regular price, costs and minimum price are given by  $q^0 = 100, c_1 = c_2 = 50$  and  $q^K = 50$  respectively. This illustrates the fact that the POP is a nonlinear IP and is in general hard to solve to optimality even for very special instances. Even getting a high-quality approximation may not be an easy task. First, even if we were able to relax the prices to take non-integer values, the objective is in general non-linear (neither concave nor convex) due to the cross time dependence between prices (see Figure 2-3). Second, even if the objective is linear, there is no guarantee that the problem can be solved efficiently using an LP solver because of the integer variables. We propose in the next section an approximation based on a linear programming reformulation of the POP.

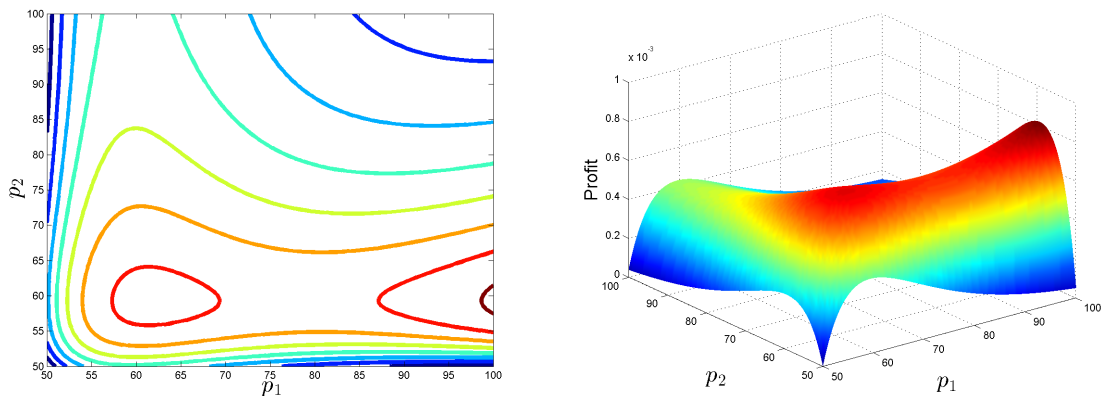


Figure 2-3: Profit function for demand with promotion fatigue effect

## 2.5 IP Approximation

By looking carefully at several data sets, we have seen that for many products, promotions often last only for one week, and two consecutive promotions are at least 3 weeks apart. If the promotions are subject to a separating constraint as in equation (2.3.5), then the interaction between successive promotions is fairly weak. Therefore, by ignoring the second-order interactions between promotions and capture only the direct effect of each promotion, we introduce a linear IP formulation that should give us a “good” solution. More specifically, we approximate the nonlinear POP objective by a linear approximation based on the sum of unilateral deviations. In order to derive the IP formulation of the POP, we first introduce some additional notation. For a given price vector  $\mathbf{p} = (p_1, \dots, p_T)$ , we define the corresponding total profits throughout the horizon:

$$POP(\mathbf{p}) = \sum_{t=1}^T (p_t - c_t) d_t(\mathbf{p}_t).$$

Let us now define the price vector  $\mathbf{p}_t^k$  as follows:  $(\mathbf{p}_t^k)_\tau = \begin{cases} q^k; & \text{if } \tau = t \\ q^0; & \text{otherwise} \end{cases}$ . In other words, the vector  $\mathbf{p}_t^k$  has the promotion price  $q^k$  at time  $t$  and the regular price  $q^0$  (no promotion) is used at all the remaining periods. We denote the regular price vector by  $\mathbf{p}^0 = (q^0, \dots, q^0)$ , for which the regular price is set at all the periods. Let us define the coefficients  $b_t^k$  as:

$$b_t^k = POP(\mathbf{p}_t^k) - POP(\mathbf{p}^0). \quad (2.5.1)$$

These coefficients represent the unilateral deviations in total profits by applying a single promotion. One can compute these *TK* coefficients before starting the optimization procedure. Since these calculations can be done off-line, they do not affect the complexity of the optimization. We are now ready to formulate the IP approxi-

mation of the POP:

$$\begin{aligned}
& POP(\mathbf{p}^0) + \max_{\gamma_t^k} \sum_{t=1}^T \sum_{k=1}^K b_t^k \gamma_t^k \\
\text{s.t.} \quad & \sum_{t=1}^T \sum_{k=1}^K \gamma_t^k \leq L \\
& \sum_{\tau=t}^{t+S} \sum_{k=1}^K \gamma_\tau^k \leq 1 \quad \forall t \\
& \sum_{k=1}^K \gamma_t^k = 1 \quad \forall t \\
& \gamma_t^k \in \{0, 1\} \quad \forall t, k
\end{aligned} \tag{IP}$$

**Remark 2.1.** *One can condense the above IP formulation in a more compact way. In particular, since at most one of the decision variables  $\{\gamma_t^k : k = 1, \dots, K\}$  is equal to one, one can define  $\tilde{b}_t = \max_{k=1, \dots, K} b_t^k$ ;  $\forall t = 1, \dots, T$  and replace the double sums by single sums. As a result, we obtain a knapsack type formulation. Since both formulations are equivalent, we consider the (IP) above.*

As we discussed, the IP approximation of the POP is obtained by linearizing the objective function. More specifically, we approximate the POP objective by the sum of the unilateral deviations by using a single promotion. Note that this approximation neglects the pairwise interactions of two promotions but still captures the promotion fatigue effect. We observe that the constraint set remains unchanged, so that the feasible region of both problems is the same. We also note that all the business rules from the constraint set are modeled as linear constraints. Consequently, the IP formulation is a linear problem with integer decision variables. As we mentioned, the IP approximation becomes more accurate when the number of separating periods  $S$  becomes large. In addition, the IP solution is optimal when there is no correlation between the time periods, or when  $L = 1$ . The instances where the IP is optimal are summarized in the following Proposition.

**Proposition 2.1.** *Under either of the following four conditions, the IP approximation coincides with the POP optimal solution. a) Only a single promotion is allowed, i.e.,*

$L = 1$ . b) Demand at time  $t$  depends only on the current price  $p_t$  and not on past prices (i.e.,  $M = 0$ ). c) The number of separating periods is at least equal to the memory (i.e.,  $S \geq M$ ).

In general, solving an IP can be difficult from a computational complexity standpoint. In our numerical experiments, we observed that Gurobi solves (IP) in less than a second. The reason is that (IP) has an integral feasible region and therefore can be solved efficiently as an LP, as we show in the following Theorem. The feasible region of both (POP) and (IP) is given by:

$$\left\{ \gamma_t^k : \sum_{t=1}^T \sum_{k=1}^K \gamma_t^k \leq L \quad \forall t; \sum_{\tau=t}^{t+S} \sum_{k=1}^K \gamma_\tau^k \leq 1; \sum_{k=0}^K \gamma_t^k = 1 \quad \forall t \right\}. \quad (2.5.2)$$

**Proposition 2.2.** *Every basic feasible solution of (2.5.2) is integral.*

Using Proposition 2.2, one can solve (IP) efficiently by solving its LP relaxation, given by:

$$\begin{aligned} POP(\mathbf{p}^0) &+ \max_{\gamma_t^k} \sum_{t=1}^T \sum_{k=1}^K b_t^k \gamma_t^k \\ \text{s.t.} & \sum_{t=1}^T \sum_{k=1}^K \gamma_t^k \leq L \\ & \sum_{\tau=t}^{t+S} \sum_{k=1}^K \gamma_\tau^k \leq 1 \quad \forall t \\ & \sum_{k=0}^K \gamma_t^k = 1 \quad \forall t \\ & 0 \leq \gamma_t^k \leq 1 \quad \forall t, k \end{aligned} \quad (\text{LP})$$

This allows us to obtain an approximation solution for the POP efficiently. From now on, we refer to (IP) as the LP approximation and denote its optimal solution by  $\gamma^{LP}$ . In addition,  $LP(\mathbf{p})$  (or equivalently  $LP(\gamma)$ ) denotes the objective function of (LP) evaluated at the vector  $\mathbf{p}$  (or equivalently  $\gamma$ ). The question is how does this LP approximation compare relative to the optimal POP solution. To address this question, we next consider two cases depending on the demand structure. First

though, we propose some “reasonable” demand models in this application area.

## 2.6 Demand Models

In this section, we introduce two classes of demand functions. They incorporate the promotion fatigue effect we previously discussed. We next analyze supermarket sales data to support and validate the existence of the promotion fatigue effect in some items and categories. We report only a brief analysis here but a detailed description of the data will be presented in Section 2.8.

We divide the 117 weeks of data into a training set of 82 weeks and a testing set of 35 weeks. Below we consider a log-log demand model (see (2.8.1)). The latter is commonly used in industry (for example, by Oracle Retail) and in academia (see [19]). We then estimate two versions of the model. Model 1 is estimated under the assumption that there is no promotion fatigue effect, i.e., the memory parameter  $M = 0$  in (2.8.1), so that the current demand  $d_t$  depends only on the current price  $p_t$  and not on past prices. Model 2 includes the promotion fatigue effect with a memory of two weeks, i.e.,  $M = 2$  in (2.8.1) so that the current demand  $d_t$  depends on the current price  $p_t$  and the prices in the two prior weeks  $p_{t-1}$  and  $p_{t-2}$ .

We summarize the regression results for a particular brand of coffee (the exact name of the brand cannot be explicitly unveiled due to confidentiality). We find that the estimated price elasticity coefficients of  $p_{t-1}$  and  $p_{t-2}$  for Model 2 are statistically significant. As a result, this supports the existence of the promotion fatigue effect for this item. In addition, we find that Model 2 has a significantly smaller forecast error relative to Model 1 (see Table 2.1). Note that the different forecast metrics used in Table 2.1 are defined in Section 2.8. The estimated demand model for this coffee brand follows the following relation:

$$\log d_t = \beta^0 + \beta^1 t + \beta^2 WEEK_t - 3.277 \log p_t + 0.518 \log p_{t-1} + 0.465 \log p_{t-2}. \quad (2.6.1)$$

Here,  $\beta^0$  and  $\beta^1$  denote the brand intercept and the trend coefficient respectively.  $\beta^2 = [\beta_t^2]$ ;  $t = 1, \dots, 52$  is a vector with seasonality coefficients for each week of the year.

	Model 1	Model 2
MAPE	0.145	0.116
OOS $R^2$	0.827	0.900
Revenue Bias	1.069	1.059

Table 2.1: Forecast accuracy of two regression models for a brand of coffee

In the remainder of this section, motivated by the above finding, i.e., that there are promotion fatigue effects in the demand, we introduce and study more general classes of demand models inspired by equation (2.6.1). We introduce the following notation that will be used in the sequel. Let  $A = \{(t_1, k_1), \dots, (t_N, k_N)\}$  with  $N \leq L$  be a set of promotions with  $1 \leq t_1 < t_2 < \dots < t_N \leq T$ . In other words, at each time period  $t_n$ ;  $\forall n = 1, \dots, N$  the promotion price  $q^{k_n}$  is used, whereas at the remaining time periods, the regular price  $q^0$  (no promotion) is set. We define the price vector associated with  $A$  as:

$$(\mathbf{p}_A)_t = \begin{cases} q^{k_n} & \text{if } t = t_n \text{ for some } n = 1, \dots, N; \\ q^0 & \text{otherwise.} \end{cases}$$

To further illustrate the above definition, consider the following example.

**Example.** Suppose that the price ladder is given by  $\mathcal{Q} = \{q^0 = 5 > q^1 = 4 > q^2 = 3\}$ , and the time horizon is  $T = 5$ . Suppose that the set of promotions  $A = \{(1, 1), (3, 2)\}$ , that is we have two promotions at times 1 and 3 with prices  $q^1$  and  $q^2$  respectively. Then,  $\mathbf{p}_A = (q^1, q^0, q^2, q^0, q^0) = (4, 5, 3, 5, 5)$ . We define the indicator variables corresponding to the set of promotions  $A$  as follows:  $(\gamma_A)_t^k =$

$\begin{cases} 1 & \text{if } (\mathbf{p}_A)_t = q^k; \\ 0 & \text{otherwise} \end{cases}$ . Note that matrix  $(\gamma_A)_t^k$  has dimensions  $(K + 1) \times T$ . In the

previous example, we have:  $\gamma_A = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$ . Recall that the LP objective

function is given by:

$$LP(\gamma) = POP(\mathbf{p}^0) + \sum_{t=1}^T \sum_{k=1}^K b_t^k \gamma_t^k, \quad (2.6.2)$$

where  $b_t^k$  is defined in (2.5.1). Finally, we denote by  $\tilde{L}$  the effective maximal number of promotions:  $\tilde{L} = \min\{L, \tilde{N}\}$ , where  $\tilde{N} = \lfloor \frac{T-1}{S+1} \rfloor + 1$ . We assume that  $L \geq 1$  (the case of  $L = 0$  is not interesting as no promotions are allowed). Since  $\tilde{N} \geq 1$ , we also have  $\tilde{L} \geq 1$ .

## 2.6.1 Multiplicative Demand

In this section, we assume that past prices have a multiplicative effect on current demand, so that the demand at time  $t$  can be expressed by:

$$d_t = f_t(p_t) \cdot g_1(p_{t-1}) \cdot g_2(p_{t-2}) \cdots g_M(p_{t-M}). \quad (2.6.3)$$

Note that the current price elasticity along with the seasonality and trend effects are captured by the function  $f_t(p_t)$ . The function  $g_k(p_{t-k})$  captures the effect of a promotion  $k$  periods before the current period, i.e., the effect of  $p_{t-k}$  on the demand at time  $t$ .  $M$  represents the memory of consumers with respect to past prices and can be estimated from data. As we verify in Section 2.8 from the actual data, it is reasonable to assume the following for the functions  $g_k(\cdot)$ .

*Assumption 1.* 1. Past promotions have a multiplicative reduction effect on current demand, i.e.,  $0 < g_k(p) \leq 1$ .

2. Deeper promotions result in larger reduction in future demand, i.e., for  $p \leq q$ ,

we have:  $g_k(p) \leq g_k(q) \leq g_k(q^0) = 1$ .

3. The reduction effect is non-increasing with time after the promotion:  $g_k$  is non-decreasing with respect to  $k$ , i.e.,  $g_k(p) \leq g_{k+1}(p)$ .

We assume that for  $k > M$ ,  $g_k(p) = 1 \forall p$ , so that no effects are present after  $M$  periods.

**Remark 2.2.** Equation (2.6.3) represents a general class of demand models, which admits as special cases several models that are used in practice. For example, the demand model of [19] with only pre-promotion effects that is of the form:  $\log d_t = a_0 + a_1 \log p_t + \sum_{u=1}^{\tau} \log \beta_u \log p_{t-u}$ .

Next, we present upper and lower bounds on the performance guarantee of the LP approximation relative to the optimal POP solution for the demand model in (2.6.3).

### Bounds on Quality of Approximation

**Theorem 2.1.** Let  $\gamma^{POP}$  be an optimal solution to (POP) and  $\gamma^{LP}$  be an optimal solution to (LP). Then:

$$1 \leq \frac{POP(\gamma^{POP})}{POP(\gamma^{LP})} \leq \frac{1}{\underline{R}}, \quad (2.6.4)$$

where  $\underline{R}$  is defined by:  $\underline{R} = \prod_{i=1}^{\tilde{L}-1} g_{i(S+1)}(q^K)$ , with  $\underline{R} = 1$  by convention, if  $\tilde{L} = 1$ .

Theorem 2.1 relies on the following two results.

**Lemma 2.1** (Submodular effect of the last promotion on profits).

1. Let  $A = \{(t_1, k_1), \dots, (t_n, k_n)\}$  be a set of promotions with  $t_1 < t_2 < \dots < t_n$  ( $n \leq L$ ) and let  $B \subset A$ . Consider a new promotion  $(t', k')$  with  $t_n < t'$ . If the new promotion  $(t', k')$ , when added to  $A$ , yields larger profits than  $\mathbf{p}_A$ , that is:

$$POP(\gamma_{A \cup \{(t', k')\}}) \geq POP(\gamma_A), \quad (2.6.5)$$

then the promotion  $(t', k')$  yields a larger marginal profit increase for  $\mathbf{p}_B$  than for  $\mathbf{p}_A$ , that is:

$$POP(\gamma_{A \cup \{(t', k')\}}) - POP(\gamma_A) \leq POP(\gamma_{B \cup \{(t', k')\}}) - POP(\gamma_B). \quad (2.6.6)$$

2. Let  $\gamma^{POP}$  be an optimal solution for the POP. Then:  $POP(\gamma^{POP}) \leq LP(\gamma^{POP})$ .

Note that if (2.6.5) is not satisfied, the sub-additivity property of Lemma 2.1 does not necessarily hold for any feasible solution. However, the required condition in (2.6.5) is automatically satisfied for the optimal POP solution. Lemma 2.1 states that for a multiplicative demand model as in (2.6.3), the POP profits are submodular in promotions (for certain relevant sets of promotions). Consequently, it supports intuitively the fact that the LP approximation overestimates the POP objective, i.e.,  $POP(\gamma^{POP}) \leq LP(\gamma^{POP})$ .

**Proposition 2.3.** *For any feasible vector  $\gamma$ , we have:  $POP(\gamma) \geq \underline{R} \cdot LP(\gamma)$ .*

Proposition 2.3 provides a lower bound for the POP objective function by applying the linearization and compensating by the worst case aggregate factor, that is  $\underline{R}$ .

Using the results of Theorem 2.1, one can solve the LP approximation (efficiently) and obtain a guarantees relative to the optimal POP solution. These bounds are parametric and can be applied to any general demand model in the form of equation (2.6.3). In addition, as we illustrate in Section 2.6.1, these bounds perform well in practice for a wide range of parameters. We next show that the bounds of Theorem 2.1 are tight.

**Proposition 2.4** (Tightness of the bounds for multiplicative demand).

1. *The lower bound in Theorem 2.1 is tight. More precisely, for any given price ladder,  $L, S$  and functions  $g_k$ , there exist  $T$ , costs  $c_t$  and functions  $f_t$  such that:*

$$POP(\gamma^{POP}) = POP(\gamma^{LP}).$$

2. The upper bound in Theorem 2.1 is asymptotically tight. For any given price ladder,  $S$  and functions  $g_k$ , there exists a sequence of promotion optimization problems  $\langle POP^n \rangle_{n=1}^\infty$ , each with a corresponding LP solution  $\gamma_n^{LP}$  and optimal POP solution  $\gamma_n^{POP}$  such that:

$$\lim_{n \rightarrow \infty} \frac{POP^n(\gamma_n^{POP})}{POP^n(\gamma_n^{LP})} = \frac{1}{\underline{R}_\infty}.$$

### Illustrating the bounds

We summarize the main findings regarding the behavior and quality of the bounds we have developed in the previous section. Recall that solving the POP can be hard in practice and one can instead implement the LP solution. The resulting profit is equal to  $POP(\gamma^{LP})$ , whereas in theory, we could have obtained a maximum profit equal to the optimal POP profits denoted by  $POP(\gamma^{POP})$ . In our numerical experiments, we examine the gap between  $POP(\gamma^{LP})$  and  $POP(\gamma^{POP})$  as a function of various parameters of the problem. In addition, we compare the ratio between  $POP(\gamma^{POP})$  and  $POP(\gamma^{LP})$  relative to the upper bound in Theorem 2.1 equal to  $1/\underline{R}$ . As we previously noted, the bounds depend on four different parameters: the number of separating periods  $S$ , the number of promotions allowed  $L$ , the effect of past prices (i.e., the value of the memory parameter  $M$  as well as the magnitude of the functions  $g_k(\cdot)$ ) and the minimal price  $q^K$ . Below, we summarize the effect of each of these factors for the following demand:  $\log d_t(\mathbf{p}) = \log(10) - 4 \log p_t + 0.5 \log p_{t-1} + 0.3 \log p_{t-2} + 0.2 \log p_{t-3} + 0.1 \log p_{t-4}$  with  $T = 9$ .

One can see that: a) In most cases, the LP solution achieves a profit that is very close to the optimal profit. In particular, the actual optimality gap (between the POP objective at optimality versus evaluated at the LP solution) seems to be of the order of 1-2 % and is smaller than the upper bound in Theorem 2.1. b) The upper bound  $1/\underline{R}$  varies between 0 and 33% depending on the values of the parameters. c) As  $S$  increases, the upper bound  $1/\underline{R}$  improves. Indeed, the promotions are further apart in time, reducing the interaction between promotions and improving the quality of

the LP approximation. For values of  $S \geq 1$ , the upper bound is at most 23% in this example. In practice, typically the number of separating periods is at least 1 but often 2-4 weeks. *d)* For values of  $L$  between 1 and 8, the upper bound is at most 23% in this example. *e)* The upper bound decreases with  $q^K$  and is at most 33% when a 50% promotion is allowed. If we restrict to a maximum of 30% promotion price, the bound becomes 14%. *f)* The upper bound increases with the memory parameter  $M$  and is at most 23% in this example. A more detailed illustration of the bounds can be found in [8].

## 2.6.2 Additive Demand

Our analysis of the sales data suggests that for some products, one needs to consider a demand model where the effect of past prices on current demand is additive. Motivated by this observation, we also propose and study a class of additive demand functions. Suppose that past prices have an additive effect on current demand, so that the demand at time  $t$  is given by:

$$d_t = f_t(p_t) + g_1(p_{t-1}) + g_2(p_{t-2}) + \cdots + g_M(p_{t-M}). \quad (2.6.7)$$

We extend the treatment for the additive model in (2.6.7) by developing upper and lower tight bounds as in Section 2.6.1. Finally, we also consider a unified demand model that captures a pool of consumers with different segments identified from the data. More specifically, consumers are partitioned into segments, such as loyal and non-loyal members and each group can have a multiplicative or an additive demand function. The analysis and results for these two cases can be found in [8].

## 2.7 Multiple Items

In several product categories, cross-item effects on demand can be significant. More precisely, a promotion in a particular item does not only affect the sales of this item

but also the sales of some other items in the category. One can distinguish between two different types of cross-item effects: substitutability (sometimes also called cannibalization) and complementarity (sometimes also called the halo effect). Two products are substitutable if for example, there are two competing brands of the same item, e.g., Coke and Pepsi. In this case, it is clear that a promotion in Coke potentially increases the Coke’s sales but also decreases the Pepsi’s sales. This follows from the fact that some customers might be indifferent between the two products and are likely to switch from one brand to the other, when it is on promotion. Mathematically, one can assume that if items  $i$  and  $j \neq i$  are substitutable,  $\frac{\partial d_i}{\partial p_j} \geq 0$  and  $\frac{\partial d_j}{\partial p_i} \leq 0$ . Two products  $i$  and  $j$  are complements if the consumption of  $i$  induces purchasing item  $j$  (and vice versa), e.g., shampoo and conditioner. In this case, it is clear that when the shampoo is on promotion, it potentially increases its own sales but also increases the sales of the conditioner, as people typically buy both items together. Mathematically, one can assume that if items  $i$  and  $j \neq i$  are complements,  $\frac{\partial d_i}{\partial p_j} \leq 0$  and  $\frac{\partial d_j}{\partial p_i} \leq 0$ . In what follows, we formulate the promotion optimization problem for multiple items that incorporate cross-item effects such as substitution and complementarity.

### 2.7.1 Problem Formulation

We consider a setting with a category composed of  $N$  products and denote by  $d_t^i(\cdot)$  the demand of item  $i$  at time  $t$ . We assume without loss of generality that all the items have the same regular price  $q^0$ , as one can normalize it for each item and assume conveniently that  $q^0 = 1$ . We still consider that demand equals sales (equivalently, that the retailer carries enough inventory for each product). In addition, we impose the following assumption on the demand function.

*Assumption 2.* The demand function depends explicitly on self past and current prices, and on cross current prices.

Assumption 2 implies that the demand does not depend explicitly on cross past prices. In other words, the demand of item  $i$  does not depend on the past prices of

the other items in the category. This assumption is supported by the fact that most consumers may be loyal to a particular brand. As a result, they can stockpile for this item while on promotion. They are also aware of all the prices of the other items at time  $t$ , so they can decide to potentially switch and purchase another product. However, consumers usually do not remember the past prices of the other items in the category. In addition, we tested this assumption using actual data from a large retailer and observed that in the vast majority of the cases we tested, the parameters for the cross past prices are not statistically significant. As a result, this assumption is supported by both the psychology of consumers and the actual data from several stores and categories of products we have considered. The demand of item  $i$  at time  $t$  can be any non-linear and time dependent function of the form:  $d_t^i(p_t^i, p_{t-1}^i, \dots, p_{t-M_i}^i, \mathbf{p}_t^{-i})$ , where,  $M_i$  represents the memory parameter of item  $i$  and  $\mathbf{p}_t^{-i}$  denotes the vector of prices at time  $t$  of all the items in the category but  $i$ .

One can now formulate the POP for multiple items (labeled as Multi-POP):

$$\begin{aligned}
& \max_{\gamma_t^{ik}} \sum_{i=1}^N \sum_{t=1}^T (p_t^i - c_t^i) d_t^i(p_t^i, p_{t-1}^i, \dots, p_{t-M_i}^i, \mathbf{p}_t^{-i}) \\
& \text{s.t.} \quad p_t^i = \sum_{k=0}^{K_i} q^{ik} \gamma_t^{ik} \quad \forall i \\
& \quad \sum_{t=1}^T \sum_{k=1}^{K_i} \gamma_t^{ik} \leq L_i \quad \forall i \\
& \quad \sum_{\tau=t}^T \sum_{k=1}^{K_i} \gamma_\tau^{ik} \leq 1 \quad \forall i, t \\
& \quad \sum_{k=0}^{K_i} \gamma_t^{ik} = 1 \quad \forall i, t \\
& \quad \gamma_t^{ik} \in \{0, 1\} \quad \forall i, t, k
\end{aligned} \tag{Multi-POP}$$

Note that all the variables and parameters are extensions of the single item formulation and follow a similar notation as in Section 5.3. In particular,  $L_i$ ,  $S_i$  and  $K_i$  denote the limitation of promotions for item  $i$ , the no-touch period for item  $i$  and the number of prices in the price ladder, respectively. The binary decision variable  $\gamma_t^{ik}$  is equal to

1 if the price of item  $i$  at time  $t$  is selected to be  $q^{ik}$  and 0 otherwise. Note that the objective is to maximize the total profits for the  $N$  items in the category over the time horizon  $T$ . In the formulation (Multi-POP), all the constraints follow from business rules for each item separately (for more details, see Section 2.3.1). Nevertheless, the  $N$  items are linked through the cross-item effects captured in the demand functions. In addition, one can consider various business requirements related to multiple items (called cross business rules), as we discuss next.

1. *Total limited number of promotions.* The supermarket may want to limit the total number of promotions in the category throughout the selling season. This requirement may be motivated by the fact that retailers wish to preserve the image of the store. For example, it may be required to have at most  $L_T = 20$  promotions during the quarter. Mathematically, one can impose the following constraint:

$$\sum_{i=1}^N \sum_{t=1}^T \sum_{k=1}^{K_i} \gamma_t^{ki} \leq L_T. \quad (2.7.1)$$

Note that  $L_T$  should satisfy  $L_T < \sum_{i=1}^N L_i$  for this constraint to be relevant.

2. *Inter-item ordinal constraints.* Several price relations can be dictated by business rules. For example, smaller size items should have a lower price relative to similar products with a larger size, and national brands must be more expensive relative to private labels. This type of constraints can be captured by linear inequalities between the prices (e.g., if item  $i$  should be priced no higher than  $j$ , we have:  $p_t^i \leq p_t^j \forall t$ ).

3. *Simultaneous promotions.* It can be required from the category manager to promote a set of items simultaneously as part of a manufacturer incentive or as part of a special promotional event. If items  $i$  and  $j$  should be promoted at the same time, one can impose the following constraint:  $\gamma_t^{0i} = \gamma_t^{0j} \forall t$ , where  $\gamma_t^{0i}$  is a binary variable that is equal to 1 if item  $i$  is not promoted at time  $t$ .

4. *Limited number of promotions at each period.* One can impose a limitation on the number of promotions at each time period. For example, promoting at most

$C^t = \frac{N}{10}$ , i.e., only at most 10% of the items at each time (note that this limitation can vary with time). Mathematically, we have:

$$\sum_{i=1}^N \sum_{k=1}^{K_i} \gamma_t^{ki} \leq C^t \quad \forall t.$$

5. *Cross no-touch constraints.* An additional requirement can be to space out promotions for a set of similar items by a minimal number of separating periods, denoted by  $S^c$ . Indeed, if successive promotions are too close, this may hurt the store image and incentivize consumers to behave as deal-seekers. In this case, we need to separate successive promotions for two (or more) products. Mathematically, one can impose the following:  $\sum_i \sum_{\tau=t}^{t+S^c} \sum_{k=1}^{K_i} \gamma_\tau^{ki} \leq 1 \quad \forall t$ , where the sum on  $i$  can be on any subset of items chosen by the category manager. Note that when  $S^c = 0$ , this corresponds to never promoting the items simultaneously in order to impose an exclusive offer (very common in practice).

## 2.7.2 Solution Approach

In the optimization problem (Multi-POP), one can observe two types of effects. First, we have cross-time effects (from the stockpiling of each item), where the memory parameter for each product  $M_i$  is estimated from data. Second, we have the presence of cross-item effects (from substitution and complementarity of the different products in the category). Consequently, the problem becomes significantly harder as one may need to solve a very large scale non-linear IP. Our first attempt was to extend the previous linear IP approximation based on unilateral deviations, developed in Section 2.5. For the case of multiple items, it approximates the objective by the sum of unilateral promotions of each item separately. As a result, it fails to capture the cross-item effects and this approach may provide a very poor performance guarantee, when the cross-item effects are significant. For instance, by ignoring the substitution effects, the approximation method can promote all the items whereas the optimal

solution was not to promote at all. In order to develop a better approach to solve problem (Multi-POP), one needs to find a way to incorporate the cross-item effects. We introduce the following sequence of methods, for any given  $k = 1, 2, \dots, N$ .

- $App(1)$  is the IP approximation applied to (Multi-POP) in a similar way as in Section 2.5. In particular, it approximates the objective function by the sum of the marginal contributions of each time and item separately. As we previously discussed, it might provide a poor performance guarantee relative to the optimal POP solution.
- $App(2)$  is an alternative IP approximation applied to (Multi-POP) that includes the marginal contributions (same as  $App(1)$ ) but also the pairwise contributions (i.e., having two items promoted at the same time).  $App(2)$  is described in details below.
- $App(N)$  is an alternative IP approximation that includes the marginal contributions, the pairwise contributions, and so on up to all the possible combinations of having the  $N$  items promoted simultaneously.

One can also naturally consider any intermediate method for  $2 < k < N$ . Note that there exists a trade-off between simplicity (as well as speed) and performance (in terms of accuracy of the approximation relative to the optimal solution). On one extreme,  $App(1)$  is a simple approach that only requires computing the marginal contributions of having one promotion at a time, but can perform poorly as it does not capture the cross-item effects at all. On the other extreme,  $App(N)$  is clearly more accurate, as it successfully captures all the cross-item effects. But this benefit comes at the expense of being more complex as one needs to compute all the combinations of promoting every subset of items simultaneously. In particular, it requires to compute an exponential number of coefficients and to solve an IP that grows exponentially with the number of items. Note that for  $T = 1$  or  $M_i = 0 \forall i$ ,  $App(N)$  is exact as it captures accurately all the cross-item effects. Nevertheless, for a general dynamic

problem with  $T$  periods and non-zero memory parameters,  $App(N)$  is not exact as it still approximates the time effects from the stockpiling. Between  $App(1)$  and  $App(N)$ , any intermediate method  $App(k)$ ;  $k = 2, 3, \dots, N-1$  yields a method with a different trade-off between speed and performance. Next, we describe  $App(2)$  in more details as it will be used in the remaining of this section.

As we previously mentioned,  $App(2)$  approximates the objective of (Multi-POP) by the sum of the unilateral deviations (i.e., having a single promotion at a time) as well as the pairwise contributions (i.e., having two items promoted simultaneously). More precisely, the approximated objective is given by:

$$POP(\mathbf{p}^0) + \max_{\gamma} \left( \sum_{i=1}^N \sum_{t=1}^T \sum_{k=1}^{K_i} b_t^{ki} \gamma_t^{ki} + \sum_{i,j:i>j}^N \sum_{t=1}^T \sum_{k=1}^{K_i} \sum_{\ell=1}^{K_j} b_t^{k\ell ij} \gamma_t^{k\ell ij} \right). \quad (2.7.2)$$

We denote the regular price vector by  $\mathbf{p}^0 = (q^0, \dots, q^0)$ , meaning that the regular price is set to all the items at all times. The first term, denoted by  $POP(\mathbf{p}^0)$ , represents the total profits for all items at all times, without any promotion. The second term captures all the unilateral contributions of having a single promotion, i.e., for one item at one time period. More precisely, we define the price vector  $\mathbf{p}_t^{\mathbf{k}j}$  as follows:

$$(\mathbf{p}_t^{\mathbf{k}j})_{\tau} = \begin{cases} q^k; & \text{if } \tau = t \text{ and } i = j \\ q^0; & \text{otherwise} \end{cases}$$

In other words, the vector  $\mathbf{p}_t^{\mathbf{k}j}$  has the promotion price  $q^k$  for item  $j$  at time  $t$ , and the regular price  $q^0$  (no promotion) is used at all the remaining periods for item  $j$ , and for all the other items at all times. The coefficient  $b_t^{kj}$  is then given by:

$$b_t^{kj} = POP(\mathbf{p}_t^{\mathbf{k}j}) - POP(\mathbf{p}^0), \quad (2.7.3)$$

and represents the marginal contribution on the total profits by having a single promotion for item  $j$  at time  $t$ , using price  $q^k$  (in a similar fashion as in Section 2.5).

The third term in equation (2.7.2) represents all the pairwise contributions of having two items on promotion at the same time. More precisely, we define the price vector  $\mathbf{p}_t^{\mathbf{k}\ell\mathbf{j}u}$  for any pair of items  $j > u$  as follows:

$$(\mathbf{p}_t^{\mathbf{k}\ell\mathbf{j}u})_\tau = \begin{cases} q^k; & \text{if } \tau = t \text{ and } i = j \\ q^\ell; & \text{if } \tau = t \text{ and } i = u \\ q^0; & \text{otherwise} \end{cases}$$

In other words, the vector  $\mathbf{p}_t^{\mathbf{k}\ell\mathbf{j}u}$  has the promotion price  $q^k$  for item  $j$  at time  $t$ , the promotion price  $q^\ell$  for item  $u$  at time  $t$ , and the regular price  $q^0$  is used at all the remaining periods for items  $j$  and  $u$ , and for all the other items at all times. The coefficient  $b_t^{\mathbf{k}\ell\mathbf{j}u}$  is given by:

$$b_t^{\mathbf{k}\ell\mathbf{j}u} = POP(\mathbf{p}_t^{\mathbf{k}\ell\mathbf{j}u}) - POP(\mathbf{p}_t^{\mathbf{k}j}) - POP(\mathbf{p}_t^{\ell u}) + POP(\mathbf{p}^0), \quad (2.7.4)$$

and represents the pairwise contribution on the total profits by having two simultaneous promotions. Note that one can also write:  $b_t^{\mathbf{k}\ell\mathbf{j}u} = POP(\mathbf{p}_t^{\mathbf{k}\ell\mathbf{j}u}) - POP(\mathbf{p}^0) - b_t^{\mathbf{k}j} - b_t^{\ell u}$ . Using this representation,  $b_t^{\mathbf{k}\ell\mathbf{j}u}$  corresponds to the marginal contribution of having two simultaneous promotions (for items  $j$  and  $u$  at time  $t$ ) relative to the case where the promotions are done separately. Finally, one needs to incorporate some consistency constraints for each pair of items  $i$  and  $j < i$  in order to capture the fact that  $\gamma_t^{\mathbf{k}i} = \gamma_t^{\ell j} = 1$  if and only if  $\gamma_t^{\mathbf{k}\ell\mathbf{i}j} = 1$  for each  $t$  and  $k, \ell$ . In order to have a consistent formulation, we should ensure that when both items  $i$  and  $j$  are on promotion, we count the pairwise contribution but also both unilateral deviations. One can encode this set of conditions by incorporating the following constraints to the formulation for each pair of items  $i, j < i$ , each  $t$  and each promotion prices  $k$

and  $\ell$ :

$$\begin{aligned}\gamma_t^{k\ell ij} &\leq \gamma_t^{ki} \\ \gamma_t^{k\ell ij} &\leq \gamma_t^{\ell j} \\ \gamma_t^{k\ell ij} &\geq 0 \\ \gamma_t^{k\ell ij} &\geq \gamma_t^{ki} + \gamma_t^{\ell j} - 1\end{aligned}$$

Note that by construction,  $App(2)$  is exact for  $N = 2$  in the sense that it captures accurately all the cross-item effects. For  $N > 2$ ,  $App(2)$  approximates the true POP objective by the sum of unilateral and pairwise contributions and is not always exact. However, it performs significantly better than  $App(1)$  as it does take into account some of the cross-item effects. Note that one can formalize the definition of  $App(k)$ ;  $k = 3, \dots, N$  in a similar fashion.

When maximizing the objective in (2.7.2), the decisions are the binary variables  $\gamma$ . In particular, there is one such variable for each item/time/price (i.e.,  $NT(K+1)$ ), assuming for simplicity that  $K_i = K \forall i$  and one such variable for any pair of items  $i > j$  at each time/price (i.e.,  $\frac{N(N-1)}{2}TK^2$ ). As we previously mentioned, for  $App(N)$ , this number grows exponentially with  $N$  and  $K$  and therefore, it might be not practical to go beyond  $App(3)$  or  $App(4)$ . Nevertheless, as we show in the next Proposition, we only consider  $App(2)$ .

**Proposition 2.5.** *Assume that the cross effects for each item are additively separable, i.e.,*

$$d_t^i(p_t^i, p_{t-1}^i, \dots, p_{t-M_i}^i, \mathbf{p}_t^{-i}) = h_t^i(p_t^i, p_{t-1}^i, \dots, p_{t-M_i}^i) + \sum_{j \neq i} H_t^{ji}(p_t^j). \quad (2.7.5)$$

- Then, we have:  $App(2) = App(3) = \dots = App(N)$ .
- Consequently, the  $App(2)$  solution coincides with the POP optimal solution under either of the following four conditions: a)  $T = 1$ , b)  $M_i = 0; \forall i$ ,

c)  $L_i = 1; \forall i$ , d)  $S_i \geq M_i; \forall i$ .

The assumption that the cross effects are additively separable allows us to show that  $App(2)$  is equivalent to  $App(N)$ . As a result, it is sufficient to consider only unilateral and pairwise deviations to capture all the cross-item effects accurately. This assumption is satisfied by several demand functions such as the linear cross elasticities model and the semi-log model. In addition, by testing this structure using actual data for several stores and product categories, we observed that it fits well the data and yields a good out of sample prediction accuracy. That being said, one can still consider alternative demand models that do not satisfy this assumption. In this case,  $App(2)$  is not necessarily equal to  $App(N)$ . However, we observed computationally that the performance of  $App(2)$  is very often satisfactory, as it captures most of the cross-item effects. Note that the second part of Proposition 2.5 is similar to Proposition 2.1 from Section 2.5.

We conclude that for additively separable demand functions, one can solve the POP for multiple items by applying  $App(2)$  and capture accurately all the cross-item effects. Next, we address the following two questions.

1. As we previously discussed, in order to solve  $App(2)$ , one needs to solve an IP with a quadratic number of variables with respect to  $N$  and  $K$ . Can we show that the LP relaxation of  $App(2)$  is integral so that one can solve  $App(2)$  efficiently as an LP? Is the feasible region totally unimodular as it is for  $App(1)$ ?
2. Since  $App(2) = App(N)$  (for additively separable demands), the cross effects are captured accurately. However,  $App(2)$  is generally not exact, as there is an approximation error from the time effects (stockpiling). Can we derive a bound on the performance guarantee of the solution obtained from  $App(2)$  relative to the optimal POP solution?

Recall that the  $App(2)$  formulation requires to incorporate four additional consistency constraints for each time and each pair of items and promotion prices. As expected, adding this set of constraints modifies the feasible region of the initial POP

formulation. As we have shown in Theorem 2.2, the single item POP has a tight LP relaxation, as the constraint matrix is totally unimodular. One can show that if we do not have cross-item constraints, this property is preserved for the multiple item POP as the feasible region for each item is totally unimodular. As a result,  $App(1)$  leads to an integral formulation (in the absence of cross constraints). However, by adding the consistency constraints for  $App(2)$ , the feasible region is not totally unimodular anymore. Nevertheless, one can show the following result.

**Proposition 2.6.** *Consider an additively separable demand as in (2.7.5). For substitutable items, the  $App(2)$  formulation is always integral in the absence of business rules.*

The result in Proposition 2.6 admits the following geometrical interpretation. As we discussed, for the  $App(2)$  formulation, the matrix of the feasible region is not totally unimodular. As a result, some of the extreme points can be fractional. By considering an additively separable demand with substitution effects, we ensure that the objective will induce the optimal solution to always lie on integer extreme point. Consequently, one can solve the LP relaxation and obtain an integer solution efficiently.

Note that for most categories of items in a supermarket, the products within a category are either independent (i.e., no cross-item effects) or substitutable. In particular, for categories such as coffee, tea and chocolate, we could not find any complementarity effects in the data we analyzed. Note also that even if some of the products are complement, we observed by extensive testing that  $App(2)$  yields an optimal integer solution very often. In particular, we considered a computational experiment with  $N = 5$  items and 10,000 instances in which the cross elasticity values are randomly generated. We then observed the following: (i) the  $App(2)$  solution was integral 99% of the time (in the absence of cross constraints); and (ii) 91% of the time in the presence of cross-item constraints. In addition, recall that in most applications, we have substitution effects only so that  $App(2)$  is always integral in the absence of business rules. Next, we observed computationally that  $App(2)$  was always integral

in the presence of business rules for each item separately (i.e., any business rule from Section 2.3.1, but no cross-item constraints). Finally, we observed by extensive numerical testing that even in the presence of cross constraints as the ones in Section 2.7.1,  $App(2)$  yields most of the time an integral solution.

Next, we address the second question by presenting the following result on the performance guarantee of the  $App(2)$  solution relative to the optimal POP solution. We consider a multiplicative demand form as in Section 2.6.1. More precisely, the demand of item  $i$  at time  $t$  can be written as:

$$d_t^i = f_t^i(p_t^i) \cdot g_1^i(p_{t-1}^i) \cdot g_2^i(p_{t-2}^i) \cdots g_{M_i}^i(p_{t-M_i}^i) + \sum_{j \neq i} H_t^{ji}(p_t^j), \quad (2.7.6)$$

where the first product of  $M_i + 1$  functions is similar to equation (2.6.3) and the second term represents the additive cross-item effects of the other items  $j \neq i$ . For the demand model in (2.7.6), we develop the following bound.

**Theorem 2.2.** *Let  $\gamma^{POP}$  be an optimal solution to (Multi-POP) and  $\gamma^{App(2)}$  be an optimal solution of the  $App(2)$  approximation. Then, when all the items are substitutable, we have:*

$$1 \leq \frac{POP(\gamma^{POP})}{POP(\gamma^{App(2)})} \leq \frac{1}{\underline{R}_M}, \quad (2.7.7)$$

where:  $\underline{R}_M = \min_{j=1,2,\dots,N} \prod_{i=1}^{\tilde{L}_i-1} g_{i(S_i+1)}^j(q^{K_j})$ , with  $\underline{R}_M = 1$  by convention, if  $\tilde{L}_i = 1$  for all  $i = 1, 2, \dots, N$ .

$\tilde{L}_i$  is defined exactly as  $\tilde{L}$  in Section 2.6. Note that the good performance of the bound from the single item setting is preserved. Indeed, the bound in (2.7.7) is characterized by the worst case of  $\underline{R}_i$  for each item  $i = 1, 2, \dots, N$ . As we discussed in Section 2.6.1 for the single item POP, this bound yields a good performance guarantee for a wide range of practical settings. Consequently, it also performs well for the multiple item setting. In addition, recall that the bound in (2.7.7) is only a theoretical

performance guarantee and in practice, the ratio  $\frac{POP(\gamma^{POP})}{POP(\gamma^{App(2)})}$  is much closer to 1. One can also show that the bounds in Theorem 2.2 are tight in a similar fashion as in Proposition 2.4. Finally, one can also derive similar bounds for the additive demand model.

### 2.7.3 Insights

In this section, we present several insights drawn from our promotion optimization problem. First, we show that our model is able to capture the loss leader effect that is very often observed in practice. Second, we consider a simple symmetric setting to study the effect of substitution and complementarity on the optimal promotion schedule.

#### Loss Leader

The loss leader is a common phenomenon in which one product is priced below cost in order to extract significant profits on complementary items. Examples include a printer and cartridges as well as a video game console and games. We next consider a small stylized example to illustrate that our promotion optimization model can capture the loss leader effect. In particular, we consider an example of complementary products and show that the optimal promotion profile will induce a loss leader behavior. We consider a small example with  $N = 4$  items, where item 1 is the leader (e.g., the printer) and items 2, 3 and 4 are the complements (e.g., cartridges). We then consider a linear demand model  $d(\mathbf{p}) = \bar{d} - \beta M \mathbf{p}$  with a price sensitivity matrix given by:

$$M = \begin{bmatrix} 1 & r/2 & r/2 & r/2 \\ r/2 & 1 & 0 & 0 \\ r/2 & 0 & 1 & 0 \\ r/2 & 0 & 0 & 1 \end{bmatrix},$$

where  $r$  represents the degree of complementarity between the leader and the other items. In addition, in this example we assume that the other items are independent. We then solve the unconstrained (Multi-POP) for  $T = 1$  and plot the optimal solution in Figure 2-4. One can see that when  $r = 0$  (i.e., independent items), it is optimal not

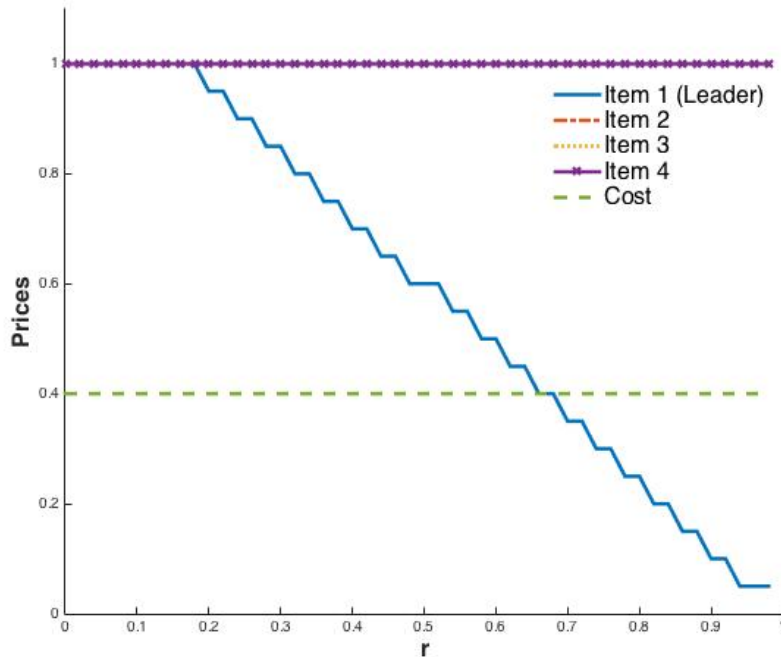


Figure 2-4: Illustration of the loss leader effect

to promote any item (the optimal prices are all equal to the normalized regular price). When the degree of complementarity  $r$  increases, it becomes optimal to promote item 1 (the leader) but never the other complementary items. In this example, the normalized cost of all the items is set to  $c = 0.4$ . One can see in Figure 2-4 that when  $r$  is large enough (in this case,  $r > 0.66$ ), it becomes optimal to set the price of the leader below cost. Consequently, our model captures the fact that it may be optimal to sell an item below cost in order to extract positive profits on complementary items.

## Cross-item Effects

We next present several insights on the impact of cross-item effects on promotion planning. In particular, we consider solving the (Multi-POP) in order to study the impact of cross-item effects on the optimal solution. For simplicity, we consider a setting with  $N$  identical items (i.e., same demand function, cost and price ladder) and no constraints in the optimization formulation. In addition, we assume that there is no stockpiling (i.e., all the memory parameters are zero,  $M_i = 0 \forall i$ ). These simplifying assumptions allow us to isolate and study the impact of the cross-item effects (complementarity versus substitution). The results are summarized in Table 2-5.

Scenario	Mild cross item effects	Strong cross item effects
<b>N Substitutes</b>	Promote all items at all times	No promotions
<b>N Complements</b>	Promote <b>only a subset of items</b> , constant number of promotions at each time	Promote all items at all times

Figure 2-5: Summary of POP insights regarding cross-item effects

One can see from Table 2-5 that for  $N$  substitutable items, it is either optimal to promote all the items at all times, or not to promote at all. In particular, it depends on the magnitude of the substitution effect. When substitution is strong, promoting one item yields larger sales for this item but at the same time, decreases the sales of the other products. As a result, one is better off by not promoting at all. When the substitution effects are mild, the benefit of promoting one product overcomes the loss from the cannibalization on the other items. Consequently, it becomes preferable to promote all the items.

For  $N$  items that are complements (and identical), the conclusions are opposite. When cross effects are strong, promoting one item yields larger sales for this item and at the same time, increases the sales of the other products. As a result, one is better off by promoting all the items at all times. This result is surprising as one

would think that when two products are complement, it will be enough to promote only one of them. However, recall that we assume a symmetric setting where any pair of items  $i$  and  $j \neq i$  are complement. Then, promoting item  $i$  enhances the sales of both products, whereas promoting item  $j$  has the same effect. In other words, by promoting both items, some additional buyers will be incentivized to purchase the complementary product. When cross-item effects are mild, promoting all the products is not optimal anymore. In this case, the benefit induced by the complementarity is not sufficiently large relative to the price decrease. Therefore, it becomes optimal to promote only a subset of items (the size of the subset depends on the magnitude of the complementarity effect).

In practice, supermarkets solve the POP for a very large scale scenario that involves asymmetries, seasonality, stockpiling, substitutes and complements as well as several business rules. It is not easy to optimally design the promotion planning as too many tradeoffs are interacting. This provides the need for an optimization tool, such as the one we develop in this work. Our tool can take into account all the different tradeoffs and compute an optimal solution for the promotion planning problem. In addition, since our methods are solved efficiently, one can perform a sensitivity analysis in order to see how the promotion planning is affected by a modification in the demand parameters or in some of the business rules. This allows category managers to reach a better understanding of several effects (such as cross-item and stockpiling effects) and how they affect promotion planning.

## 2.8 Computational Results

In order to quantify the value of our promotion optimization model, we perform an end-to-end experiment where we start with data from an actual retailer (supermarket), estimate the demand model we introduce, validate it, compute the optimized prices from our LP model and finally compare them with actual prices implemented by the retailer. In this section, following the recommendation of our industry col-

laborators, we perform detailed computational experiments for the log-log demand, which is a special case of the multiplicative model (2.6.3) and often used in practice.

### 2.8.1 Estimation Method

We obtained customer transaction data from a grocery retailer. The structure of the raw data is the customer loyalty card ID (if applicable), a timestamp, and the purchased items during that trip. In this chapter, we focus on the coffee category at a particular store. For the purposes of demand estimation, we first aggregated the sales at the brand-week level. It seems natural to aggregate sales data at the week level as we observe that typically, a promotion starts on a Monday and ends on the following Sunday. Our data consists of 117 weeks from 2009 to 2011. For ease of interpretation and to keep the prices confidential, we normalize the regular price of each product to 1.

To predict demand as a function of prices, we estimate a log-log (power function) demand model incorporating seasonality and trend effects (similarly as in (2.6.1)):

$$\log d_{it} = \beta^0 BRAND_i + \beta^1 t + \beta^2 WEEK_t + \sum_{m=0}^M \beta_{im}^3 \log p_{i,t-m} + \epsilon_t, \quad (2.8.1)$$

where  $i$  and  $t$  denote the brand and time indices,  $d_{it}$  denotes the sales (which we assume is equal to the demand, as we discussed in Section 2.3.2) of brand  $i$  in week  $t$ ,  $BRAND_i$  and  $WEEK_t$  denote brand and week indicators,  $p_{it}$  denotes the average per-unit selling price of brand  $i$  in week  $t$ .  $\beta_0$  and  $\beta_2$  are vectors with components for each brand and each week respectively, whereas  $\beta_1$  is a scalar that captures the trend. Note that the seasonality parameters  $\beta_2$  for each week of the year are jointly estimated across all the brands in the category. The additive noises  $\epsilon_t$ ;  $\forall t = 1, \dots, T$  account for the unobserved discrepancies and are assumed to be normally distributed and i.i.d. Similar demand models have been used in the literature, e.g., [16] and [19].

The model in (2.8.1) is a multiplicative model, which assumes that the brands share a common multiplicative seasonality; but each brand depends only on its own

current and past prices; and the independent variables are assumed to have multiplicative effects on demand. In particular, the model incorporates a trend effect  $\beta^1$ , weekly seasonality  $\beta^2$ , and price effects  $\beta^3$ . When the memory parameter  $M = 0$ , then only the current price affects the demand in week  $t$ . When the memory parameter  $M = 2$ , then the demand in week  $t$  depends not only on the price in the current week  $p_t$  but also in the price of the two previous weeks  $p_{t-1}$  and  $p_{t-2}$ . We note that our model does not account explicitly for cross-brand effects, i.e., we assume that the demand for brand  $i$  depends only on the prices of brand  $i$ . This assumption is reasonable for certain products such as coffee because people are loyal about the brand they consume and do not easily switch between brands. In addition, the high predictive accuracy of our model validates this assumption.

For ease of notation, from this point, we drop the brand index  $i$  since we estimate and optimize for a single item model. Observe that one can define:

$$f_t(p_t) = \exp(\beta^0 + \beta^1 t + \beta^2 WEEK_t + \beta_0^3 \log p_t),$$

$$g_m(p_{t-m}) = (p_{t-m})^{\beta_m^3}, \quad m = 1, \dots, M,$$

and therefore, equation (2.8.1) is in fact a special case of the multiplicative demand model in (2.6.3).

Based on our intuition, one expects to find the following from the estimation:

1. Since demand decreases as the current price increases, we would expect that the self-elasticity parameter is negative, i.e.,  $\beta_0^3 < 0$ .
2. Since a deeper past promotion leads to a greater reduction in current demand, we would expect that the past elasticity parameters are positive, i.e.,  $\beta_m^3 \geq 0$  for  $m > 0$ .
3. Holding the depth of promotion constant, a more recent promotion leads to a greater reduction in current demand than the same promotion earlier in time. Therefore, we would expect that the past-elasticity parameters are decreasing

in time, i.e.,  $\beta_m^3 > \beta_{m+1}^3$ , for  $m = 1, \dots, M - 1$ .

We note that the conditions above are a special case of Assumption 1 for the log-log demand.

We divide the data into a training set, which comprises the first 82 weeks and a test set which comprises the second 35 weeks. We use the training set to estimate the demand model and then predict the out-of-sample sales to test our predictions. In order to measure forecast accuracy, we use the following forecast metrics. In the sequel, we use the notation  $s_t$  for the actual sales (or equivalently demand) and  $\hat{s}_t$  be the forecasted values.

- The *mean absolute percentage error* (MAPE) is given by:

$$\text{MAPE} = \frac{1}{T} \sum_{t=1}^T \frac{|s_t - \hat{s}_t|}{s_t}.$$

The MAPE captures the average relative forecast error in absolute value. If the forecast is perfect, then the MAPE is equal to zero.

- The  $R^2$  is given by:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}},$$

where  $\bar{s} = \sum_{t=1}^T s_t / T$ ,  $SS_{tot} = \sum_{t=1}^T (s_t - \bar{s})^2$  and  $SS_{res} = \sum_{t=1}^T (s_t - \hat{s}_t)^2$ . We distinguish between in-sample (IS) and out-of-sample (OOS)  $R^2$ . If the forecast is perfect, then  $R^2 = 1$ . In addition, one can consider the adjusted  $R^2$  as it is common in demand estimation. The latter adjusts the regular  $R^2$  to account for the number of explanatory variables in the model relative to the number of data points available and is given by:

$$R_{adj}^2 = 1 - (1 - R^2) \cdot \frac{n - 1}{n - p - 1},$$

where  $p$  is the total number of independent variables in the model (not counting the constant term), and  $n$  is the sample size.

	Variable	Coefficient	Std Error	<i>p</i> -value
Brand1	$\log p_t$	-3.277	0.231	2e-16***
	$\log p_{t-1}$	0.518	0.229	0.024*
	$\log p_{t-2}$	0.465	0.231	0.045*
Brand2	$\log p_t$	-4.434	0.427	2e-16***
	$\log p_{t-1}$	1.078	0.423	0.011*
	$\log p_{t-2}$	0.067	0.413	0.870
Brand1	MAPE		0.116	
	OOS $R^2$		0.900	
	Revenue bias		1.059	
Brand2	MAPE		0.097	
	OOS $R^2$		0.903	
	Revenue bias		1.017	

Table 2.2: A subset of the estimation results for two coffee brands

- The *revenue bias* is measured as the ratio of the forecasted to actual revenue, and is given by:

$$\text{revenue bias} = \frac{\sum_{t=1}^T p_t \hat{s}_t}{\sum_{t=1}^T p_t s_t}.$$

## 2.8.2 Estimation Results and Discussion

### Coffee Category

The coffee category is an appropriate candidate to test our model as it is common in promotion applications (see e.g., [15] and [27]). We use a linear regression to estimate the parameters of the demand model in equation (2.8.1) for five different brands of coffee. For conciseness, we only present a subset of the estimation results for two coffee brands in Table 2.2. We compare the actual and predicted sales in Figure 2-6. Remember that our data consists of 117 weeks which we split into 82 weeks on training and 35 weeks of testing.

On one hand, brand 1 is a private-label brand of coffee which has frequent promotions (approximately once every 4 weeks). The price-elasticity coefficients for the

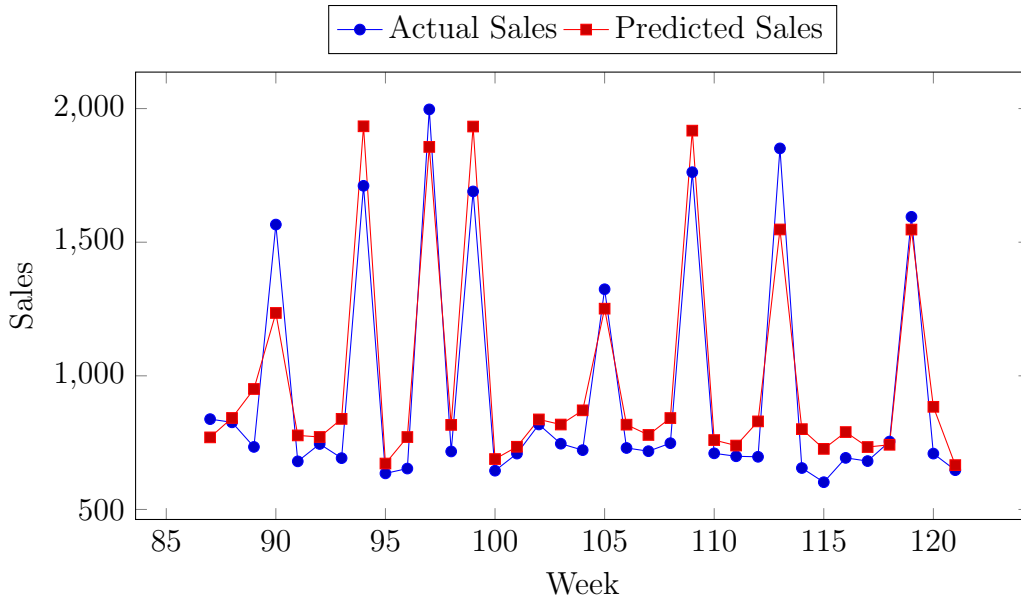


Figure 2-6: Actual versus forecasted sales over the 35 test weeks for Brand1

current price and two previous prices are statistically significant suggesting that for this brand, the memory parameter  $M = 2$ .

On the other hand, brand 2 is a premium brand of coffee which has also frequent promotions (approximately once every 5 weeks). The price-elasticity coefficients for both the current price and the price in the prior week are statistically significant, but the coefficient for the price two weeks ago is not. This suggests that for this brand, the memory parameter  $M = 1$ .

By observing the statistically significant price coefficients, one can observe that they agree with the expected findings mentioned previously. Furthermore, given the high accuracy as measured by low MAPEs, we expect that cross-brand effects are minimal.

#### Four Categories

In the same spirit, we estimate the log-log demand model for several brands for the chocolate, tea and yogurt categories. The results are summarized in Table 2.3. We do not report the individual product coefficients but note that they follow our

expectations in terms of sign and ordering. We wish to highlight that the forecast error is low as evidenced by the high in-sample and out-of-sample  $R^2$ , and the low MAPE values and revenue bias being close to 1.

Category	IS Adj $R^2$	MAPE	OOS $R^2$	Revenue Bias	Product Memory
Coffee	0.974	0.115	0.963	1.000	0, 1, 2
Chocolate	0.951	0.185	0.872	0.990	0, 1, 2
Tea	0.984	0.187	0.759	1.006	0, 1
Yogurt	0.983	0.115	0.964	1.073	0, 1

Table 2.3: Summary of the estimation results for four categories

We next observe the following regarding the effect of the memory parameter.

1. The memory parameter differs across products within a category. In general, basic products have higher memory ( $M = 1$  or  $2$ ) whereas premium items have lower memory ( $M = 0$ ).
2. The memory parameters are estimated from data and differ depending on the category. Products in the yogurt and tea categories have memory of zero or one; whereas products in the coffee and chocolate categories have memory of zero, one or two. This agrees with our intuition that for perishable items (such as yogurt), consumers do not stock-pile and therefore, the memory parameter is zero. However, coffee is clearly a less perishable category so that stock-piling is more significant.

### 2.8.3 Optimization Results and Discussion

Having validated the forecasting demand model, we next perform a computational experiment to compute and test the optimized promotion prices. We assume that the demand forecast is the true demand model and use it as an input into our promotion optimization model from (POP).

**Experimental setup:** We compute the LP optimized prices for a single item (Brand1) over the horizon of the test weeks, which is  $T = 35$  weeks. During the planning horizon, the retailer used  $L = 8$  promotions with at most  $S = 1$  separating weeks (i.e., consecutive promotions are separated by at least 1 week). As stated earlier, the regular price is normalized to be one unit. Due to confidentiality, we do not reveal the exact costs of the product, i.e., the parameters  $c_t$  in (POP). For the purpose of this experiment, we assume the cost of the product to be constant,  $c_t = 0.4$ . Since the lowest price charged by the retailer was 0.75, the set of permissible normalized prices is chosen to be  $\{0.75, 0.80, 0.85, \dots, 1\}$ .

The LP optimization results are shown in Figure 2-7. Before discussing the results, we first want to make the following observations:

- The predicted profit using the actual prices implemented by the retailer (and not chosen optimally) together with the forecast model is \$18,425. All the results will be compared relative to this benchmark value.
- The predicted profit using only the regular price (i.e., no promotions) is \$17,890. This is a 2.9% loss relative to the benchmark. Therefore, the estimated log-log model predicts that the actual prices yield a 2.9% gain relative to the case without promotions, even if the actual promotions are not chosen optimally.
- The predicted profit using the optimized LP prices imposing the same number of promotions as a business requirement ( $L = 8$  during a period of 35 weeks) is \$19,055. This is a 3.4% gain relative to the benchmark. Therefore, the estimated log-log model predicts that the optimized LP prices with the same number of promotions yield a 3.4% gain relative to the actual implemented profit. In other words, by only carefully planning the same number promotions, our model (and tool) suggests that the retailer can increase its profit by 3.4% in this case.
- The predicted profit using the optimized LP prices and allowing three additional promotions ( $L = 11$ ) is \$19,362. This is a 5.1% gain relative to the benchmark.

Therefore, the estimated log-log model predicts that optimized prices with three additional promotions yield a 5.1% gain relative to the actual profit. Therefore, the retailer can easily test the impact of allowing additional promotions within the horizon of  $T = 35$  weeks.

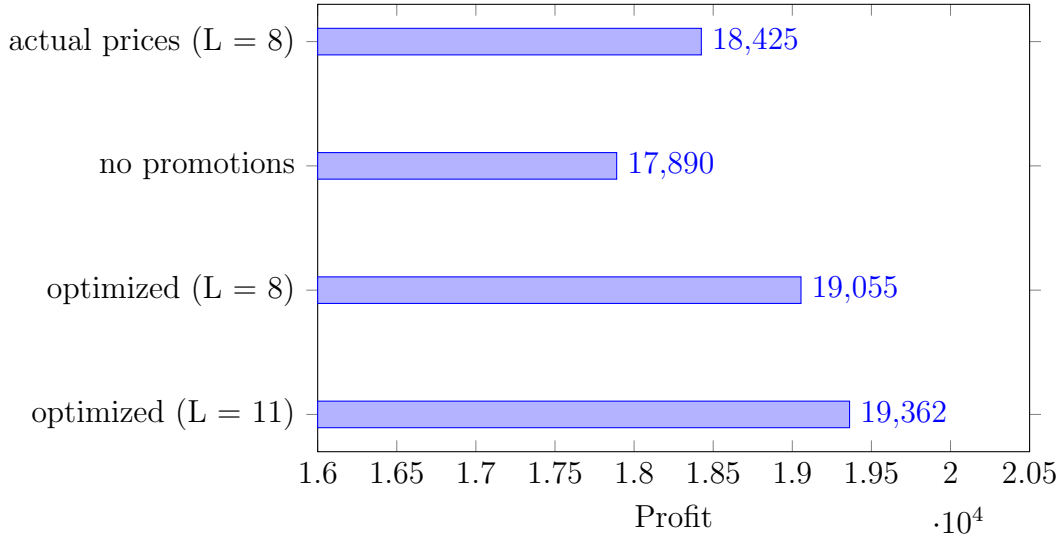


Figure 2-7: Profits for different scenarios using a log-log demand model

We next compute the bound from Theorem 2.1 for the actual data we have been using in our computations above. The lower bound can be rewritten as  $\underline{R} \cdot POP(\gamma^{POP}) \leq POP(\gamma^{LP})$ , where  $\underline{R} = \prod_{i=1}^{\tilde{L}-1} g_{i(S+1)}(q^K)$  and therefore, depends on the parameters of the problem. We compute  $\underline{R}$  for both coffee brands from Table 2.2. We have  $q^K = 0.75$ ,  $L = 8$  and test the bound,  $\underline{R}$ , for various values of  $S$ . When  $S \geq 2$ , we observe that  $\underline{R} = 1$  and therefore the method is optimal for both brands. For  $S = 1$ , we obtain that for Brand1,  $\underline{R} = 0.8748$ , whereas for Brand2,  $\underline{R} = 1$ . Finally, we consider  $S = 0$  as it is the worst case scenario. In other words, no requirement on separating two successive promotions is imposed (not very realistic). We have for Brand1 and Brand2,  $\underline{R} = 0.7538$  and  $\underline{R} = 0.733$  respectively. We note that the above bounds outperform the approximation guarantees from the literature on submodular maximization. In particular, the problem of maximizing an arbitrary non-monotone submodular function subject to no constraints admits a  $1/2$  approximation algorithm

(see for example, [5] and [11]). In addition, the problem of maximizing a monotone submodular function subject to a cardinality constraint admits a  $1 - 1/e$  approximation algorithm (e.g., [21]). However, our bounds are not constant guarantees for every instance of the POP with multiplicative demand, as it depends on the values of the parameters. Recall also that in practice the LP approximation usually performs better than the bounds.

Next, we compare the running time of the LP to a naïve approach of using an exhaustive search method in order to find the optimal prices of the POP. Note that the POP problem is neither convex nor concave. The results are shown in Figure 2-8. The experiments were run using a desktop computer with an Intel Core i5 680 @ 3.60GHz CPU with 4 GB RAM. The number of the price ladder elements is  $|\mathcal{Q}|$ . The LP was solved using the Java interface to Gurobi 5.5.0. The LP formulation requires 0.01–0.05 seconds to solve, regardless of the value of the promotion limit  $L$ . However, the exhaustive search running time grows exponentially in  $L$ . In addition, for a simple instance of the problem with only 2 prices in the price ladder, it requires one minute to solve when  $L = 8$ . The running time of the exhaustive search method also grows exponentially in the number of elements of the price ladder. For example, with 3 elements in the price ladder and  $L = 8$ , it requires 3 hours to solve, whereas the LP solution solves within milliseconds. We note that since we are considering non-linear demand functions with integer variables, general methods to solve this problem do not exist in commercial solvers and hence are not practical.

The above results show that the exhaustive search method is clearly not a viable option in practice. Note that the LP formulation solves very fast. An important feature of our method relies on the fact that in practice, one can implement it on a platform such as Excel. For a category manager in charge of around 300 SKUs, solving the POP for each item independently would require only about 15 seconds. An additional advantage of short running times is that it allows category managers to perform a sensitivity analysis with respect to the business requirements and to the model parameters. For example, if the optimization is embedded into a decision-

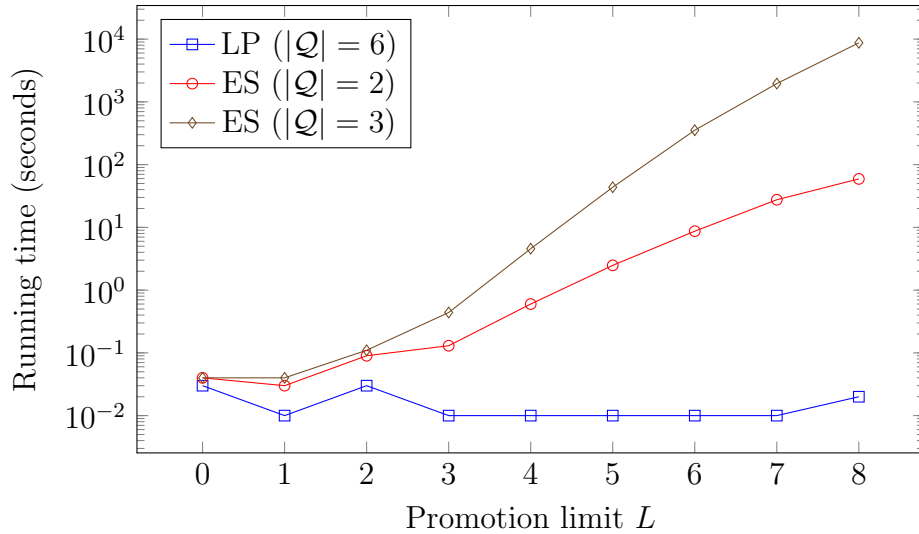


Figure 2-8: Running times of the LP formulation and the exhaustive search method for the POP

support tool, category managers could perform interactive “what-if” analysis. In practice, this would not be possible in the case where the optimization running times exceed a few minutes. In addition, as we have shown in this chapter, the LP formulation yields a solution that is accurate relative to the POP optimal prices and one can compute the upper and lower bounds as a guarantee.

## 2.9 Conclusions

In many important settings, promotions are a key instrument for driving sales and profits. We introduce and study an optimization formulation for the POP that captures several important business requirements as constraints (such as separating periods and promotion limits). We propose two general classes of demand functions depending on whether past prices have a multiplicative or an additive effect on current demand. These functions capture the promotion fatigue effect emerging from the stock-piling behavior of consumers and can be easily estimated from data. We show that for multiplicative demand, promotions have a supermodular effect (for some sub-

sets of promotions) which leads to the LP approximation being an upper bound on the POP objective; whereas for additive demand, promotions have a submodular effect which leads to the LP approximation being a lower bound on the POP objective. The objective is nonlinear (neither convex nor concave) and the feasible region has linear constraints with integer variables. Since the exact formulation is “hard”, we propose a linear approximation that allows us to solve the problem efficiently as an LP by showing the integrality of the IP formulation. We develop analytical results on the LP approximation accuracy relative to the optimal (but intractable) POP solution and characterize the bounds as a function of the problem parameters. We also show computationally that the formulation solves fast using actual data from a grocery retailer and that the accuracy is high.

Together with our industry collaborators from Oracle Retail, our framework allows us to develop a tool which can help supermarket managers to better understand promotions. We test our model and solution using actual sales data obtained from a supermarket retailer. For four different product categories, we estimate from transactions data the log-log demand model. Our estimation results provide a good fit and explain well the data but also reveal interesting insights. For example, non-perishable products exhibit longer memory in the sense that the sales are affected not only by the current price but also by the past prices. This observation validates the hypothesis that demand exhibits a promotion fatigue effect for certain items. We test our approach for solving the promotion optimization problem, by first estimating the demand model from data. We then solve the POP by using our LP approximation method. In this case, using the LP optimized prices would lead about 3% profit gain for the retailer, with even 5% profit gain by slightly modifying the number of promotions allowed. In addition, the running time of our LP is short ( $\sim 0.05$  seconds) making the method attractive and efficient. The naïve optimal exhaustive search method is several orders of magnitude slower. The fast running time allows the LP formulation to be used interactively by a category manager who may manage around 300 SKUs in a category. In addition, one can conveniently run a large number

of instances allowing to perform a comprehensive sensitivity analysis translated into “what-if” scenarios. We are currently in the process of conducting a pilot experiment with an actual retailer, where we test our model in a real-world setting by optimizing promotions for several items and stores.

## Bibliography

- [1] Kusum L Ailawadi, Karen Gedenk, Christian Lutzky, and Scott A Neslin. Decomposition of the sales impact of promotion-induced stockpiling. *Journal of Marketing Research*, pages 450–467, 2007.
- [2] Dimitris Bertsimas and Romy Shioda. Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43(1):1–22, 2009.
- [3] Daniel Bienstock. Computational study of a family of mixed-integer quadratic programming problems. *Mathematical programming*, 74(2):121–140, 1996.
- [4] R. C. Blattberg and S. A. Neslin. *Sales promotion: Concepts, Methods, and Strategies*. Prentice Hall, 1990.
- [5] Niv Buchbinder, Moran Feldman, Joseph Seffi Naor, and Roy Schwartz. A tight linear time  $(1/2)$ -approximation for unconstrained submodular maximization. In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*, pages 649–658. IEEE, 2012.
- [6] Gérard P Cachon and Martin A Lariviere. Supply chain coordination with revenue-sharing contracts: strengths and limitations. *Management science*, 51(1):30–44, 2005.
- [7] Maxime C Cohen, Jeremy Kalas, Kiran Panchamgam, and Georgia Perakis. Optimizing promotions for multiple items in supermarkets. *Working paper*, 2015.

- [8] Maxime C Cohen, Ngai-Hang Z Leung, Kiran Panchamgam, Georgia Perakis, and Anthony Smith. The impact of linear optimization on promotion planning. *Available at SSRN 2382251*, 2015.
- [9] L. G. Cooper, P. Baron, W. Levy, M. Swisher, and P. Gogos. Promocast: A new forecasting method for promotion planning. *Marketing Science*, 18(3):301–316, 1999.
- [10] Daniel Corsten and Thomas Gruen. Stock-outs cause walkouts. *Harvard Business Review*, 82(5):26–28, 2004.
- [11] Uriel Feige, Vahab S Mirrokni, and Jan Vondrak. Maximizing non-monotone submodular functions. *SIAM Journal on Computing*, 40(4):1133–1153, 2011.
- [12] G. Fibich, A. Gavious, and O. Lowengart. Explicit solutions of optimization models and differential games with nonsmooth (asymmetric) reference-price effects. *Operations Research*, 51(5):721–734, 2003.
- [13] Eijte W Foekens, Peter SH Leeflang, and Dick R Wittink. Varying parameter models to accommodate dynamic promotion effects. *Journal of Econometrics*, 89(1):249–268, 1998.
- [14] Ignacio E Grossmann. Review of nonlinear mixed-integer and disjunctive programming techniques. *Optimization and Engineering*, 3(3):227–252, 2002.
- [15] Sunil Gupta. Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing Research*, pages 342–355, 1988.
- [16] H. J. Van Heerde, P. S. H. Leeflang, and D. R. Wittink. The estimation of pre- and post- promotion dips with store-level scanner data. *Journal of Marketing Research*, 37:383–395, 2000.
- [17] Raymond Hemmecke, Matthias Köppe, Jon Lee, and Robert Weismantel. Non-linear integer programming. In *50 Years of Integer Programming 1958-2008*, pages 561–618. Springer, 2010.

- [18] P. K. Kopalle, A. G. Rao, and J. L. Assuncao. Asymmetric reference price effects and dynamic pricing policies. *Marketing Science*, 15(1):60–85, 1996.
- [19] S. Macé and S. A. Neslin. The determinants of pre- and postpromotion dips in sales of frequently purchased goods. *Journal of Marketing Research*, 41:339–350, 2004.
- [20] S Thomas McCormick. Submodular function minimization. *Handbooks in operations research and management science*, 12:321–391, 2005.
- [21] George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. An analysis of approximations for maximizing submodular set functions. *Mathematical Programming*, 14(1):265–294, 1978.
- [22] I. Popescu and Y. Wu. Dynamic pricing strategies with reference effects. *Operations Research*, 55(1):98–112, 2007.
- [23] Alexander Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *Journal of Combinatorial Theory, Series B*, 80(2):346–355, 2000.
- [24] Hanif D Sherali and Warren P Adams. *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*, volume 31. Springer, 1998.
- [25] K. T. Talluri and G. J. van Ryzin. *The Theory and Practice of Revenue Management*. Springer, 2005.
- [26] K Van Donselaar, Tom Van Woensel, RACM Broekmeulen, and J Fransoo. Inventory control of perishables in supermarkets. *International Journal of Production Economics*, 104(2):462–472, 2006.
- [27] J Miguel Villas-Boas. Models of competitive price promotions: Some empirical evidence from the coffee and saltine crackers markets. *Journal of Economics & Management Strategy*, 4(1):85–107, 1995.

# Chapter 3

## Designing Price Incentives in a Network with Social Interactions

### 3.1 Introduction

The recent ubiquity of social networks has revolutionized the way people interact and influence each other. The overwhelming success of social networking platforms such as Facebook and Twitter allows firms to collect unprecedented volumes of data about their customers, their buying behavior including their social interactions with other customers. The challenge that confronts every firm, from big to small, is how to process this data and turn it into actionable policies so as to improve their competitive advantage. In this chapter, we focus on the design of effective pricing strategies to improve profitability of a monopolistic firm that sells indivisible goods or services to agents embedded in a social network.

*Word-of-mouth* communication between agents has always been an effective marketing tool for several businesses. In recent times, word of mouth communication is just as likely to arise from social networks or smart phone applications as from a neighbor across the fence. Consultants at The Conversation Group report that 65% of consumers who receive a recommendation from a contact on their social media sites have purchased a product that was recommended to them. In particular, personal-

ized referrals from friends and family have been more effective in encouraging such purchases. Finally, nearly 93% of social media users have either made or received a recommendation for a product or service. Academic research on consumer behavior shows that consumers' purchasing decisions and product evaluations are influenced by their reference groups (see e.g. [17]). The previous examples clearly indicate that people influence their connections. They not only guide their purchasing behavior but more importantly alter their willingness-to-pay for various items. For example, when an individual buys a product and posts a positive review on his Facebook page, he not only influences his peers to purchase the same item or service but also increases their valuation for the item. The increase in the valuation can sometimes be non-linear in that if many other people have already recommended the item to the peers then the increase in the peer's valuation can be relatively small (or big).

An important feature of the products or services we consider in this chapter is that they exhibit *local non-linear positive externalities*. This means people *positively* influence each other's willingness-to-pay for an item. In addition, the item becomes more valuable to a person if many of his friends buy it (e.g., monotonic) even though there can be a decreasing (or increasing) marginal effect (e.g., submodular or supermodular). Examples of products with such effects include smart phones, tablets, certain fashion items and cell phone plan subscriptions. Such positive externalities may be even more significant when a new generation of products is introduced in the market and people use social networks as a way to accelerate their friends' awareness about the item.

It is common practice that a very small number of highly influential people (e.g., certified bloggers on certain websites) receive the item nearly for free to increase the awareness of the remaining population. Mark W. Schaefer, the author of: "Return on Influence" reports that: "For the first time, companies large and small can find these passionate influencers (using social networks), connect to them and turn them into brand advocates". Therefore, it can be very valuable for firms to identify these influential agents. As an example, many online sellers let consumers sign in with their

Facebook account. Consequently, they have access to their personal information such as age, gender, geographical location, number of friends and more importantly, their network. Various sellers even build a Facebook page to advertise their firm through social platforms. For example, the large US corporation Macy's has more than 14.6M of fans that liked their Facebook page (December 2014). These fans can claim offers using the social platform that allows for a certain degree of publicity and thus directly influence their friends about purchasing. This interaction between the seller and the fan club not only allows the seller to keep their fan club engaged and interested in the brand, but also enables the seller to identify the influencers, provide them with personalized prices or incentives and turn them into influencers, thereby increasing overall profitability.

We consider a setting where a monopolist sells an indivisible good to consumers embedded in a social network. Our goal is to develop a model that incorporates local non-linear interactions as well as positive social externalities between potential buyers and design efficient algorithms to compute the optimal prices that maximize the seller's profitability. We formulate the optimal pricing problem as a two-stage game between seller and the agents (buyers in the network) where the seller first offers prices and the agents then choose whether to purchase the item or not at the offered prices. The main contributions of the chapter are as follows:

1. **Non-linear additive influence model.** We introduce and study a non-linear additive model of influence that captures an agent's valuation for an item. In particular, the total value the agent derives when purchasing the item is the sum of the own valuation and the valuation derived from the influences of all the subsets of his friends who own the item. This model captures a broad class of linear and non-linear influence functions and can model submodular and supermodular influence structures with respect to the number of influencing agents.
2. **Existence of a pure strategy Nash equilibrium that is strictly pre-**

**ferred by all buyers in the network.** We show the existence of a pure strategy Nash equilibrium for the second stage game given any vector of prices. We observe that there can exist multiple equilibria but one of them is preferred by the seller. We also show that using a small price perturbation, the seller can ensure that all agents strictly prefer the same equilibrium.

3. **Reformulate into a single stage operational MIP.** Using duality theory, we derive equilibrium constraints and reformulate the two-stage problem faced by the seller into a non-convex integer program. We then transform it into an equivalent mixed-integer program (MIP) using reformulation techniques from integer programming. This resulting MIP can be viewed as an operational pricing tool as any firm can easily incorporate business rules on prices and constraints on network segmentation. For example, the seller can identify bounded and tiered prices for the members in its fan club based on their loyalty class such as platinum or gold.
4. **Efficient algorithms for discriminative and uniform pricing strategies.** We develop efficient and scalable methods to optimally solve the MIP for two distinct pricing strategies using the linear programming (LP) relaxation of the MIP. We consider the discriminative and the uniform pricing strategies and present a solution method that is efficient (polynomial in the number of agents) and scalable to large networks.
5. **Insights about the structure of the optimal solution in the discriminative pricing strategy.** We show that the price of an agent that buys in the optimal discriminative pricing solution is the sum of its own value and a markup term that corresponds to the influence by the network of agents that buy. The seller offers this price to each agent depending on its own valuation for the item and the influence the agent exerts on the network. The seller's profit from network externalities in essence comes from two types of customers: high valued customers who influence their neighbors and low valued customers

who are highly influential. In addition, when comparing submodular, linear and supermodular influence models, we show that as we move from the submodular to the linear and then to the supermodular influence models, additional agents will buy. In addition, the buyers will pay a higher price so that it induces a larger profit for the seller.

6. **Uncertain or partially known influences.** We study a stylized model where the interaction terms are uncertain and modeled as random variables. We draw interesting insights and study the impact of several parameters on the optimal price. In particular, we show that for large networks with a large number of edges, the optimal price converges to the solution of the deterministic problem with the random interaction terms replaced by their mean.
7. **Model with incentives that guarantee influence.** In the above model, we consider an online setting and assume that anyone who buys the item influences their peers. This assumption is not realistic in many practical settings. Indeed, after purchasing an item, it is sometimes not entirely natural to influence friends about the product unless one takes some effort to do so. This for example could be writing a review, endorsing the item on their (Facebook) wall or at the very least announcing that they have purchased the item. In practice, it is popular for firms to offer cash rewards for recommending or referring friends. For example, Groupon, a popular deal-of-the-day website that features discounted gift certificates offers a \$10 reward to users who refer a friend. Interestingly, online booking service sites like HolidayCheck.com provides incentives for people to just share their experience on the website.

We extend our pricing model and results to optimally design both prices and incentives so as to guarantee influence among agents by soliciting influence actions. Examples of incentives can be a small discount in exchange for a simple action such as liking the product, and a more significant discount by writing a detailed review. With this more general setting, the seller can ensure

the influence among the agents so that the network externalities effects are guaranteed to occur. Interestingly, the methods and results we develop for our previous model extend to this model.

## Literature review

Models that incorporate local network externalities find their origins in the works of [12] and [18]. These early papers assume that consumers are affected by the global consumption of all other players. In other words, the network effects are of global nature, i.e., the utility of a consumer depends directly on the behavior of the entire set of agents in the network. In our model, consumers directly interact only with a subset of agents, also known as their neighbors. Although interactions are of a local nature, the utility of each player may still depend on the global structure of the network, because each agent potentially interacts indirectly with a much larger set of agents than just his neighbors.

Models of local network externalities which explicitly take into account the network structure have been proposed in several papers including [3], [5] and [22]. [3] proposes a model in which individuals located in a network choose actions (criminal activities) that affect the payoffs of other individuals in the network using linear influence models that we adopt as well. [5] studies a setting where firms sell to consumers located on a network with local adoption externalities. They characterize networks that can sustain different technologies in equilibrium and show that even if rival firms engage in Bertrand competition, this form of network externalities permits strong market segmentation. The paper by [22] presents a model of local network effects in which agents connected in a social network value the adoption of a product by a heterogeneous subset of other agents in their neighborhood and have incomplete information about the adoption complementarities between all other agents. Another related area of research is network games. Our second stage problem takes the form of a network game where agents in a network interact with each other. A recent series of papers that study network games include [14] and the references therein.

Several recent papers explicitly include the interactions among agents in social networks to study the social network's effects on various marketing problems. The first among these are the works on influence maximization by [11] and [19] which aimed to identify influential agents in a network. Several recent papers such as those by [16], [1] and [2] extend these works to study optimal pricing strategies in networks. [16] focuses on viral marketing strategies for revenue maximization where the agents are offered the product in a sequential manner and show simple two-price strategies (the *Influence-and-Exploit* strategy where the seller chooses a set of consumers who get the product for free and use the optimal myopic pricing for the rest) performs very well relative to the optimal strategy which is NP-hard. [1] extends this approach to a multi-stage model where the seller sets different prices for each stage. [2] allows buyers to buy the product with a certain probability if the product is recommended by their friends who purchased the item. The main difference in the approaches in these papers and ours stems from the timing of purchasing decisions. These papers consider sequential purchases where myopic consumers base their consumption decisions on the number of consumers who have already bought the product. In our work, we consider a simultaneous purchasing decision for all the agents in the network, who are fully rational.

In the literature and in our work, pricing with simultaneous purchasing decisions is set up as a two-stage game, where the seller designs the prices in the first stage and agents respond by playing their purchasing decisions. Rational behavior in this case is captured by the Nash equilibrium (or Bayesian Nash equilibrium if the information is incomplete). Three papers in this context closely related to our work are [9], [8] and [10]. [9] studies optimal pricing strategies for a divisible good with linear social influence functions so as to maximize the profits of the seller who has complete information of the network. They provide efficient algorithms to compute fully discriminative prices as well as the uniform optimal price and show that the problem is NP-hard when the monopolist is restricted to two pre-specified prices. [8] and [10] study the optimal pricing problem of an indivisible item with linear utility functions

under incomplete information. [8] studies externalities resulting either from local network interactions or from prices and distinguish between a single global monopoly and several local monopolies. [10] assumes a uniform prior and propose an efficient uniform pricing algorithm for revenue maximization but show that the general discriminative pricing problem is NP-hard. Our work is in the similar light of the three aforementioned papers for the case of an indivisible item under complete information. However, the model and techniques required to address our setting differ significantly from these papers. In particular, the second-stage equilibrium in our setting can be characterized only with a system of equilibrium constraints that happen to be highly non-linear and non-convex with integer variables and this entirely alters the simple quadratic form of optimal pricing problem that could be solved in closed form as observed by [9] and [8]. In fact, in our setting the optimal pricing problem is cast as a mathematical program with equilibrium constraints (MPEC) (see [20]). We refer the reader to the books by [21] and [6] for the integer programming reformulation techniques that we use in this work to address the non-convexities and arrive at a MIP.

Due to the modeling flexibility of our approach, to the best of our knowledge, our work is the first to provide an explicit optimization formulation for the pricing problem that can incorporate business rules on prices and constraints on market segmentation. Moreover, we are also able to extend the model and results to a practical setting where potential buyers are explicitly given incentives, potentially different for each agent, and a choice to influence their neighbors in addition to a price.

Finally, another recent paper on incorporating the effects of social network influence but unrelated to pricing is by [15]. The latter investigates social network influence in the context of product design, in particular, in the share-of-choice problem, and construct a genetic algorithm to solve the problem.

The remaining of the chapter is organized as follows. In Section 5.3, we describe the model and our assumptions as well as the dynamics of the two-stage game. In Section 3.3, we show the existence of a pure strategy Nash equilibrium for the second-

stage purchasing game. We use duality theory to formulate the problem as a MIP in Section 3.4. In Section 5.4, we derive efficient algorithms to solve the MIP for the discriminative and uniform pricing strategies and derive insights about the optimal solution. In Section 3.6, we draw some insights for the case where cross valuations are uncertain. In Section 3.7, we extend our model to the case when the seller can design both price and incentives to guarantee the influence among agents in the network. In Section 3.8, we present computational experiments to draw some qualitative insights. Finally, we present our conclusions in Section 3.9. The proofs of the different propositions and theorems are provided in the Appendix.

## 3.2 Model

Consider a monopolistic firm selling an indivisible product to  $N$  agents denoted by the set  $\mathcal{I} = \{1, \dots, N\}$ , embedded in a social network. We denote the set of *value interaction* constants for this product for any agent  $i \in \mathcal{I}$  by  $G_i = \{g_{S,i} | S \subset \mathcal{I} \setminus \{i\}\}$ , where the element  $g_{S,i}$  represents the marginal increase in value that agent  $i$  obtains by owning the product when all the agents in  $S$  also own the product. In particular,  $g_{\emptyset,i}$  (also denoted by  $g_i$ ) is the marginal value that agent  $i$  derives by owning the product. If agent  $j$  does not influence agent  $i$  then all the terms  $g_{S,i}$  where  $j \in S$  are zero. On the other hand, if  $j$  influences agent  $i$  then at least one of the terms  $g_{S,i}$  where  $j \in S$  is non-zero. In this case, we refer to  $j$  as a *neighbor* of agent  $i$ .

*Assumption 3.* We make the following assumptions regarding the elements in  $G_i \forall i \in \mathcal{I}$  and the corresponding model of influence.

- a. The firm and the agents have perfect knowledge of the externalities i.e., everyone knows  $G_i$ .
- b. **Additive influence model:** The total value that an agent derives when purchasing the item is assumed to be the sum of the agent's own valuation and the valuation derived from the influences of all the subsets of his friends who own

the item. Mathematically, agent  $i$ 's valuation model is given by:

$$v_i(\alpha_i, \boldsymbol{\alpha}_{-i}) = \alpha_i \left[ \sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \prod_{j \in S} \alpha_j \right], \quad (3.2.1)$$

where  $\alpha_i \in \{0, 1\}$  is a binary variable that represents the purchasing decision of agent  $i$  and  $\boldsymbol{\alpha}_{-i}$  represents the vector of purchasing decisions of all the agents but  $i$  in the network. Here,  $\prod$  denotes the product operator.

- c. **Generalized non-negativity:** The total influence of any neighbor  $j$  (including the empty set) on agent  $i$  is positive. Mathematically, for each agent  $i \in \mathcal{I}$ , we have:

$$\sum_{\substack{S \ni j, \\ S \setminus j \subset T}} g_{S,i} \geq 0 \quad \forall j \in \mathcal{I} \setminus \{i\}, T \subset \mathcal{I} \setminus \{i, j\} \quad (3.2.2)$$

$$g_i \geq 0. \quad (3.2.3)$$

Here,  $S \ni j$  denotes the set  $S$  that contains agent  $j$ . The set  $T$  can be interpreted as the set of agents who buy the item.

The perfect information assumption is made for simplicity and tractability of the analysis to obtain insights. This is later relaxed in Section 3.6. The additive influence model in (3.2.1) captures a broad class of linear and non-linear influence functions using a set theoretic approach. We highlight a few with examples below. The generalized non-negativity assumption assumes a net positive externality of any agent on another. This interestingly does not reduce to all the  $g_{S,i}$  being positive. This will be illustrated in one of the examples below. Note that this assumption results in the total valuation of an agent being positive (but not the vice versa), i.e.,

$$\sum_{S \subset T} g_{S,i} \geq 0 \quad \forall T \subset \mathcal{I} \setminus \{i\}. \quad (3.2.4)$$

This is obtained by applying the generalized non-negativity assumption multiple times

by adding agents sequentially to form the set  $T$ .

In general, terms  $g_{S,i}$  with a large value of  $|S|$  are usually negligible and this motivates the following definition for a class of influence models based on the maximum subset size that influence an agent.

**Definition 3.1.**  *$K$ -wise influence model is a model where  $g_{S,i} = 0$  for all subsets  $|S| > K - 1$ , where  $|\cdot|$  refers to the cardinality of a set.*

These models capture different degrees of influence in a network and can model different types of influence behavior. We now present two examples of the  $K$ -wise influence models ( $K = 2, 3$ ) that satisfy the generalized non-negativity assumption and can capture linear and non-linear (e.g., submodular, supermodular) influence effects.

1. **Pairwise influence:** This is the 2-wise model where  $g_{S,i} = 0$  for all subsets  $|S| > 1$ , and  $g_i, g_{j,i} \geq 0 \forall i, j \in \mathcal{I}$ . More specifically, we have:

$$v_i(\alpha_i, \boldsymbol{\alpha}_{-i}) = \alpha_i \left[ g_i + \sum_{j \in \mathcal{I} \setminus \{i\}} \alpha_j g_{j,i} \right]. \quad (3.2.5)$$

This utility function captures only the independent and marginal influence of each neighbor on an agent and is linear and additive across neighbors. These type of influence models has been the focus of several earlier papers such as [9], [8] and [10].

2. **Triple-wise Influence:** This is the 3-wise model where  $g_{S,i} = 0$  for all subsets  $|S| > 2$ . In particular, we have:

$$v_i(\alpha_i, \boldsymbol{\alpha}_{-i}) = \alpha_i \left[ g_i + \sum_{j \in \mathcal{I} \setminus \{i\}} \alpha_j g_{j,i} + \sum_{j,k \in \mathcal{I} \setminus \{i\}; j \neq k} \alpha_j \alpha_k g_{\{j,k\},i} \right]. \quad (3.2.6)$$

In addition, we can have either of the following:

- a.  $g_i, g_{j,i}, g_{\{j,k\},i} \geq 0 \forall i, j, k \in \mathcal{I}$ ; or,

- b.  $g_i, g_{j,i} \geq 0$  and  $g_{\{j,k\},i} \leq 0 \forall i, j, k \in \mathcal{I}$  such that  $g_{j,i} + \sum_{k \in \mathcal{I} \setminus \{i,j\}} g_{\{j,k\},i} \geq 0$ .

The model in (a) is a special case of a supermodular influence model, where the marginal effect of an additional neighbor increases with the set of existing influencers. The model in (b) is an example of a submodular influence model, where the marginal effect of an additional neighbor decreases with the set of existing influencers. Observe that these effects are characterized by influence terms that do not decompose by agents and are hence non-linear but additive on the subsets of neighbors. In general, one can consider a mixed model where some of the influences are supermodular and some are submodular, as long as the generalized non-negativity assumption is satisfied. Note also that the generalized non-negativity assumption does not restrict to positive values of influence as indicated in model (a).

Let the vector  $\mathbf{p} \in \mathbf{P}$  denote the prices offered to all the agents by the firm. In particular,  $p_i \in \mathbb{R}_+$  is the  $i^{\text{th}}$  element of the vector  $\mathbf{p}$  and represents the price offered to agent  $i$ . Here,  $\mathbf{P}$  is assumed to be a polyhedral set that represents the feasible pricing strategies of the firm, which possibly includes several business constraints on prices and on network segmentation. For example, the firm can adopt a discriminative pricing strategy where each agent may potentially receive a different price. In this case,  $\mathbf{P} = \mathbb{R}_+^N$ . In addition, one can restrict the values of these prices to lie between  $p_L$  and  $p_U \geq p_L$ , i.e.,  $\mathbf{P} = \{\mathbf{p} \in \mathbb{R}_+^N \mid p_L \leq p_i \leq p_U \forall i\}$ . A common pricing strategy is to adopt a single uniform price for all the agents across the network. Here,  $\mathbf{P} = \{\mathbf{p} \in \mathbb{R}_+^N \mid p_i = \bar{p} \forall i, \bar{p} \in \mathbb{R}_+\}$ . In a similar fashion, depending on the application, the firm can select some appropriate business constraints to impose on the pricing strategy. Finally,  $\mathbf{P}$  can also incorporate specific constraints on the network segmentation. For example, motivated by business practices, a particular segment of agents should be offered the same price. Alternatively, special members (such as loyal customers) should receive a lower price than regular customers.

Our goal is to develop a general and efficient optimal pricing method for the

firm that incorporates the different business rules as constraints. In addition, we are interested in deriving some properties of the optimal prices depending on the different parameters of the model. Before we mathematically formulate the problems of the potential buyers and the firm, we summarize our assumptions.

*Assumption 4.* We assume the following about each agent  $i \in \mathcal{I}$  in the network.

- a. Each agent has a utility model as described below in (3.2.8).
- b. Each agent is assumed to be rational and a utility maximizer.
- c. Each agent can buy at most one unit of the item and either fully purchases the item or does not purchase it at all.
- d. If the utility of an agent is exactly zero, the tie is broken assuming this agent buys the item.

We discuss in the next section how the last assumption is not restrictive in our problem. In particular, the seller can always reduce the offered price by a very small amount  $\epsilon \geq 0$  to break the tie so that the agent is strictly better off by buying the item. We next formalize the assumption for the seller.

*Assumption 5.* We assume that the seller is a profit maximizer as described in (3.2.10) with a linear manufacturing cost.

For a given set of prices chosen by the seller, the agents in the network aim to collectively maximize their utility from purchasing the item. We capture the utility model of an agent as the total value minus the price paid. In particular, the utility of agent  $i$  is given by:

$$u_i(\alpha_i, \boldsymbol{\alpha}_{-i}, p_i) = v_i(\alpha_i, \boldsymbol{\alpha}_{-i}) - \alpha_i p_i \quad (3.2.7)$$

$$= \alpha_i \left[ \sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \prod_{j \in S} \alpha_j - p_i \right]. \quad (3.2.8)$$

If  $\alpha_i = 1$ , agent  $i$  purchases the item. In this case, agent  $i$  derives a utility equal to  $\sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \prod_{j \in S} \alpha_j - p_i$ . If  $\alpha_i = 0$ , the agent does not purchase the item and derives zero utility.

The utility maximization problem of agent  $i$  can be written as follows:

$$\max_{\alpha_i \in \{0,1\}} u_i(\alpha_i, \boldsymbol{\alpha}_{-i}, p_i). \quad (3.2.9)$$

The profit maximizing problem of the seller is given by:

$$\max_{\mathbf{p} \in \mathbf{P}} \sum_{i \in \mathcal{I}} \alpha_i (p_i - c), \quad (3.2.10)$$

where the vector  $\boldsymbol{\alpha}$  represents the purchasing decisions of the agents obtained from the utility maximization problems (3.2.9) and  $c$  is the unit manufacturing cost. If agent  $i$  decides to buy the product at the offered price  $p_i$ ,  $\alpha_i$  is equal to 1 and the firm incurs a profit of  $p_i - c$ . If agent  $i$  decides not to purchase the item, it incurs zero profit to the seller.

We view the entire problem as a two-stage Stackelberg game and refer to it as the pricing-purchasing game. First, the seller leads the game by choosing the vector of prices  $\mathbf{p}$  to be offered to the buyers. Second, the agents follow by deciding whether or not to purchase the item at the offered prices. In other words, the firm sets the prices  $\mathbf{p} \in \mathbf{P}$  and the network of agents collectively follow with their decisions,  $\alpha_i \forall i \in \mathcal{I}$ . We are interested in *subgame perfect equilibria* of this two-stage pricing-purchasing game (see e.g., [13]). For a fixed vector of prices offered by the seller, the equilibria of the second stage game, referred to as the *purchasing equilibria*, are defined by:

$$\alpha_i^* \in \arg \max_{\alpha_i \in \{0,1\}} u_i(\alpha_i, \boldsymbol{\alpha}_{-i}^*, p_i) \quad \forall i \in \mathcal{I}. \quad (3.2.11)$$

We note that this definition is similar to the consumption equilibria for a divisible item (or service) in [9]. However, in our case the decision variables  $\alpha_i$  are restricted to be binary so that agents cannot buy fractional amounts of the item.

We also note that the overall two-stage problem is non-linear and non-convex as it includes terms of the form  $\alpha_i p_i$  in the seller's objective function and  $\alpha_i \prod_{j \in S} \alpha_j$  in the objective functions of the agents (which we will shortly see appear as constraints in the seller's problem). In addition, the discrete nature of the purchasing decisions increases the complexity of the problem, as it yields a non-convex integer program. In the next section, we start by considering the second stage purchasing game and show the existence of an equilibrium such as in (3.2.11), for any given vector of prices. We then characterize the equilibria by a set of constraints for any price vector. In Section 3.4, we use this characterization to formulate the optimal pricing problem as a MIP.

### 3.3 Purchasing equilibria

In this section, we consider the second stage purchasing game and show the existence of a pure Nash equilibrium (PNE) strategy, given any vector of prices  $\mathbf{p}$ . We observe that there could be multiple pure Nash equilibria for this game but we characterize all these equilibria through a system of constraints using duality theory. We also identify a mild condition that allows us to focus on the purchasing equilibrium that is preferred by both the seller and the network of agents. We later show that our optimization formulation naturally induces this preferred purchasing equilibrium.

**Theorem 3.1.** *Consider the second stage game played collectively by the network of agents.*

1. *The second stage game has at least one pure Nash equilibrium for any given vector of prices  $\mathbf{p}$  chosen by the seller.*
2. *There exists a small  $\epsilon \geq 0$  such that a price perturbation  $p_i - \epsilon \forall i \in \mathcal{I}$  does not change any of the PNEs. In addition, it ensures that all the agents in all the PNEs have no ties in their utility (between buying and not buying).*

3. *Among the multiple PNEs at the perturbed prices, there exists a Pareto optimal PNE where every agent's utility is at least as large as in any other PNE and is strictly better for at least one agent. This implies all the agents that buy in any PNE will also buy in the Pareto optimal PNE while deriving a strictly positive and higher utility.*

We note that the first part of Theorem 3.1 guarantees the existence of a PNE but not necessarily its uniqueness. Consider the following simple example in which two distinct PNEs arise. Assume a network with two symmetric agents that mutually influence one another:  $g_1 = g_2 = 2$  and  $g_{21} = g_{12} = 1$ . Consider the given price vector:  $p_1 = p_2 = 2.5$ . In this case, we have two PNEs: buy-buy and no buy-no buy. In other words, if player 1 buys, player 2 should buy but if player 1 does not buy, player 2 will not either. Therefore, uniqueness is not guaranteed. More precisely, for any price strictly larger than 3 or strictly smaller than 2, we have a unique equilibrium but for any price between 2 and 3, we have multiple equilibria. Nevertheless, for any price between 2 and  $3 - \epsilon$ , the purchasing equilibrium is preferred by the agents as they each derive a positive utility from buying. In particular, for any price but  $3 - \epsilon$  can be taken to be 0 and for  $p = 3$ , any small positive number will work. Note that even though there exist multiple equilibria, by reducing the price by  $\epsilon$ , the purchasing equilibria is strictly preferred by both agents.

The purpose of  $\epsilon$  as can be noted from the above example is to break ties and encourage agents with ties to buy (see Assumption 4 part d) without affecting other agents' decisions. Note also that the value of  $\epsilon$  can be taken very small so that it does not affect the revenue of the seller significantly. For example, by assuming that all the influence values  $g_{S,i}$  and the prices are integers, the value of  $\epsilon$  can be set to any positive number lower than 1.

In the last part of Theorem 3.1, we show that the seller's preferred equilibrium (in terms of revenue<sup>1</sup>) is preferred by all buyers (in any of the equilibria) with strictly

---

<sup>1</sup>Note that in the first stage game, the seller aims to maximize profits. If he does not want an agent to buy (e.g., an agent whose price is less than cost), he can always increase the prices to a

positive and higher utility after a minute price perturbation. The generalized non-negativity assumption (3.2.2) is crucial to show this result. With a less restrictive assumption such as (3.2.4), if one agent switches from not buying to buying, it can decrease the utility of some other agents. Nevertheless with the generalized non-negativity assumption, the last part of Theorem 3.1 shows that there is a unique preferred equilibrium by both the seller and the buyers. Without loss of generality, the analysis in the rest of the chapter focuses only on this equilibrium where the seller uses a price perturbation as a tactic to break ties.

We note that the existence of a PNE for the second stage game is not always guaranteed for the case with negative externalities. Nevertheless, it may be possible to find sufficient conditions on the valuation functions to ensure the existence. This direction is beyond the focus of this work and hence not pursued.

In Section 3.4, we determine that the nature of the first stage game always induces the purchasing equilibrium of interest (in the above example buy-buy, i.e., ties are broken by buying).

## Characterization of the purchasing equilibria

The natural next step is to characterize the purchasing equilibria as a function of the prices. In other words, we would like to characterize the functions  $\alpha_i(\mathbf{p}) \forall i \in \mathcal{I}$ . This will allow us to reduce the two-stage problem to a single optimization formulation, where the only variables are the prices. In our setting, a closed form expression for  $\alpha_i(\mathbf{p})$  is not straightforward. Instead, by using duality theory, we characterize the set of constraints the equilibria should satisfy for any given vector of prices. We begin by making the following observation regarding the utility maximization problem of each agent.

**Observation 3.1.** Given a vector of prices  $\mathbf{p}$ , let us consider subproblem 3.2.9 for agent  $i$ . If the decisions of the other agents  $\alpha_{-i}$  are given, the problem of agent  $i$  has

---

very large value so that the agent does not buy. But if he indeed does want the agent to buy, then he is willing to incur a small loss in revenue.

a tight linear programming (LP) relaxation.

In fact, for fixed values of  $\mathbf{p}$  and  $\boldsymbol{\alpha}_{-i}$ , the subproblem faced by agent  $i$  happens to be a very simple assignment problem. More specifically, let us consider the LP obtained by the continuous relaxation of the binary constraints  $\alpha_i \in \{0, 1\}$  to  $0 \leq \alpha_i \leq 1$ . One can view this LP as a relaxation purchasing game where agents can purchase fractional amounts of the item and therefore adopt mixed strategies. If the quantity  $\left( \sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \prod_{j \in S} \alpha_j - p_i \right)$  (which is exactly known since  $\mathbf{p}$  and  $\boldsymbol{\alpha}_{-i}$  are given) is positive,  $\alpha_i^* = 1$  and if this quantity is negative,  $\alpha_i^* = 0$ . Finally, if this quantity is equal to zero,  $\alpha_i^*$  can be any number in  $[0, 1]$  so that the agent is indifferent between buying and not buying the item. Therefore, the LP relaxation of the subproblem of agent  $i$  for fixed values of  $\mathbf{p}$  and  $\boldsymbol{\alpha}_{-i}$  is integral, meaning that all the extreme points are integer.

Observation 3.1 allows us to transform subproblem 3.2.9 for agent  $i$  into a set of constraints by using duality theory of linear programming. More specifically, this set of constraints consists of primal feasibility, dual feasibility and strong duality conditions. For agent  $i$ , the constraints can be written as follows:

$$\text{Primal feasibility:} \quad 0 \leq \alpha_i \leq 1 \quad (3.3.1)$$

$$\text{Dual feasibility:} \quad y_i \geq \sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \prod_{j \in S} \alpha_j - p_i \quad (3.3.2)$$

$$y_i \geq 0 \quad (3.3.3)$$

$$\text{Strong duality:} \quad y_i = \alpha_i \left( \sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \prod_{j \in S} \alpha_j - p_i \right) \quad (3.3.4)$$

Here, the variable  $y_i$  represents the dual variable of subproblem 3.2.9 for agent  $i$ . Combining the above constraints (3.3.1–3.3.4) for all the agents  $i \in \mathcal{I}$  characterizes all the equilibria (mixed and pure) of the second stage game as a function of the prices. In order to restrict our attention to the pure Nash equilibria (that the existence is guaranteed by Theorem 3.1), one can impose  $\alpha_i \in \{0, 1\} \forall i$ . Observe that this characterization has reduced  $N + 1$  interconnected problems to be compactly written

as a single optimization formulation. We note that the number of variables increases by  $N$ , as we add a dual continuous variable for each agent's subproblem.

### 3.4 Optimal pricing: MIP formulation

In this section, we use the existence and characterization of the PNEs to transform the two-stage optimal pricing problem into a single optimization formulation. This formulation happens to be a non-convex integer program but exhibits some interesting properties. We then reformulate the problem to arrive at a MIP with linear constraints.

We next formulate the pricing problem faced by the seller (denoted by problem Z) by incorporating the second stage PNE characterized by the set of constraints (3.3.1–3.3.4) for each agent. The class of optimization problems with equilibrium constraints is referred to as MPEC (Mathematical Program with Equilibrium Constraints) and is well studied in the literature (see e.g., [20]). The equilibrium constraints in our case include constraints (3.3.2–3.3.4) and  $\alpha_i \in \{0, 1\}$  instead of constraint (3.3.1) for all agents to restrict to the pure Nash equilibria. The formulation is given by:

$$\begin{aligned}
 & \max_{\substack{\mathbf{p} \in \mathbf{P} \\ \mathbf{y}, \boldsymbol{\alpha}}} \sum_{i \in \mathcal{I}} \alpha_i (p_i - c) & \text{(Z)} \\
 & \text{s.t.} \\
 & \left. \begin{aligned}
 y_i &= \alpha_i \left( \sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \prod_{j \in S} \alpha_j - p_i \right) \\
 y_i &\geq \sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \prod_{j \in S} \alpha_j - p_i \\
 y_i &\geq 0 \\
 \alpha_i &\in \{0, 1\}
 \end{aligned} \right\} \forall i \in \mathcal{I}
 \end{aligned}$$

In addition to the presence of binary variables, one can see that the above optimization problem is non-linear and non-convex as it includes terms of the form  $\alpha_i \prod_{j \in S} \alpha_j$  and

$\alpha_i p_i$ . Therefore, problem Z is not easily solvable by commercially available solvers. We next show an interesting tightness result for problem Z. This allows us to consider a non-convex continuous optimization problem instead of an integer program. This also provides insight to one of our main results presented in Theorem 3.2.

**Proposition 3.1.** *Problem Z admits a tight continuous relaxation.*

The main intuition behind the proof is the following. At optimality,  $\alpha_i$  is fractional only if there are ties between buying and not buying. Therefore, the buyer's decision can be altered to buying (i.e., made integral) without impacting the decision of the other buyers or decreasing the seller's profit. We next show that by introducing a few additional continuous variables, one can reformulate problem Z into an equivalent MIP with the same number of binary variables. To simplify the illustration, we present the MIP for the  $K$ -wise model of influence. We first define the following additional variables:

$$\begin{aligned} z_i &= \alpha_i p_i & \forall i \in \mathcal{I} \\ \alpha_S &= \prod_{j \in S} \alpha_j & \forall 1 < |S| \leq K, S \subset \mathcal{I}. \end{aligned}$$

By using the binary nature of the variables and adding certain linear constraints, one can replace all the non-linear terms in problem Z. This yields the following MIP formulation denoted by Z-MIP:

$$\max_{\substack{\mathbf{p} \in \mathbf{P} \\ \mathbf{y}, \mathbf{z}, \boldsymbol{\alpha}}} \sum_{i \in \mathcal{I}} (z_i - c\alpha_i) \quad (\text{Z-MIP})$$

s.t.

$$\left. \begin{aligned} y_i &= \sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \alpha_{S \cup i} - z_i \\ y_i &\geq \sum_{S \subset \mathcal{I} \setminus \{i\}} g_{S,i} \alpha_S - p_i \\ y_i &\geq 0 \end{aligned} \right\} \forall i \in \mathcal{I} \quad (3.4.1)$$

$$\left. \begin{aligned} z_i &\geq 0 \\ z_i &\leq p_i \\ z_i &\leq \alpha_i p^{max} \\ z_i &\geq p_i - (1 - \alpha_i) p^{max} \end{aligned} \right\} \forall i \in \mathcal{I} \quad (3.4.2)$$

$$\left. \begin{aligned} \alpha_S &\geq 0 \\ \alpha_S &\leq \alpha_{S \setminus \{i\}} \quad \forall i \in S \end{aligned} \right\} \forall 1 < |S| \leq K, S \subset \mathcal{I} \quad (3.4.3)$$

$$\alpha_{S \cup \{i,j\}} \geq \alpha_{S \cup \{i\}} + \alpha_{S \cup \{j\}} - \alpha_S \quad \forall |S| < K - 1, S \subset \mathcal{I} \setminus \{i, j\}, i \neq j \in \mathcal{I} \quad (3.4.4)$$

$$\alpha_i \in \{0, 1\} \quad \forall i \in \mathcal{I} \quad (3.4.5)$$

$$\alpha_S = 0 \quad \forall |S| > K, S \subset \mathcal{I} \quad (3.4.6)$$

$$\alpha_\emptyset = 1 \quad (3.4.7)$$

In the above formulation,  $p^{max}$  denotes the maximal price allowed and is typically known from the context. For example, one can take:  $p^{max} = \max_i \left\{ \sum_{S \subset \mathcal{I}, g_{S,i} > 0} g_{S,i} \right\}$  without affecting the problem at all, since no agent would ever pay a price beyond that value. The set of constraints (3.4.2) aims to guarantee the definition of the

variable  $z_i$ , whereas the sets of constraints (3.4.3) and (3.4.4) ensure the correctness of the variable  $\alpha_S$ . For example, the constraint (3.4.4) for the pair of agents  $i$  and  $j$  is given by:  $\alpha_{i,j} \geq \alpha_i + \alpha_j - 1$  (here, we took the set  $S$  to be the null set). One can note that in the above Z-MIP formulation with a  $K$ -wise influence model, we have a total of (at most)  $4N + \sum_{k=2}^K \binom{N}{k}$  variables ( $4N$  for the  $\alpha_i$ ,  $\mathbf{p}$ ,  $\mathbf{y}$  and  $\mathbf{z}$  and the rest are the continuous  $\alpha_S$ ). Note that only  $N$  variables are binary, while the remaining are all continuous. In particular, in the pairwise influence setting we have  $4N + \binom{N}{2}$  variables and in the triple-wise influence setting we have  $4N + \binom{N}{2} + \binom{N}{3}$  variables. In other words, for small values of  $K$  (e.g., 3 which is sufficiently rich to capture submodular or supermodular influence), the number of variables is a small degree polynomial in  $N$ .

We conclude that our problem of designing prices for selling an indivisible good to agents embedded in a social network can be formulated as a MIP. This MIP is equivalent to the two-stage non-convex IP game we started with. As a result, one can easily incorporate various business constraints such as pricing policies, market segmentation, inter-buyer price constraints, just to name a few. In other words, this formulation can be viewed as an operational tool to solve the optimal pricing problem of the seller. This is in contrast to previous approaches that proposed tailored algorithms for the problem where one cannot easily incorporate business rules. However, solving a MIP may not be very scalable. If the size of the network is not very large, one can still solve it efficiently using commercially available MIP solvers. Moreover, it is possible to solve the problem offline (before launching a new product for example) so that the running time may not be of the highest consideration. Potentially, one can also consider network clustering methods to aggregate or coalesce several nodes to a single virtual agent in order to reduce the size of the network. If the size of the network is very large, one needs to find more efficient methods to solve the Z-MIP problem. In the next section, we derive efficient methods (polynomial in the number of agents and very fast in practice) to solve the problem to optimality for two popular pricing strategies.

## 3.5 Efficient algorithms

### 3.5.1 Discriminative Prices

In this section, we consider the general pricing strategy where the firm offers discriminative prices that potentially differ for each agent, depending on his influence in the network. In particular,  $\mathbf{P} = \mathbb{R}_+^N$  in problem Z-MIP. This scenario is of interest in various practical settings where the seller gathers the purchasing history of each potential buyer, his geographical location as well as other attributes. It can also be used by the seller to understand who are the influential agents in the network and what is the maximal profit he can potentially achieve if he were to discriminate prices at the individual level. The prices can then be implemented by setting a ticket price that is the same for all the agents and sending out coupons with discriminative discounts to the potential buyers. In fact, in practice, it often occurs that people receive different deals for the same item depending on the loyalty class, purchase history and geographical location. It is also common that a very small number of highly influential people (e.g., certified bloggers) receive the item for free or at a very low price. The method we propose aims to provide a systematic and automated way of finding the prices (equivalently, the discounts) to offer to the agents embedded in a social network based on their influence so as to maximize the total profit of the seller.

As we discussed, solving the Z-MIP problem presented in the previous section using an optimization solver may be impractical for a large network. We next show that solving the LP relaxation of the Z-MIP problem yields the desired optimal integer solution. Consequently, one can solve the problem efficiently (polynomial in the number of agents and very fast in practice) and obtain an optimal solution even for very large networks. Recall that the linearization of problem Z was possible due to the integrality of the decision variables. In other words, in order to reformulate problem Z into Z-MIP, the binary restriction was crucial. As a result, by introducing the new variables  $z_i$  and  $\alpha_S$ , one may potentially obtain fractional solutions that cannot be implemented in practice. However, the following Theorem shows that the optimal

solutions of Z-MIP can be identified using its continuous relaxation.

**Theorem 3.2.** *The optimal discriminative pricing solution of Z-MIP can be obtained efficiently (polynomial in the number of agents). In particular, problem Z-MIP admits a tight LP relaxation.*

At a high level, we order agents in increasing order of their  $\alpha_i$  values and sequentially increase each agent's  $\alpha_S$  values (where  $i \in S$ ) to the next agent's  $\alpha_i$  value (equal to 1 in the last iteration) without impacting the decision of the other buyers or decreasing the seller's profit while maintaining feasibility of the LP relaxation. This process is repeated until all agents who bought fractional values finally purchase the item. The proof can thus be viewed as a constructive method of rounding the fractional LP solution to obtain an integer solution with at least the same profit level. One can employ this constructive method or use a method like simplex to arrive at the optimal extreme points which are guaranteed to be integer.

The result of Theorem 3.2 suggests an efficient method of solving the problem that we formulated as a two stage non-convex integer program. The LP based method inherits all the complexity properties of linear programming and is thus scalable and applicable to large networks.

We next explore some properties of the optimal solution. Interestingly, if we know the agents that buy the item, the corresponding optimal prices can be obtained in closed form as summarized in the following observation.

**Observation 3.2.** If  $T^*$  is the set of agents that buy at optimality, i.e.,  $T^* = \{i \in \mathcal{I} | \alpha_i^* = 1\}$ , then the optimized prices are given by:

$$p_i^* = \begin{cases} \sum_{S \subset T^* \setminus \{i\}} g_{S,i} & \forall i \in T^* \\ p^{max} & \text{otherwise.} \end{cases} \quad (3.5.1)$$

In other words, the price of an agent that buys is the sum of the agent's own value for the item ( $g_i$ ) and a markup term that corresponds to the influence of the "buying"

network on this agent. Membership into the buying class itself depends on the self value  $g_i$  and the influence the agent exerts on the network. The corresponding optimal profit is given by:

$$\Pi^* = \sum_{i \in T^*} \left( \sum_{S \subset T^* \setminus \{i\}} g_{S,i} - c_i \right) \quad (3.5.2)$$

$$= \underbrace{\sum_{\{i \in T^* | g_i \geq c\}} (g_i - c)}_{\text{Profit in the absence of network effects}} + \underbrace{\sum_{\{i \in T^* | g_i < c\}} (g_i - c) + \sum_{\substack{i \in T^* \\ \emptyset \neq S \subset T^* \setminus \{i\}}} g_{S,i}}_{\text{Incremental profit due to network effects}}. \quad (3.5.3)$$

We divide the profit into two components. First, the profitable component of individual valuations which the seller perceives even in the absence of network effects. Second, the incremental profit due to network effects which balances the seller's profit from unprofitable individual valuations with the revenue gain from network interactions.

These prices are optimal because they result in the maximal profit the seller can extract given that the agents in  $T^*$  are buying. One can see that  $T^*$  includes all the agents in  $\mathcal{I}$  whose individual valuations are profitable (also in part due to generalized non-negativity). In addition, some agents that buy can be offered a price below cost. One can view these agents as influencers who get membership into the buying class because of their strong influence in the network. The seller can tap into this source of profits only by taking advantage of the network effects.

We next use this closed form solution to compare the prices and profit between the linear, supermodular and submodular influence models provided in (3.2.5), (3.2.6)a, and (3.2.6)b respectively. Consider a setting where  $g_i \geq 0$  and  $g_{j,i} \geq 0$  are exactly the same across the three models but differ in the  $g_{\{j,k\},i}$  values for any  $i, j, k \in \mathcal{I}$ , while satisfying the generalized non-negativity conditions. In particular,  $g_{\{j,k\},i}$  values are negative in the submodular model, zero in the linear model and positive in the supermodular model. We denote the problem and optimized solutions of the submodular, linear and supermodular models with subscripts *sub*, *lin* and *sup* respectively and

make the following observations.

**Observation 3.3.** The optimized solutions satisfies the following trends:

- $T_{sub}^* \subset T_{lin}^* \subset T_{sup}^*$  where  $T^* = \{i \in \mathcal{I} | \alpha_i^* = 1\}$ .
- $p_{i,sub}^* \leq p_{i,lin}^* \forall i \in T_{sub}^*$  and  $p_{i,lin}^* \leq p_{i,sup}^* \forall i \in T_{lin}^*$ .
- $\Pi_{sub}^* \leq \Pi_{lin}^* \leq \Pi_{sup}^*$  where  $\Pi^*$  denotes the optimal profit.

In other words, as we move from the submodular to the linear and then to the supermodular influence models, additional agents will buy. In addition, the buyers will pay a higher price so that it induces a larger profit for the seller. In particular, as the  $g_{\{j,k\},i}$  term increases, it results in a larger number of agents buying the item. The seller can also charge higher prices by using equation (3.5.1). Consequently, the seller gathers at least as much profit in a linear (or supermodular) model as the submodular (or linear) model from the agents  $T_{sub}^*$  (or  $T_{lin}^*$ ). The additional agents that buy, i.e., the agents in  $T_{sub}^* \setminus T_{lin}^*$  or  $T_{lin}^* \setminus T_{sup}^*$  can only increase the profit of the seller.

### 3.5.2 Uniform Price

In this section, we consider the case where the seller offers a uniform price across the network while incorporating the effects of social interactions. This scenario may arise when the firm may not want to price discriminate due to fairness or ethical reasons and prefers to offer a uniform price. We observe that a similar result to Theorem 3.2 for the uniform pricing case does not hold. In other words, by adding the linear uniform price constraint:  $p_1 = p_2 = \dots = p_N$  to the Z-MIP formulation, the corresponding LP relaxation is no longer tight and we obtain fractional solutions that cannot be implemented in practice. Geometrically, this means that incorporating such a constraint in the Z-MIP formulation is equivalent to adding a cut that violates the integrality of the extreme points of the feasible region. Therefore, we propose an alternative approach for solving the problem to optimality, by using an efficient algorithm (polynomial in the number of agents). The algorithm is based on iteratively

solving the relaxed Z-MIP, which is an LP. We summarize this result in the following theorem.

**Theorem 3.3.** *The optimal solution of the Z-MIP problem for the case of a single uniform price can be obtained efficiently (polynomial in the number of agents) by applying Algorithm 1.*

---

**Algorithm 1** Procedure for finding the uniform optimal price

---

**Input:**  $c$ ,  $N$  and  $G$

**Procedure**

1. Set the iteration number to  $t = 1$ , solve the relaxed Z-MIP (an LP) and obtain the vector of discriminative prices  $\mathbf{p}^{(1)}$ .
  2. Find the minimal discriminative price defined as  $p_{min}^{(t)} = \max \left\{ c, \min_{i \in \mathcal{I}} p_i^{(t)} \right\}$  and evaluate the objective function  $\Pi^{(t)}$  with  $p_i = p_{min}^{(t)} \forall i \in \mathcal{I}$  using formula (B.0.12).
  3. Remove all the nodes that receive prices less than or equal to the minimal discriminative price from the network (including all their edges). If there are no more agents in the network, go to step 5. If not, go to step 4.
  4. Re-solve the relaxed Z-MIP for the new reduced network and denote the output by  $\mathbf{p}^{(t+1)}$ . Set  $t := t + 1$  and go to step 2.
  5. The optimal uniform price is equal to  $p_{min}^{(\hat{t})}$ , where  $\hat{t} = \arg \max \Pi^{(t)}$  i.e., the price that yields the larger profits.
- 

We show the termination in finite time of Algorithm 1 and prove its correctness in the Appendix. At a high level, the procedure in Algorithm 1 iteratively reduces the size of the network by eliminating agents with low valuations (at least one per iteration). As a result, it suffices to consider only a finite selection of prices (at least as high as cost) to identify the optimal uniform price.

## 3.6 Extension to uncertain valuations

In this section, we present some analysis and insights for the case where the value interaction terms are uncertain. We would like to compare settings where the prices  $p_i^*$  are set according to the lowest realizations of uncertainties (a safe guaranteed profit) as opposed to the mean of the uncertainties. It is clear that the answer depends on the distributional assumptions as well as on the network topology. A stochastic pricing approach can result in the seller setting riskier prices (e.g., the prices based on the mean of the uncertainties). Below, we focus on presenting such insights under simple symmetric topologies with uncertain influence of neighbors.

The main drawback in the stochastic pricing approach as we will see is that the method is hard, let alone tractable, even with an independence assumption under the discriminative price setting. In Section 3.5.1 we show that the discriminative pricing problem Z-MIP under deterministic value interaction terms can be solved efficiently as an LP. As a result, one can apply techniques from robust optimization to this MIP (see e.g., [7]) when the value interaction terms are modeled using correlated uncertainty sets. We refer the reader to [4] for examples of high dimensional correlated uncertainty sets. This direction of analysis is possible because of an optimization formulation such as Z-MIP but is beyond the scope of this work and hence not pursued.

### Insights under stochastic influence

In this section, we focus on the pairwise influence model in (3.2.5) that captures the influence of individual neighbors only (i.e.,  $K = 2$ ). Our goal is to study the setting where the influences of neighbors are not exactly known and are modeled as random variables. For simplicity and to keep the analysis tractable, we consider the following setting in the remaining of this section.

*Assumption 6.* We assume the following.

- a. The network is composed of  $N$  symmetric agents with a deterministic and symmetric topology (edges). In particular, each agent has  $L$  neighbors.

- b. The self valuations are deterministic, i.e.,  $g_i = \theta \geq 0 \forall i \in \mathcal{I}$ .
  
- c. The pairwise influence parameters  $g_{ij} = g \forall i, j (\neq i) \in \mathcal{I}$  are stochastic and are independent and identically distributed (i.i.d) in  $[g_1, g_2]$  with CDF  $F(\cdot)$  and PDF  $f(\cdot)$  (where  $g_2 > g_1 \geq 0$ ).

In this symmetric setting, without loss of generality the seller offers a uniform price  $p$  that is the same for all the agents. Using the structure in (3.5.1), the price set by the seller can be written as  $p = \max\{\theta + Lg, c\}$ , for some value of  $g$  such that  $g_1 \leq g \leq g_2$ . Suppose the seller sets the price  $p = \theta + Lg_1$ , it is clear that all the agents will buy and the profit of the seller is equal to  $N(\theta + Lg_1 - c)$ . The other extreme case is when the price is set to  $p = \theta + Lg_2$ . In this case, all the agents do not buy with probability 1, and the expected profit of the seller is equal to zero. In the sequel, we consider the case where  $g_1 \leq g^* \leq g_2$ . More precisely, we assume that the seller sets a price given by:

$$p^* = \theta + Lg^*, \tag{3.6.1}$$

and our goal is to find the optimal value of  $p^*$  (equivalently  $g^*$ ) so as to maximize the expected profit. To avoid trivial solutions, without loss of generality we assume that  $\theta + Lg_1 \geq c$ . We refer to the case  $g^* = g_1$  as the safe price (i.e., the seller is not taking any risk). In addition, a higher  $g^*$ , refers to a larger risk taken by the seller. This measure of risk instead of a larger price enables an easy and fair comparison between symmetric graphs with different values of  $N$  and  $L$ .

The expected profit at the optimal price can be written as follows:

$$\Pi(g^*) = (\theta + Lg^* - c)D(\theta, g^*) \quad \text{where} \quad (3.6.2)$$

$$D(\theta, g^*) = \text{Expected demand at the offer price } p^* = \theta + Lg^* \quad (3.6.3)$$

$$= \sum_{k=1}^N k \sum_{\substack{S_k \subset \mathcal{I} \\ |S_k|=k}} \mathbb{P}(\text{All agents in } S_k \text{ buy and others do not buy at } p^* = \theta + Lg^*) \quad (3.6.4)$$

$$= \sum_{k=1}^N k \sum_{\substack{S_k \subset \mathcal{I} \\ |S_k|=k}} \left[ \prod_{i \in S_k} \mathbb{P} \left( \sum_{m \in T_i \cap S_k} g_{mi} \geq Lg^* \right) \prod_{i \in \mathcal{I} \setminus S_k} \mathbb{P} \left( \sum_{m \in T_i \cap S_k} g_{mi} < Lg^* \right) \right] \quad (3.6.5)$$

$$= D(g^*) \quad (\text{i.e., independent of } \theta), \quad (3.6.6)$$

where  $T_i$  represents the  $L$  influencers of agent  $i$ . As a result, the expected profit is given by:  $\Pi(g^*) = (\theta + Lg^* - c)D(g^*)$ . The optimal  $g^*$  can be obtained by the first order condition as follows:

$$\theta + Lg^* - c = -\frac{D(g^*)}{D'(g^*)}. \quad (3.6.7)$$

(3.6.7) is a non-linear fixed point equation in a single dimension. The optimal  $g^*$  can be solved with the knowledge of the network topology and the distribution information. We next impose the following assumption about  $D(g^*)$ .

*Assumption 7.*  $\frac{D(g^*)}{D'(g^*)}$  is an increasing function of  $g^*$ .

This assumption guarantees that (3.6.7) has a unique solution. We will see below that in special cases, the above ratio is related to the inverse hazard rate which satisfies the above assumption for IFR distributions. Also, note that because every term in the function  $D(g)$  depends only on the incoming edges, the optimal price does not depend on the exact network topology of out-going edges. More precisely, the optimal price given in (3.6.1) is identical for any symmetric topology with the same

values of  $N$  and  $L$ .

**Proposition 3.2.** *The following properties hold in general.*

- a. *For sufficiently large  $L$  and  $N$  (i.e.,  $N, L \rightarrow \infty$ ), the optimal price converges to:  $p^* = \max\{\theta + L\mathbb{E}(g), c\}$ .*
- b. *The optimal price  $p^*$  increases with the number of neighbors  $L$ .*
- c. *The optimal price  $p^*$  decreases with the self valuation  $\theta$  and converges to the lowest realization  $g_1$  when  $\theta \rightarrow \infty$ .*

The first part offers the following interesting insight. For very large networks with a large number of edges, the optimal price converges to the solution of the deterministic problem with the random valuations replaced by their mean. As a result, one can solve the deterministic version by using the results of Section 5.4. In this case, the seller is willing to take some risk, exploiting the fact that the overall uncertainty from the different neighbors averages out. Note that in various practical social networks, the values of  $L$  and  $N$  are very large and therefore this scenario may be relevant in practice. The second part shows that even in a setting with stochastic influence, the optimal price chosen by the seller grows with the number of neighbors. The last part is easy to deduce from the fixed point equation in (3.6.7) when the RHS is decreasing in  $g^*$  (which is ensured by Assumption 7). As a result, the seller should take less risk as the self valuation increases, as expected.

We next consider a symmetric directed cyclic network (i.e.,  $L = 1$ ). It allows us to solve the problem, relate the optimal solution to the inverse hazard rate and draw additional insights. In this case, if one agent does not buy, no one buys and hence  $D(g^*) = N [1 - F(g^*)]^N$ . Therefore, (3.6.7) simplifies to:

$$\theta + g^* - c = \frac{1 - F(g^*)}{Nf(g^*)}. \quad (3.6.8)$$

Observe that the optimal price is expressed as an inverse hazard rate normalized by the number of agents  $N$  (RHS of (3.6.8)). For example, when the influence factor  $g$

is uniformly distributed, we obtain  $g^* = \frac{g_2 - N(\theta - c)}{N+1}$ . Recall that the expression of  $g^*$  in (3.6.8) is projected onto  $[g_1, g_2]$  and lies in the boundaries otherwise. We next make the following observation based on equation (3.6.8).

**Observation 3.4.** For  $L=1$ , the optimal  $g^*$  converges to  $g_1$  as  $N \rightarrow \infty$ .

This means that for a very large cyclic network, the retailer should not take any risk by setting the safe price. This makes sense as the probability that at least one agent does not buy increases and approaches 1 as  $N$  becomes very large, no matter what is the price. Consequently, the larger the number of agents, the less risk the retailer should take in order to make sure to capture the buyers and attain a positive profit.

### 3.7 Price-incentives to guarantee influence

Thus far, we have assumed that consumers always influence their peers as long as they purchase the item. This assumption is not realistic in many practical settings. Indeed, after purchasing an item, it is sometimes not entirely natural to influence friends about the product unless one takes some effort to do so. This, for example, could be by writing a review, endorsing the item on their social media page (e.g., Facebook wall), or at the very least announcing the purchase. However, to the best of our knowledge, most previous works impose such an assumption that purchasing is equivalent to influencing with the exception of [2]. In the latter, the authors study a cash back setting where the seller offers an exogenous specified uniform cash reward to any recommender if he influenced at least one of his neighbors and they buy the item. We consider a variant of this model but study it in the context of purchasing equilibrium and the optimization framework we have proposed in this work.

Consider a setting where the seller offers both a price and a discount (also referred to as an incentive) to each agent in the network. Each agent can then decide whether to buy the item or not. If the agent decides to purchase the item, he can claim a

fraction of the discount offered by the seller in return for influence actions. These can include liking the product or a wall post in an online social platform such as Facebook or writing a review so that these actions can be digitally tracked by the seller. The agent receives a small discount in exchange for a simple action such as liking the product and a more significant discount by taking an action such as writing a detailed review. Such incentive mechanisms are commonplace these days. For example, online booking agencies request reviews of booked hotels on their website in return for certain loyalty benefits. Using such a model, the seller can now ensure the influence among the agents so that the network externalities effects are guaranteed to occur. In particular, the profits obtained through the optimization are guaranteed for the seller since each agent claims the discounted price as soon as the influence action is taken. In the previous setting where externalities are assumed to always occur, the actual profits may be far from the value predicted by the optimization. In fact, we show using a computational example in Section 3.8 that even if a single agent does not impart his influence, it can significantly reduce the total profits of the seller. We now extend our model and results to this more general setting where the seller can design price-incentives to guarantee social influence.

For simplicity, we focus on the utility model from (3.2.5) that captures the individual influence of neighbors only (i.e.,  $K = 2$ ). We consider a model with a continuum of actions to influence ones' neighbors. Let  $t_i \geq 0$  denote the utility equivalent of the maximal effort needed by agent  $i$  to claim the entire discount offered by the seller. If agent  $i$  decides to purchase the item, we assume that  $\gamma_i t_i$  is the effort required by agent  $i$  to claim a fraction  $\gamma_i$  of the discount, where  $0 \leq \gamma_i \leq 1$ . We view  $t_i$  as the *influence cost* for agent  $i$  and the variable  $\gamma_i$  as the *influence intensity* chosen by agent  $i$ . The parameter  $t_i$  can be estimated from historical data using the intensity of online activity for past purchases, the number of reviews written, the corresponding incentives needed and data from cookies. For a given set of prices  $\mathbf{p}$  and discounts  $\mathbf{d}$

chosen by the seller, we extend the utility function of agent  $i$  in (3.2.8) as follows:

$$u_i(\alpha_i, \gamma_i, \boldsymbol{\alpha}_{-i}, \boldsymbol{\gamma}_{-i}, p_i, d_i) = \alpha_i \left( g_i + \sum_{j \in \mathcal{I} \setminus \{i\}} \gamma_j g_{ji} - p_i \right) + \gamma_i (d_i - t_i), \quad (3.7.1)$$

where  $\gamma_i \leq \alpha_i$  and  $\alpha_i$  is the binary purchasing decision of agent  $i$ . So, if agent  $i$  does not purchase the item,  $\alpha_i = 0$  and  $\gamma_i = 0$  as well. In other words, the constraint  $\gamma_i \leq \alpha_i$  captures the fact that only buyers can influence friends. But if agent  $i$  purchases the item, then  $\alpha_i = 1$  and  $\gamma_i$  can be any number in  $[0, 1]$  as chosen by agent  $i$ . Here,  $\boldsymbol{\alpha}_{-i}$  and  $\boldsymbol{\gamma}_{-i}$  are the decisions of all the other agents but  $i$ . Similarly to problem 3.2.9, the utility maximization problem for agent  $i$  can be written as follows:

$$\begin{aligned} \max_{\alpha_i, \gamma_i} \quad & u_i(\alpha_i, \gamma_i, \boldsymbol{\alpha}_{-i}, \boldsymbol{\gamma}_{-i}, p_i, d_i) & (3.7.2) \\ \text{s.t.} \quad & 0 \leq \gamma_i \leq \alpha_i \\ & \alpha_i \in \{0, 1\} \end{aligned}$$

In a similar way as problem 3.2.10, the seller's profit maximization problem can be written as:

$$\max_{\mathbf{p}, \mathbf{d} \in \mathbf{P}} \sum_{i \in \mathcal{I}} \left[ \alpha_i (p_i - c) - \gamma_i d_i \right]. \quad (3.7.3)$$

Here, the decision variables of the seller are  $\mathbf{p}$  and  $\mathbf{d}$  which are two vector of prices and discounts with an element for each agent in the network. These vectors can be chosen according to different pricing strategies. For example, one can consider a fully discriminative or a fully uniform pricing strategy or more generally, an hybrid model where the regular price is uniform across the network ( $p_i = p_j$ ) but the discounts are tailored to the various agents. This hybrid setting corresponds to a common practice of online sellers that offer a standard posted price for the item but design personalized discounts for different classes of customers that are sent via e-mail coupons. Finally, similarly to the previous setting, one can incorporate various polyhedral business rules

on prices, discounts and constraints on network segmentation. The variables  $\alpha_i$  and  $\gamma_i$  are decided according to each agent's utility maximization problem given in (3.7.2). If agent  $i$  decides to buy the product, then the seller incurs a profit of  $p_i - \gamma_i d_i - c$ .

We note that in the special case where  $\alpha_i = \gamma_i$  and  $t_i = 0 \forall i \in \mathcal{I}$ , we recover the previous model where the seller offers a single price to each agent and any buyer is assumed to always influence his peers. In addition, by adding the constraint  $\gamma_i \in \{0, 1\}$  we have an interesting setting where each agent can only buy at two different prices: a full price  $p_i$  that does not require any action and a discounted price  $p_i - d_i$  that requires some action to influence. Note that one can easily extend the model in this section to more than two prices so as to incorporate a finite but discrete set of different actions specified by the seller.

Our goal is to extend the results from previous sections for this new setting with incentives to guarantee influence. We begin by studying the purchasing equilibria of the second stage game. By using a similar methodology as in Section 3.3, one can show that for any given prices and discounts there exists a PNE for the second stage game.

**Theorem 3.4.** *The second stage game has at least one pure Nash equilibrium for any given vector of prices  $\mathbf{p}$  and discounts  $\mathbf{d}$  chosen by the seller. A small perturbation in prices and discounts to break ties between buying and not buying as well as influencing and not influencing results in a Pareto optimal PNE that is preferred by both the seller and the network of buyers.*

The proof of Theorem 3.4 is in a similar spirit of the best response dynamics in Theorem 3.1 and is not presented due to space limitations. In this case, a PNE is defined where purchasing decisions  $\alpha_i$  are strictly 0 or 1. However, one can also note that there always exists an equilibrium for which the variables  $\gamma_i$  are all integer as well. More precisely, if  $d_i - t_i > 0$  (remember that the prices and discounts are given),  $\gamma_i$  can be increased to 1 and otherwise  $\gamma_i = 0$ . We therefore have the existence of a PNE with  $\gamma_i$  integer as well.

One can see that a result similar to Observation 3.1 still holds and therefore one can characterize the equilibria (mixed and pure) as a set of constraints where the binary variables are relaxed to be continuous. In this case, one can transform subproblem 3.7.2 of agent  $i$  to a set of feasibility constraints using duality theory as follows:

$$\text{Primal feasibility:} \quad 0 \leq \alpha_i \leq 1 \quad (3.7.4)$$

$$0 \leq \gamma_i \leq \alpha_i \quad (3.7.5)$$

$$\text{Dual feasibility:} \quad y_i - w_i \geq g_i + \sum_{j \in \mathcal{I} \setminus \{i\}} \gamma_j g_{ji} - p_i \quad (3.7.6)$$

$$w_i \geq d_i - t_i \quad (3.7.7)$$

$$y_i, w_i \geq 0 \quad (3.7.8)$$

$$\text{Strong duality:} \quad y_i = \alpha_i \left( g_i + \sum_{j \in \mathcal{I} \setminus \{i\}} \gamma_j g_{ji} - p_i \right) + \gamma_i (d_i - t_i) \quad (3.7.9)$$

We now have two continuous dual variables  $y_i$  and  $w_i$ , together with two dual feasibility constraints for each agent  $i$ . Similar to the earlier setting, in order to restrict to the pure Nash equilibria (that is necessary for the problem of optimal pricing), we need to impose  $\alpha_i$  to be binary variables for all agents  $i \in \mathcal{I}$ . We can then formulate the optimal pricing problem faced by the seller, similar to problem Z, that maximizes the profits given in (3.7.3) with the equilibrium constraints (3.7.4)-(3.7.9), where the constraints on  $\alpha_i$  are replaced by the binary versions as follows:

$$\begin{aligned} \max_{\substack{\mathbf{p}, \mathbf{d} \in \mathbf{P} \\ \mathbf{y}, \mathbf{w}, \boldsymbol{\alpha}, \boldsymbol{\gamma}}} \quad & \sum_{i \in \mathcal{I}} \left[ \alpha_i (p_i - c) - \gamma_i d_i \right] & (Zi) \\ \text{s.t.} \quad & \text{constraints (3.7.5) - (3.7.9), } \alpha_i \in \{0, 1\} \quad \forall i \in \mathcal{I} \end{aligned}$$

We denote this problem by  $Zi$  where  $i$  represents the model with incentives to guarantee influence of this present section. It is easy to make the following observation.

**Observation 3.5.** Every optimal solution of problem  $Zi$  satisfies  $d_i \leq t_i$ .

Indeed, the seller can always reduce  $d_i$  to be equal to  $t_i$  while maintaining feasibility and strictly increasing the objective function. This implies that the constraint (3.7.7) is redundant in the optimal pricing problem. Consequently and by using the constraints (3.7.6–3.7.8), one can always assign  $w_i = 0$  in the pricing problem while maintaining feasibility and without altering the objective function. This observation allows us to simplify problem  $Z_i$  by removing all the dual variables  $w_i \forall i \in \mathcal{I}$ . We next extend Proposition 3.1 for this setting.

**Proposition 3.3.** *Problem  $Z_i$  admits a tight continuous relaxation. Moreover, there always exists an optimal solution to problem  $Z_i$  where all the variables  $\gamma$ 's are integer as well.*

The second result in this Proposition is interesting because it implies that even though the seller allows for a continuum of influence actions, the buyer would either fully influence or not influence at all. As a result, this is equivalent to the setting where the seller offers only two options: a full price  $p_i$  and a discounted price  $p_i - d_i$  in exchange of a specific action to influence.

Problem  $Z_i$  has non-linearities of the form  $\alpha_i \gamma_j$ ,  $\alpha_i p_i$  and  $\gamma_i d_i$ . Using the discrete nature of the variables  $\alpha_i$  and  $\gamma_i$  from Proposition 3.3, one can transform problem  $Z_i$

to the following MIP formulation, denoted by Zi-MIP:

$$\max_{\substack{\mathbf{p}, \mathbf{d} \in \mathbf{P} \\ \mathbf{y}, \mathbf{z}, \mathbf{z}^d, \mathbf{x}, \alpha, \gamma}} \sum_{i \in \mathcal{I}} (z_i - z_i^d - c\alpha_i) \quad (\text{Zi-MIP})$$

s.t.

$$\left. \begin{aligned} y_i &= \left( \alpha_i g_i + \sum_{j \in \mathcal{I} \setminus \{i\}} x_{ji} g_{ji} - z_i \right) + (z_i^d - \gamma_i t_i) \\ y_i &\geq g_i + \sum_{j \in \mathcal{I} \setminus \{i\}} \gamma_j g_{ji} - p_i \\ \gamma_i &\leq \alpha_i \\ y_i &\geq 0 \end{aligned} \right\} \forall i \in \mathcal{I} \quad (3.7.10)$$

$$\left. \begin{aligned} z_i, z_i^d &\geq 0 \\ z_i &\leq p_i \\ z_i &\leq \alpha_i p^{max} \\ z_i &\geq p_i - (1 - \alpha_i) p^{max} \\ z_i^d &\leq d_i \\ z_i^d &\leq \gamma_i p^{max} \\ z_i^d &\geq d_i - (1 - \gamma_i) p^{max} \end{aligned} \right\} \forall i \in \mathcal{I} \quad (3.7.11)$$

$$\left. \begin{aligned} x_{ji} &\geq 0 \\ x_{ji} &\leq \alpha_i \\ x_{ji} &\leq \gamma_j \\ x_{ji} &\geq \alpha_i + \gamma_j - 1 \end{aligned} \right\} \forall i \neq j \in \mathcal{I} \quad (3.7.12)$$

$$\alpha_i, \gamma_i \in \{0, 1\} \quad \forall i \in \mathcal{I} \quad (3.7.13)$$

where  $p^{max}$  is the maximum price allowed. Note that we removed the dual variables  $w_i$  by using Observation 3.5. We conclude that the problem of designing prices and

incentives for selling an indivisible item to agents embedded in a social network can be formulated as a MIP where one can incorporate business rules on prices and on constraints on network segmentation. However, solving a MIP may not be very scalable. For the case of discriminative prices and discounts, i.e., when  $\mathbf{P} = \mathbb{R}_+^N \times \mathbb{R}_+^N$ , we are able to retrieve a similar result as Theorem 3.2. The result is summarized in the following Theorem.

**Theorem 3.5.** *The optimal discriminative pricing solution of the Zi-MIP problem can be obtained efficiently (polynomial in the number of agents). In particular, problem Zi-MIP admits a tight LP relaxation.*

The proof is in a similar spirit as Theorem 3.2 and is omitted due to space limitations. However, the main idea can be folded into the following two steps. First, fix the values of  $\gamma_i, z_i^d$  and proceed in the same fashion as in Theorem 3.2 to show how to construct a solution with  $\alpha_i$  integer  $\forall i \in \mathcal{I}$ . Next, with the integer values of  $\boldsymbol{\alpha}$  obtained from the previous step, one can show that the objective does not drop by modifying any component of  $\boldsymbol{\gamma}$  to 1 by appropriately altering the prices of the neighbors so that their actions do not change as in Proposition 3.3.

In comparison to problem Z-MIP with a single price for each agent, problem Zi-MIP yields potentially lower profits for the seller. However, these profits are guaranteed whereas in the previous case, the estimated profits can be far from the actual values if people fail to influence their neighbors (i.e., the model is mis-specified). The difference in profits between both settings can be viewed as the price the seller has to pay to guarantee the influence between agents in the network and can be computed efficiently by solving both settings.

An interesting observation that we see throughout this section is that even though our model allows a continuum of influence actions for every agent, the optimal prices for any agent can be designed in such a way that only two price options suffice. More specifically, the two options for any agent are a full price with no action required and a discounted price which requires an influence action in return.

### 3.8 Computational experiments

In this section, we present computational experiments on simple example networks to draw qualitative insights about incorporating social interactions, comparing various pricing strategies including the richer model with incentives developed in this work. We consider a network with  $N = 10$  agents.

**Value of incorporating network externalities:** In Fig. 3-1, we plot the optimal prices offered by the seller to the different agents under the discriminative and uniform pricing strategies, both with and without social interactions. The circles around the markers, whenever present, depict the fact that the agent decided not to purchase the item at the offered price (agents 7, 8 and 9 for uniform price with network externalities). In this instance, each agent is randomly connected to three other agents with  $g_{j,i} = 1.25$  for any connected edge,  $g_i = 3$  and  $c = 2$ .

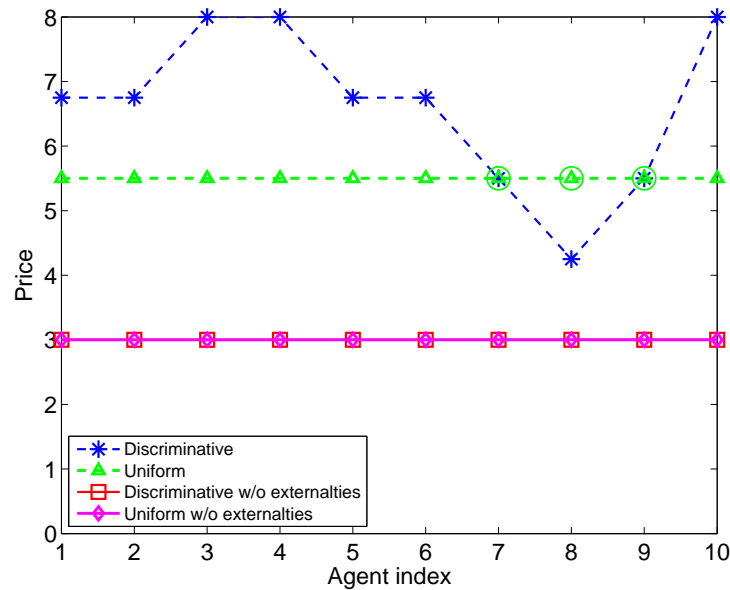


Figure 3-1: Value of incorporating network externalities for the discriminative and uniform pricing strategies

We observe that by incorporating the positive externalities between the agents, the seller earns higher profits. In this particular example, the total profits are equal to 46.25 (discriminative prices) and 24.5 (uniform price) for the case with network

externalities. In the case without network externalities, the profits in both uniform and discriminative prices are equal to 10. This result is expected because every agent's willingness-to-pay increases as their neighbors positively influence them. The seller can therefore charge higher prices and increase his profits. Fig. 3-1 also shows the added benefit by using a discriminative pricing strategy compared to a uniform single price. When the firm has the additional flexibility to price discriminate and offers a different price to each agent in the network, the total profits can increase significantly. In the example above, only one agent is offered a price that is lower than the optimal uniform price.

**Pricing an influencer:** In Fig. 3-2, we present an example where it is beneficial for the seller to earn negative profit ( $p_i < c$ ) on some influential agent  $i$  in order to extract significant positive profits on his neighbors. In particular, we consider a network where agent 5 is a very influential player with  $g_{5,5}$  being very low (0.075) while  $g_{5,j}$  is sufficiently high (1.38) for the four agents that he influences. Here,  $g_{i,j} = 0.75$  for any other connected edge,  $g_i = 1.5R \forall i \neq 5$  where  $R = U[1, 2]$  and  $c = 2$ . The

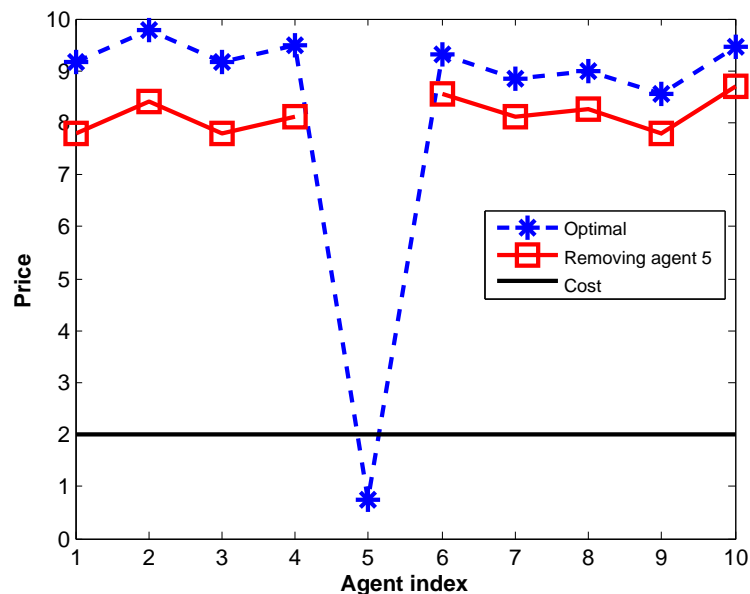


Figure 3-2: Centrality effect: losing money on an influential agent

optimal discriminative price vector includes a price for agent 5 that happens to be lower than the cost. This illustrates the fact that agent 5 has a central and influential position in the network and therefore, the seller should strongly incentivize this player. In particular, the optimal algorithm identifies this feature and captures the fact that it is profitable to offer a very low price to this person so that he can influence other people about the product. This way, the seller loses some small amount of money on the influential agent but is able to extract higher profits on his neighbors. We now compare this to an alternate strategy where the seller decides to remove agent 5 from the network due to his low valuation (this is equivalent to offer a very large price to agent 5). In this case, all the optimal prices are decreased and the overall profit drops from 63.52 to 55.5 units so that one can increase profits by about 14.5% by including player 5.

**Value of incorporating incentives that guarantee influence:** In Fig. 3-3, we compare the optimal solution for discriminative prices to the extended model introduced in Section 3.7 where the seller offers a uniform regular price ( $p = 4$ ) and designs discriminative discounts in exchange of some action to influence. If the seller does not incentivize the agents to influence, some of them would not influence. The goal of this experiment is to understand the impact of seller’s incentives without which some agents would not influence. In this instance, every agent is randomly connected to three other agents with  $g_{j,i} = 0.75$  for any connected edge and  $g_i = 1.5R$  where  $R = U[1, 4.5]$ . We assume  $t_i = U[0, 1] \forall i \neq 1$ ,  $t_1 = 6.9$  and  $c = 1$ .

We observe that the total profit using the earlier model (without incentives to influence i.e.,  $t_i = 0$ ) is equal to 27.15. This profit is not guaranteed because some agents may not influence their peers (i.e., the model might be mis-specified). In particular, in this example, suppose agents 5 and 10 who buy at full price do not influence their neighbors which includes agent 1. Agent 1 ends up not purchasing the item and consequently does not influence his neighbors either. Finally, it so happens that only agents 2, 5 and 10 buy the item yielding a profit of 9 as opposed to 27.15. Consequently, the earlier model predicts a value for the profits that is significantly

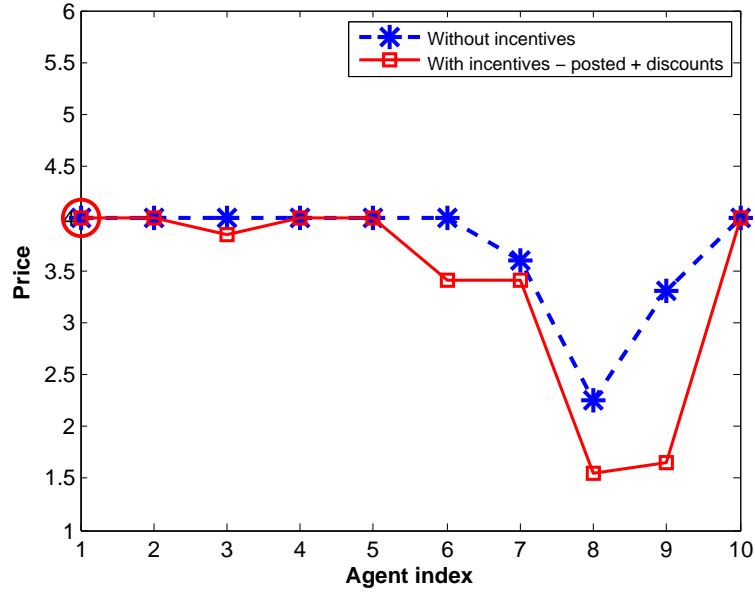


Figure 3-3: Value of incorporating incentives that guarantee influence

higher than the realized one even if a few agents do not influence. On the other hand, in the model with incentives that guarantee influence ( $t_i$  is taken into account), the total profits are equal to 20.85 and agent 1 does not purchase the item and agents 5 and 10 do not influence anyone but the other agents do. Observe that this is lower than 27.15 but significantly larger than 9. Therefore, the model with incentives provides the seller with the flexibility of using prices together with incentives that result in a higher degree of confidence on the predicted profits.

**Symmetric agents with asymmetric incentives:** In Fig. 3-4, we present a setting with symmetric agents who receive asymmetric incentives to influence their neighbors. In this instance, every agent not only has the same number of neighbors but also the same self and cross valuations. In particular, we consider a complete graph with  $g_i = 1.3$  and  $g_{i,j} = 0.3$ , a cost to influence  $t_i = 2.2$  and  $c = 0.2$ . We also assume that the item has a posted price equal to 3. We compute the optimal discriminative prices which happen to be at 3 for everyone and compare them to the case where the seller designs incentives to guarantee influence by offering two prices using problem  $Z_i$ -MIP.

Interestingly, the optimal solution for the model with incentives is not symmetric despite the fact that all the agents are homogenous. Indeed, it is sufficient for the seller to incentivize any six out of the ten agents in the network (no matter which group of six). These six agents receive a targeted discount to influence their peers that purchase at the full posted price.

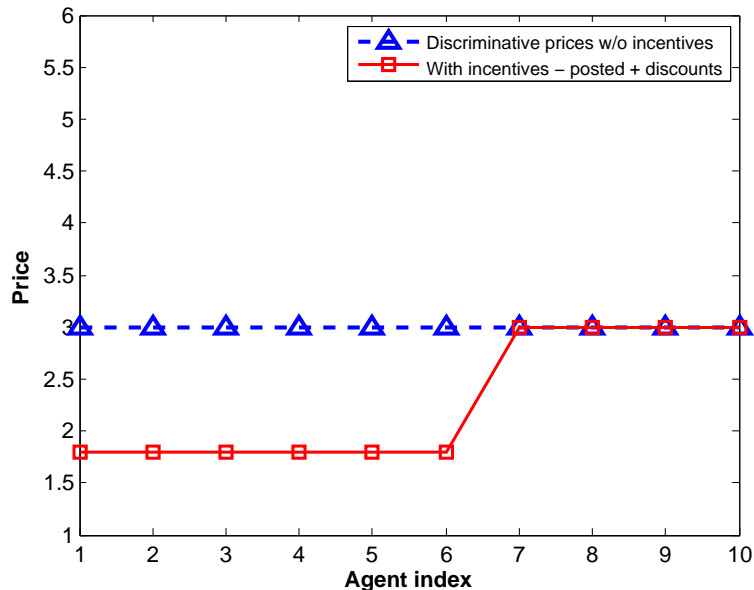


Figure 3-4: Symmetric Graph with asymmetric incentives: with and without incentives

**Effect of network topology on optimal prices:** In Fig. 3-5, we consider different network topologies and compare the optimal discriminative prices as well as the corresponding profits. In all the scenarios,  $g_i = 1.5R$  where  $R = U[1, 2]$ ,  $g_{i,j} = 0.75$  when agent  $i$  influences agent  $j$  and 0 otherwise and  $c = 2$ . For each network topology, we solve the optimal discriminative prices using the relaxation of Z-MIP. We plot the optimal price vector for the different networks in Fig. 3-5. We observe that in our example, all the agents always decide to purchase the item. In the complete graph, all the nodes are connected to each other and therefore the profits are the highest and equal to 70.15. In the intermediate topology where each agent has three neighbors,

the total profits are equal to 22.45. The cycle graph is a network where the nodes are connected in a circular fashion, where each agent has one ingoing and one outgoing edge (influences one agent and influenced by one). In this case, the total profits are equal to 8.95. Star 1 and star 2 are star graphs with a central agent being agent 5. In star 1, agent 5 influences all the other agents and in star 2 agent 5 is influenced by all the others. In both cases, the profits are equal to 8.2. This is interesting to observe that both star networks yield the same profits as the total valuations in the system are the same. In star1, agent 5 receives a small discount to influence so that the prices of the others are slightly higher. In star 2, the prices of all the agents but 5 are slightly lower so that the seller can charge a high price to agent 5. As we observe the prices for the different network topologies, we note that the value of the prices and the profits increase with the number of edges in the graph. Indeed, each additional edge corresponds to an agent increasing another agent's willingness-to-pay and therefore the more the graph is connected, the larger are the profits.

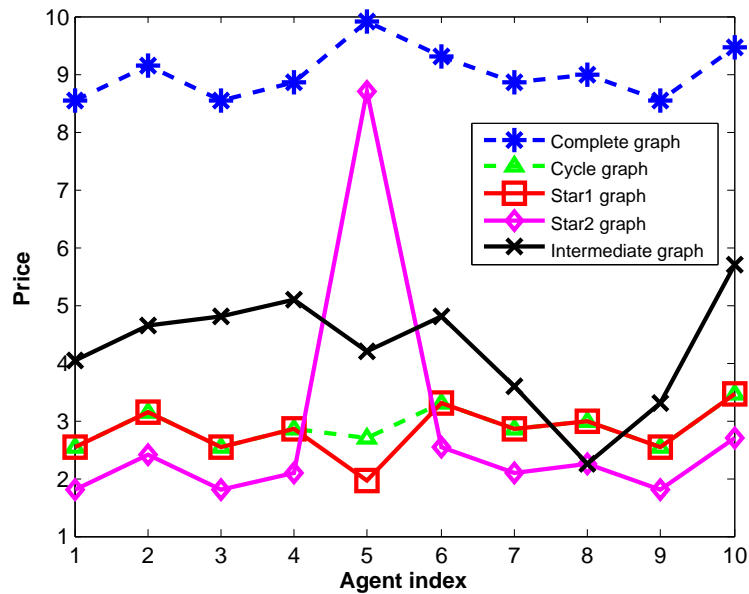


Figure 3-5: Optimal prices for various network topologies

## 3.9 Conclusions

In this chapter, we presented an optimal pricing model for a profit maximizing firm that sells an indivisible item to agents embedded in a social network. We assumed that the agents interact and positively influence each others' purchasing decisions. The local influence structures can be quite complex and are modeled using non-linear functions that capture submodular or supermodular effects in the number of friends influencing a agent. We model the problem as a two stage game where the seller first offers prices and the agents collectively follow with their purchasing decisions by taking into account their neighbors influences. Using equilibrium existential properties, linear programming duality theory and techniques from integer programming, we reformulate the two stage pricing problem as a MIP formulation with linear constraints. We view this MIP as an operational pricing tool that any firm can use by incorporating various business rules on prices and constraints on network segmentation. This allows us to cast the problem into the traditional optimization framework, where one can explore and exploit various advancements in optimization techniques to solve the problem efficiently. For the case of discriminative and uniform pricing strategies, we present efficient methods to optimally solve the MIP that are polynomial in the number of agents using its LP relaxation. We observe that price of an agent that buys in the optimal discriminative pricing solution is the sum of the agent's own value for the item and a markup term that corresponds to the influence on this agent by the network of agents that buy. The seller offers this buy price to an agent depending on the agent's own valuation for the item and/or the influence the agent exerts on the network. The gain from network externalities in essence comes from two types of customers: high valued customers who influence their neighbors but also low valued customers who are influential. We also present some analysis and insights for the case where the value interaction terms are uncertain. In particular, we show that for large scale networks with a very large number of edges, the optimal price converges to the deterministic problem with expected valuations.

We extend our proposed model and results to the case when the seller can design both prices and incentives to guarantee influence amongst agents. This extension is important because in general, agents that buy need not necessarily influence their peers. The seller can use incentives in exchange for an action such as an endorsement, a wall post or a review to guarantee influence. Finally using computational experiments, we highlight the benefits of incorporating network externalities, compare the different pricing strategies and the more general model with incentives. In particular, we experimentally see how sometimes it is beneficial for the seller to earn negative profit on an influential agent in order to extract significant positive profits on others.

As a part of future work, it would be interesting to study (both analytically and computationally) the difference in profits between the models with and without incentives depending on the input parameters. The optimization framework for optimal pricing presented in this chapter allows one to explore decomposition techniques for other complex pricing strategies, and stochastic and robust optimization methods to handle partially observable noisy social network data.

## Bibliography

- [1] H. Akhlaghpour, M. Ghodsi, N. Haghpanah, V. Mirrokni, H. Mahini, and A. Nikzad. Optimal iterative pricing over social networks. *WINE. Lecture Notes in Computer Science*, 6484:415–423, 2010.
- [2] D. Arthur, R. Motwani, S. Aneesh, and Y. Xu. Pricing strategies for viral marketing on social networks. In *The 5th International Workshop on Internet and Network Economics*, pages 101–112, 2009.
- [3] Coralio Ballester, Antoni Calvó-Armengol, and Yves Zenou. Who’s who in networks. wanted: the key player. *Econometrica*, 74(5):1403–1417, 2006.
- [4] Chaithanya Bandi and Dimitris Bertsimas. Tractable stochastic analysis in high

- dimensions via robust optimization. *Mathematical programming*, 134(1):23–70, 2012.
- [5] A. Banerji and B. Dutta. Local network externalities and market segmentation. *International Journal of Industrial Organization*, 27:605–614, 2009.
- [6] D. Bertsimas and R. Weismantel. *Optimization over integers*, volume 13. Dynamic Ideas Belmont, 2005.
- [7] Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.
- [8] F. Bloch and N. Querou. Pricing in social networks. *Games and economic behavior*, 80:243–261, 2013.
- [9] O. Candogan, K. Bimpikis, and A. Ozdaglar. Optimal pricing in networks with externalities. *Operations Research*, 60:883–905, 2012.
- [10] W. Chen, P. Lu, X. Sun, B. Tang, Y. Wang, and Z. A. Zhu. Optimal pricing in social networks with incomplete information. 2011. The 7th Workshop on Internet & Network Economics.
- [11] P. Domingos and M. Richardson. Mining the network value of customers. *Proceedings of the 7th Conference on Knowledge Discovery and Data Mining*, pages 57–66, 2001.
- [12] J. Farrell and G. Saloner. Standardization, compatibility and innovation. *RAND Journal of Economics*, 16:70–83, 1985.
- [13] D. Fudenberg and J. Tirole. *Game theory*. MIT press, 1991.
- [14] A. Galeotti, S. Goyal, M. Jackson, F. Vega-Redondo, and L. Yariv. Network games. *Review of Economics Studies*, 77(1):218–244, 2010.
- [15] Dilek Gunec and S Raghavan. Integrating social network effects in product design. 2012. Submitted to Management Science.

- [16] J. Hartline, V. S. Mirrokni, and M. Sundararajan. Optimal marketing strategies over social networks. 2008. Proceedings of the 17th International Conference on World Wide Web.
- [17] R. Iyengar, C. Van den Bulte, and T. W. Valente. Opinion leadership and social contagion in new product diffusion. *Marketing Science*, 30(2):195–212, 2011.
- [18] M. Katz and C. Shapiro. Network externalities, competition and compatibility. *American Economic Review*, 75:424–440, 1985.
- [19] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. 2003. Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
- [20] Z. Q. Luo, J. S. Pang, and D. Ralph. *Mathematical programs with equilibrium constraints*. Cambridge University Press, 1996.
- [21] G. L. Nemhauser and L. A. Wolsey. *Integer and combinatorial optimization*, volume 18. Wiley New York, 1988.
- [22] A. Sundararajan. Local network effects and complex network structure. *The BE Journal of Theoretical Economics*, 7(1):Article 46, 2007.



# Chapter 4

## Pricing without knowledge of demand

### 4.1 Introduction

Firms that introduce new products often set price with little or no knowledge of demand, and no data on which to base elasticity estimates. Examples include pharmaceutical companies introducing new types of drugs (Lilly's Prozac in 1987), technology companies introducing new products or services (Apple setting the price of music downloads when launching its iTunes store in 2002, and more recently, pricing its iWatch), or a company introducing an existing product in an emerging market (P&G launching Pampers in China in 1998). Although marginal cost may be easy to estimate (it is close to zero for most drugs and music or software downloads, and can be estimated from experience for diapers), the firms are likely to know little or nothing about the demand curves they face, and may not even be able to estimate arc elasticities. How should firms set prices in such settings?

As discussed below, this problem has been the subject of a variety of studies, most of which focus on experimentation and learning, e.g., setting different prices and observing the outcomes. Experimenting with price, however, is usually not feasible or desirable; it is more commonly the case that firms must choose and maintain a

particular price.<sup>1</sup> We examine a much simpler approach to this pricing problem that involves no experimentation.

We show that in many situations, the firm can use a remarkably simple pricing rule. The use of the rule requires that (i) the firm’s marginal cost is known and constant, (ii) the firm need not predict the initial quantity it will sell, and (iii) the firm can estimate the *maximum price* it can charge and still expect to sell at least some units. (We relax this last assumption in Section 4.3.) Denoting that maximum price by  $P_m$ , the firm sets price as though the actual demand curve were linear, i.e.,

$$P = P_m - bQ . \tag{4.1.1}$$

Assuming marginal cost,  $c$ , is constant, the firm’s quasi-optimal price is  $P^* = (P_m + c)/2$ , which we refer to as the “linear price.” This price is independent of the slope  $b$  of the linear demand curve, although the resulting quantity,  $Q_L^* = (P_m - c)/2b$ , is not. But as long as the firm does not need to invest in production capacity or otherwise plan on a particular sales level (as would be the case for most new drugs, music downloads, or software), knowledge of  $b$ , and thus the ability to predict its sales, is immaterial. More sales will be better than less, but the only problem at hand is to set the price. Denote the resulting price and profit from using eqn. (4.1.1) by  $P^*$  and  $\Pi^*$  respectively.

How well can the firm expect to do if it sets the price  $P^*$ ? Suppose that with precise knowledge of its true demand curve, the firm would have set a different price  $P^{**}$  and earn a (maximum) profit  $\Pi^{**}$ . The question we address is simple: How close can we expect  $\Pi^*$  to be relative to  $\Pi^{**}$ , i.e., how well is the firm likely to do using this simple pricing rule? As we will show, if the true demand curve is one of many commonly used demand functions, or even if it is a more complex function, the firm can expect to do very well.

Determining the maximum price  $P_m$  might not be easy, but it is a much less

---

<sup>1</sup>This has been the case when pharmaceutical companies introduced new drugs, when Apple launched the iTunes store, and when P&G launched Pampers in China.

difficult task than estimating an entire demand curve. A pharmaceutical company might estimate  $P_m$  by comparing a new drug to existing therapies (including non-drug therapies).<sup>2</sup> And when it planned to sell music through iTunes, Apple might have estimated  $P_m$  to be around \$2 or \$3 per song, as a multiple of the per-song price of compact discs.<sup>3</sup>

The basic idea behind this chapter is quite simple, and is illustrated in Figure 4-1. The demand curve labeled “Actual Demand” was drawn so it might apply to a new drug, or to music downloads in the early years of the iTunes store. A pharmaceutical company might estimate a price  $P_m$  at which some doctors will prescribe and some consumers will buy its new drug, even though insurance companies refuse to pay for it. As the price is lowered and the drug receives insurance coverage, the quantity demanded expands considerably. At some point the market saturates so that even if the price is reduced to zero there will be no further increase in sales. For music downloads, at prices above  $P_m$  it is more economical to buy the CD and “rip” the desired songs to one’s computer. At lower prices demand expands rapidly, and at some point the market saturates.

The figure also shows the corresponding marginal revenue curve,  $MR_{\text{ACTUAL}}$ . If the firm knew this curve, it would set the profit-maximizing price  $P^{**}$  and expect to sell the quantity  $Q^{**}$ , at the intersection with marginal cost. But it does not know this curve. A linear demand curve that starts at  $P_m$  and the corresponding marginal revenue curve have also been drawn, and labeled  $D_L$  and  $MR_L$ . Using this linear demand curve implies a profit-maximizing price  $P^*$  and quantity  $Q_L^*$ , where the subscript  $L$  refers to the quantity sold if the linear demand curve were the true demand curve. (Note that the slope of  $D_L$  is immaterial for the pricing decision.) The actual quantity that would be sold given the price  $P^*$  is  $Q^*$  (under the actual demand curve). How badly would the firm do by pricing at  $P^*$  instead of  $P^{**}$ ? For

---

<sup>2</sup>For example, when pricing Prilosec, the first proton-pump inhibitor anti-ulcer drug, Astra-Merck might estimate that  $P_m$  should be two or three times higher than the price of Zantac, an older generation H<sub>2</sub>-antagonist drug.

<sup>3</sup>At the time, a CD containing 10 to 12 songs might cost around \$12 to \$15, but many consumers would want only a few of those songs.

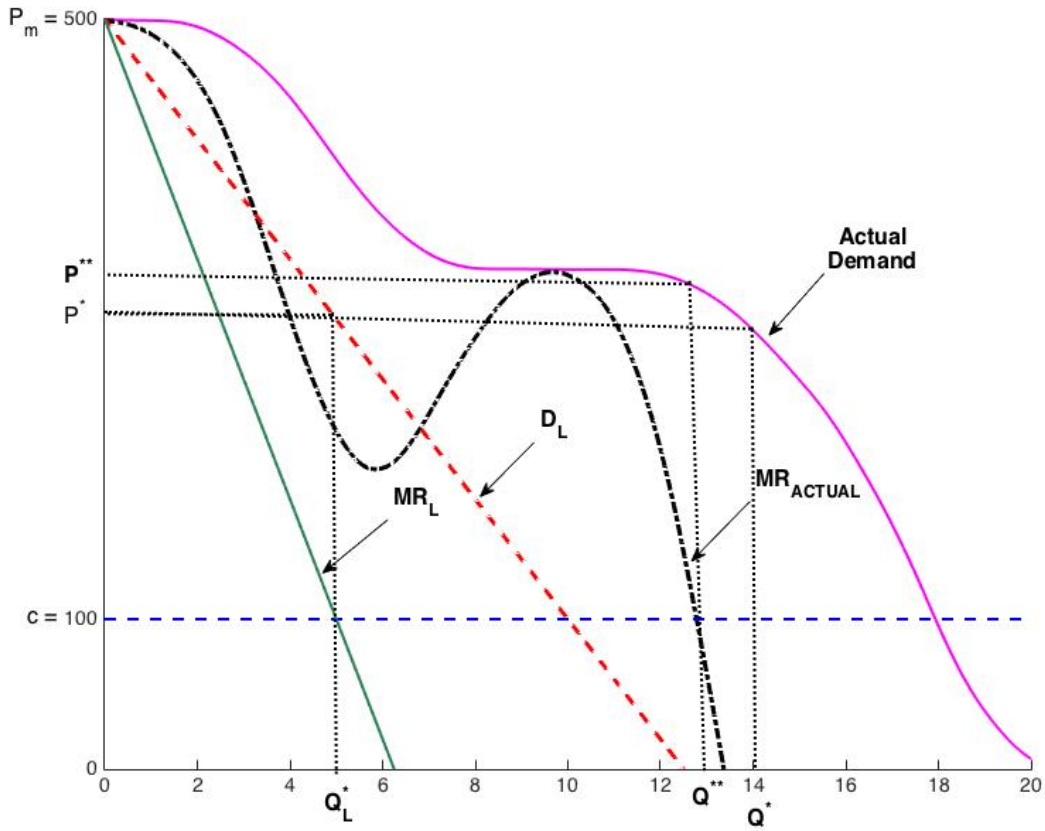


Figure 4-1: Illustration for a representative demand curve

the demand curve and marginal cost shown in Figure 4-1, the profit and price ratios (determined numerically) are  $\Pi^{**}/\Pi^* = 1.023$  and  $P^{**}/P^* = 1.069$ , i.e., the resulting profit is within a few percent of what the firm could earn if knew the actual demand curve and used it to set price. (The firm would do slightly worse if  $c = 0$ , in which case  $\Pi^{**}/\Pi^* = 1.084$ .)

There are certainly demand curves for which this pricing rule will perform very poorly. For example, suppose the true demand curve is a rectangle, i.e.,  $P = P_m$  for  $0 \leq Q \leq Q_m$  and  $P = 0$  for  $Q > Q_m$ . Then the profit-maximizing price is clearly  $P_m$ , and the resulting profit is  $\Pi^{**} = (P_m - c)Q_m$ . Setting a price  $P^* = (P_m + c)/2$  will yield a much lower profit; in fact,  $\Pi^{**}/\Pi^* = 2.0$ . We want to know how well

our pricing rule will perform — i.e., what is  $\Pi^{**}/\Pi^*$  — for alternative “true” demand curves.

There is a large literature on optimal pricing with limited knowledge of demand, much of which deals with experimentation and learning. One of the best (and earliest) examples is [15], who assumes that a firm chooses from a finite set of prices, observes outcomes, and because each trial is costly, eventually settles on the price that it thinks (perhaps incorrectly) is optimal. The firm’s choice is then the solution of a multi-armed bandit problem. (In the simplest version of the model, the firm prices high or low.) The solution does not involve estimating any demand curve.

Other studies focus on learning in either a parametric or non-parametric context. A variety of papers address the use of learning to update estimates of parameters of a known demand function using Bayesian or non-Bayesian techniques; examples include [3], [5], [14] and [11]. A second stream examines the interplay between learning and optimizing revenues over time without assuming any parametric form. Following [15], several authors assume the seller first sets a price to learn about demand, and then adjusts the price to optimize revenues (see, e.g., [6] and [2]).

The operations research literature examines dynamic pricing using robust optimization. This approach assumes that the functional form of the demand curve is known but one or more parameters are only known to lie in an “uncertainty set”. For example, the demand function might depend on two unknown parameters  $\alpha_1$  and  $\alpha_2$ , so the profit function is  $\Pi(\alpha_1, \alpha_2, p)$ . The decision variable, in this case price, is chosen to maximize the worst possible outcome over the uncertainty set, i.e.,  $\max_p \min_{\alpha_1, \alpha_2} \Pi(\alpha_1, \alpha_2, p)$ .<sup>4</sup> Although robust optimization incorporates uncertainty, it does not allow for learning, and by only considering the worst-case scenario it can yield conservative pricing strategies.

Our work is also related to studies of model misspecification. In particular, we want to show that a simple linear demand model may perform well even if the true

---

<sup>4</sup>See, e.g., [1] and [17]. An alternative is the “distributionally robust” approach, where price is robust with respect to a class of demand distributions with similar parameters such as mean and variance. See, e.g., [13], and [4].

demand curve is far from linear. Others have likewise shown that linear models can perform well (e.g., [10] in clinical prediction and [8] in contract theory). [7] study the “price of misspecification” for dynamic pricing with demand learning. They investigate the revenue loss in a multi-period setting incurred if the seller uses a simple parametric demand model that differs significantly from the underlying demand curve.

Our approach to pricing is quite different from the studies cited above, and is related to the prescriptive rules of thumb found, for example, in [16]. Managers often seek simple and robust rules for pricing (and other decisions such as levels of advertising or R&D), and other studies have shown that simple rules can be very effective.<sup>5</sup> The pricing rule we suggest is certainly simple; the extent to which it is effective is the focus of this chapter.

In the following sections, we characterize the performance of our pricing rule by deriving analytical bounds for the profit ratio  $\Pi^{**}/\Pi^*$  for several classes of demand curves: quadratic demand, monomial demand functions, semi-log and log-log demand functions. We also find bounds for the ratio  $\Pi^{**}/\Pi^*$  in the case of a general concave demand and extend our results when the maximum price is not known exactly. Finally, we examine randomly generated “true” demand curves, and determine computationally the expected profit performance of our pricing rule, and confidence bounds for the ratio  $\Pi^{**}/\Pi^*$ .

## 4.2 Common Demand Functions

Here we examine several demand models — quadratic, monomial, polynomial, semi-log and log-log. For each we derive performance bounds comparing the profits from

---

<sup>5</sup>In related work, [9] show how a technique they call “Bundle Size Pricing” (BSP) provides a close approximation to optimal mixed bundling. In BSP, a price is set for each individual good, for any bundle of two, for any bundle of three, etc., up to a bundle of all the goods produced. Profits turn out to be very close to what would be obtained from mixed bundling. Also [8] examines a principal who has only limited knowledge of what an agent can or cannot do, and wants to write a contract robust to this uncertainty. He shows that the most robust contract is a linear one — e.g., the agent is paid a fixed fraction of output. Finally, [12] provides a general treatment of robust control, i.e., optimal control with model uncertainty.

our simple pricing rule to the profits that would result if the actual demand function were known. We will see that in most cases the profit ratio is close to one.

Before proceeding, we note that the difference between the linear price  $P^*$  and the optimal price  $P^{**}$  using the actual demand function depends on the convexity properties of the function. In the Appendix we show that if the actual inverse demand curve is convex (concave) with respect to  $Q$ , the linear price is greater (smaller) than the optimal price:

**Result 4.1.** *If the actual inverse demand curve  $P_A(Q)$  is convex with respect to  $Q$ , then  $P^{**} \leq P^*$ , and if  $P_A(Q)$  is concave,  $P^{**} \geq P^*$ .*

Note that we only need  $P_A(Q)$  to be convex (or concave) in the range  $[0, Q^{**}]$  and not everywhere. The value of  $Q^{**}$  might not be known, but this result can still be useful in that it tells us whether our simple rule will over- or under-price relative to the optimal price, and it might be possible to correct for this error by adjusting the price up or down.

### 4.2.1 Quadratic Demand

Suppose the actual inverse demand function is a quadratic of the form:

$$P_A(Q) = P_m - b_1Q + b_2Q^2, \quad (4.2.1)$$

where, as before,  $P_m$  is the maximum price. We want analytical bounds for the profit ratio  $\Pi^{**}/\Pi^*$  and price ratio  $P^{**}/P^*$ . The bounds depend on the convexity properties of the function in (4.2.1) and are summarized in the following result. (Proofs are in the Appendix.)

**Result 4.2.** *For the quadratic inverse demand curve of eqn. (4.2.1), the ratios of profits and prices satisfy the following relations:*

- **Convex case:**  $b_1, b_2 \geq 0$  and  $b_2 \leq b_1^2/4P_m$

$$1 \leq \frac{\Pi^{**}}{\Pi^*} \leq \frac{8\sqrt{2}}{27(\sqrt{2}-1)} = 1.0116$$

$$\frac{8}{9} \leq \frac{P^{**}}{P^*} \leq 1$$

- **Concave case:**  $b_1 \geq 0$  and  $b_2 \leq 0$

$$1 \leq \frac{\Pi^{**}}{\Pi^*} \leq \frac{4\sqrt{2}}{3\sqrt{3}} = 1.0887$$

$$1 \leq \frac{P^{**}}{P^*} \leq \frac{2}{3} \left( \frac{2P_m + c}{P_m + c} \right) \leq \frac{4}{3} = 1.33$$

Note that the restrictions on the values of  $b_1$  and  $b_2$  are necessary and sufficient conditions to guarantee that the inverse demand curve is non-negative and non-increasing everywhere.

If the demand curve is convex, the simple pricing rule is very close to the optimal; it will yield a profit that is only about 1% less than what the firm could achieve if it knew the true demand curve. Also, this is a “worst case” result that occurs when  $c = 0$ ; if  $c > 0$ , the ratio  $\Pi^{**}/\Pi^*$  is even closer to 1. The price  $P^*$  can be as much as 12% lower than the optimal price  $P^{**}$ , but of course the concern of the firm is (or should be) its profit performance. (Also,  $P^{**}/P^*$  deviates the most from 1 when  $c = 0$ .)

In the concave case, the resulting profit  $\Pi^*$  will be within 9% of the optimal profit, irrespective of the parameters  $b_1$  and  $b_2$ . In the proof of Result 4.2 in the Appendix, we show that the largest value of  $\Pi^{**}/\Pi^*$  (1.0887) occurs when  $b_1 = 0$ ; for positive values of  $b_1$ , the profit ratio is closer to 1. The reason is that when  $b_1$  increases, the curve becomes closer to a linear function. In addition, one can show that the profit ratio becomes closer to 1 for the concave case when either  $c$  or  $b_2$  increase (recall that  $b_2 \leq 0$ ).

## 4.2.2 Monomial Demand

Now suppose the inverse demand curve is a monomial of order  $n$ :

$$P_A(Q) = P_m - \gamma Q^n, \quad \gamma > 0. \quad (4.2.2)$$

Note that all functions of the form (4.2.2) are concave and decreasing, given that  $\gamma > 0$ . Now the profit and price ratios are as follows. (Proof in Appendix.)

**Result 4.3.** *For the inverse demand curve of eqn. (4.2.2), the profit and price ratios satisfy:*

$$1 \leq \frac{\Pi^{**}}{\Pi^*} = \frac{2^{\frac{1}{n}+1}n}{(n+1)^{\frac{1}{n}+1}} \leq 2$$

$$1 \leq \frac{P^{**}}{P^*} = \frac{2(nP_m + c)}{(n+1)(P_m + c)} \leq 2$$

Note that the above results are not bounds but equalities. Also, note that for any monomial inverse demand curve, the profit ratio only depends on the order of the monomial  $n$ ; it does not depend on the values of  $P_m$ ,  $c$  or  $\gamma$ . (The price ratio does depend on  $P_m$ ,  $c$  and  $n$ , but not on  $\gamma$ .) Both ratios are monotonically increasing with the degree of the monomial  $n$  and converge to 2 and  $2P_m/(P_m + c) \leq 2$  respectively, as  $n \rightarrow \infty$ . For monomials of order 3 and 4, the profit ratios are 1.19 and 1.27 respectively.

These results can be extended to the case of polynomials with non-positive coefficients. In particular, one can show that for a polynomial inverse demand function of order  $n$  with non-positive coefficients, the ratios of profits and prices attain their worst values for the corresponding monomials. By adding some extra non-positive elements to the monomial, the inverse demand curve becomes closer to a linear function, improving the performance of our simple pricing rule. In short, if the true inverse demand function is a low-order monomial or polynomial, the firm can expect profits from our pricing rule that are reasonably close to what it could achieve if it knew the

true demand curve.

### 4.2.3 Semi-Log Demand

Now consider the semi-log inverse demand curve:

$$P_A(Q) = P_m e^{-\alpha Q}, \quad \alpha > 0. \quad (4.2.3)$$

The following result (proof in Appendix) shows the profit and price ratios for the case where marginal cost  $c = 0$  and then for  $c > 0$ .

**Result 4.4.** *For the semi-log inverse demand curve of eqn. (4.2.3),*

- *When  $c = 0$ , the ratios of profits and prices are:*

$$\Pi^{**}/\Pi^* = 2e^{-1}/\log(2) = 1.0615$$

$$P^{**}/P^* = 2e^{-1} = 0.7357$$

- *When  $c > 0$ , the ratios are closer to 1:*

$$1 \leq \Pi^{**}/\Pi^* < 1.0615$$

$$0.7357 < P^{**}/P^* \leq 1$$

If  $c = 0$  both ratios can be computed exactly and do not depend on  $\alpha$  or  $P_m$ ; in this worst case, the simple pricing rule yields a profit that differs from the optimal by only 6.15%, even though the prices differ by 26.5%. When  $c > 0$ , one cannot compute the ratios in closed form. Instead, we solve numerically for  $\Pi^{**}$  and  $P^{**}$  and present the results in Figure 4-2, where we plot both ratios as a function of  $c/P_m$ . (From numerical tests, we find that the ratios are independent of the value of  $\alpha$ .) Note from the figure that as  $c$  increases, the performance of the pricing rule improves, as both ratios approach 1.

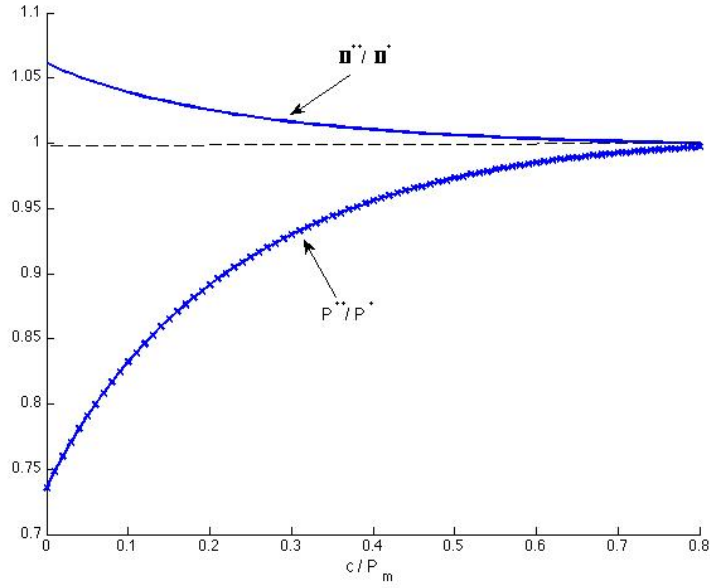


Figure 4-2: Profit and price ratios for the semi-log inverse demand curve as a function of  $c/P_m$

#### 4.2.4 Log-Log Demand

We turn now to the commonly used log-log demand model:

$$P_A(Q) = A_0 Q^{-1/\beta}; \quad \beta > 1, \quad (4.2.4)$$

where  $-\beta$  is the (constant) elasticity of demand. Because this demand curve has no maximum price, we truncate it so that  $P(0) = P_m$ . Setting  $P_A(Q_0) = P_m$ , the corresponding quantity is  $Q_0 = (P_m/A_0)^{-\beta}$ . We therefore work with the following modified version of eqn (4.2.4):

$$P_A(Q) = \begin{cases} P_m; & \text{if } Q < Q_0 \\ P_m(Q/Q_0)^{-1/\beta}; & \text{if } Q \geq Q_0 \end{cases} \quad (4.2.5)$$

We require that  $\beta > \beta_{min} = P_m/(P_m - c)$  in order for the optimal price  $P^{**}$  to be smaller than the maximum price  $P_m$ . The performance of our pricing rule in this

case is:

**Result 4.5.** *For the demand curve of eqn. (4.2.5), the profit and price ratios are:*

$$\frac{\Pi^{**}}{\Pi^*} = \frac{2}{(P_m/c - 1)(\beta - 1)} \left[ \frac{2\beta}{(P_m/c + 1)(\beta - 1)} \right]^{-\beta}$$

$$\frac{P^{**}}{P^*} = \frac{2\beta}{(P_m/c + 1)(\beta - 1)}$$

Note that these ratios are exact, and depend only on the elasticity  $\beta$  and  $P_m/c$ . Also, there is a unique value of  $\beta^* = (P_m + c)/(P_m - c)$  for which both ratios equal 1.<sup>6</sup>

There are two limiting cases to note:  $c$  large and  $c$  very small. If  $c$  is large, i.e.,  $c \rightarrow P_m$ ,  $\beta_{min} \rightarrow \infty$ . If  $\beta_{min}$ , and hence  $\beta$ , are very large, both the price and profit ratios approach 1. However, in this case the truncated demand curve may not make sense, because unless one thinks the true demand is very elastic, we will have  $P^{**} = \beta c / (\beta - 1) > P_m$ . At the other extreme, as  $c \rightarrow 0$ ,  $P^{**} \rightarrow 0$ , whereas  $P^* \rightarrow .5P_m$ , and  $\Pi^{**}/\Pi^*$  is unbounded. But again, an isoelastic demand curve would then make little sense, because  $Q^{**} \rightarrow \infty$ .

The general case is illustrated in Figure 5-3, which shows the profit and price ratios as a function of  $P_m/c$  for  $\beta = 1.5, 2.0$ , and  $2.5$ . If  $\beta = 1.5$ ,  $\Pi^{**}/\Pi^*$  is always close to 1. But if  $\beta = 2.5$ ,  $\Pi^{**}/\Pi^*$  can exceed 2 for large enough values of  $P_m/c$ .<sup>7</sup> (Note that  $P^{**}$  can be larger or smaller than  $P^*$ .) Thus if demand is very elastic (i.e.,  $\beta$  is large) or marginal cost  $c$  is small, our simple pricing rule may not perform well.

Table 4.1 summarizes the price and profit ratios for the different demand curves.

---

<sup>6</sup>If  $\beta = \beta^*$ , the elasticity of the isoelastic demand equals the elasticity of the linear demand at the optimal price. The latter elasticity is  $E_d = bP^*/Q_L^* = (P_m + c)/(P_m - c)$ , so if  $\beta = \beta^*$ , both the linear and log-log demand curves have the same profit-maximizing price and output.

<sup>7</sup>The log-log demand curve is convex, but truncating modifies its convexity properties, which affects the relationship between  $P^{**}$  and  $P^*$  (see Result 4.1). If either  $\beta$  or  $P_m/c$  is small, the optimal quantity  $Q^{**}$  is small and can lie on the truncated — and non-convex — part of the curve.

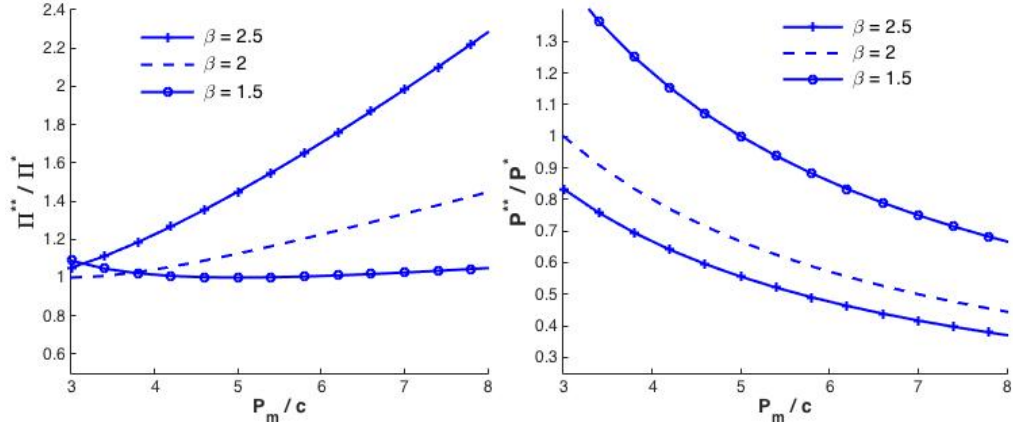


Figure 4-3: Profit and price ratios for the log-log inverse demand curve as a function of  $P_m/c$

Inverse demand function	$P^{**}/P^*$	$\Pi^{**}/\Pi^*$
<b>Quadratic convex:</b> $P_A(Q) = P_m - b_1Q + b_2Q^2$ $b_1, b_2 \geq 0$ and $b_2 < b_1^2/4P_m$	$\frac{8}{9} \leq P^{**}/P^* \leq 1$	$\leq 1.0116$
<b>Quadratic concave:</b> $P_A(Q) = P_m - b_1Q + b_2Q^2$ $b_1 \geq 0$ and $b_2 \leq 0$	$1 \leq P^{**}/P^* \leq \frac{4P_m+2c}{3P_m+3c} \leq 1.33$	$\leq 1.0887$
<b>Monomial:</b> $P_A(Q) = P_m - \gamma Q^n$ $n = 3$ $n = 4$	$2(nP_m + c)/(n + 1)(P_m + c)$ $\leq 1.5$ $\leq 1.6$	$2^{(n+1)/n}n/(n + 1)^{(n+1)/n}$ $1.19$ $1.27$
<b>Semi-log:</b> $P_A(Q) = P_m e^{-\alpha Q}$ $c = 0$ $c > 0$	$0.7357$ $< 0.7357$	$1.0615$ $< 1.0615$
<b>Log-log (truncated):</b> $P_A(Q) = \begin{cases} P_m; & \text{if } Q < Q_0 \\ P_m(Q/Q_0)^{-1/\beta}; & \text{if } Q \geq Q_0 \end{cases}$ $\beta \geq \beta_{\min} = P_m/(P_m - c)$	$2\beta/(P_m/c + 1)(\beta - 1)$	$\frac{2}{(P_m/c-1)(\beta-1)} \left( \frac{2\beta}{(P_m/c+1)(\beta-1)} \right)^{-\beta}$

Table 4.1: Price and profit ratios for several “true” demand curves

### 4.3 Uncertain Maximum Price

We have assumed that while the firm does not know its true demand curve, it *does* know the maximum price  $P_m$  it can charge and still expect to sell at least some units. Here we consider the possibility that the firm only has an estimate of the maximum

price:

$$\hat{P}_m = P_m(1 + \epsilon), \quad (4.3.1)$$

where  $\epsilon$  lies in some interval  $[-B, B]$ , with  $0 \leq B \leq 1$ . Our pricing rule is now  $P^* = (\hat{P}_m + c)/2$ , and suffers from two miss-specifications: the form of the demand curve and the value of the intercept. To see how the profit ratio  $\Pi^{**}/\Pi^*$  is affected by this additional source of uncertainty, we assume that  $\epsilon$  is a uniform random variable with  $B = 0.2$  (i.e.,  $\epsilon \sim U[-0.2, 0.2]$ ). We will show that our results are reasonably robust to misspecification of the maximum price of up to 20%.

“True” inverse demand function	$\mathbb{E}[\Pi^{**}/\Pi^*]$	$\Pi^{**}/\Pi^*(\epsilon = 0)$
<b>Linear:</b> $P_A(Q) = P_m - bQ$	1.0137	1
<b>Quadratic convex:</b> $P_A(Q) = P_m - b_1Q + b_2Q^2$ $b_1, b_2 \geq 0$ and $b_2 < b_1^2/4P_m$	1.0254	1.0116
<b>Quadratic concave:</b> $P_A(Q) = P_m - b_1Q + b_2Q^2$ $b_1 \geq 0$ and $b_2 \leq 0$	1.1023	1.0887
<b>Monomial:</b> $P_A(Q) = P_m - \gamma Q^n$ $n = 3$	1.205	1.19
$n = 4$	1.2882	1.27
<b>Semi-log:</b> $P_A(Q) = P_m e^{-\alpha Q}$	1.0748	1.0615

Table 4.2: Expected profit ratios when  $\hat{P}_m = P_m(1 + \epsilon)$  and  $\epsilon \sim U[-0.2, 0.2]$

To simplify matters, we assume in this section that  $c = 0$ . (Recall from the previous section that  $\Pi^{**}/\Pi^*$  deviates from 1 the most when  $c = 0$  for the demand curves we considered.) We derive and present in the Appendix closed form expressions

for  $\Pi^{**}/\Pi^*$  as a function of  $\epsilon$  for the demand models we examined in Section 4.2. We also compute the expected value of the profit ratio when  $\epsilon \sim U[-0.2, 0.2]$ , as well as the profit ratio that results when the true  $P_m$  is known exactly (i.e.,  $\epsilon = 0$ ). The results are shown in Table 4.2.

Note that one of the “true” demand curves in Table 4.2 is the linear demand  $P_A(Q) = P_m - bQ$ , but our pricing rule is based on  $\hat{P}_m$ . This misspecification yields an expected profit loss of about 1.4%. Observe that for the quadratic, monomial and semi-log inverse demand functions,  $\mathbb{E}[\Pi^{**}/\Pi^*]$  is very close to the ratio when the true value of  $P_m$  is known. (We show  $\Pi^{**}/\Pi^*$  for the truncated log-log curve in the Appendix, but omit its expectation as it depends on  $P_m/c$  and  $\beta$ .)

Of course, depending on the “draw” for  $\epsilon$ , the actual profit ratio could be farther from 1. To see how much farther, we used the closed-form expressions in the Appendix to plot the profit ratios as a function of  $\epsilon$  for  $-0.2 \leq \epsilon \leq 2$ . As Figure 4-4 shows, the monomial demand (with  $n = 3$ ) is most sensitive to the value of  $\epsilon$ , with  $\Pi^{**}/\Pi^*$  reaching 1.5 when  $\epsilon = -0.2$ . For the other demand curves, however,  $\Pi^{**}/\Pi^*$  is below 1.25 for the entire range of  $\epsilon$  we consider. Thus the performance of our pricing rule is degraded by a misspecification of the maximum price, but only moderately.

## 4.4 General Demand Functions

We have seen that our simple pricing rule works well for a variety of underlying demand functions — but not all. In particular, if the true demand is a truncated log-log function,  $\Pi^{**}/\Pi^*$  can deviate substantially from 1 if demand is very elastic and/or marginal cost is small. This follows from the convexity of this demand curve, and the fact that (unrealistically) the quantity demanded expands without limit as the price is reduced towards zero.

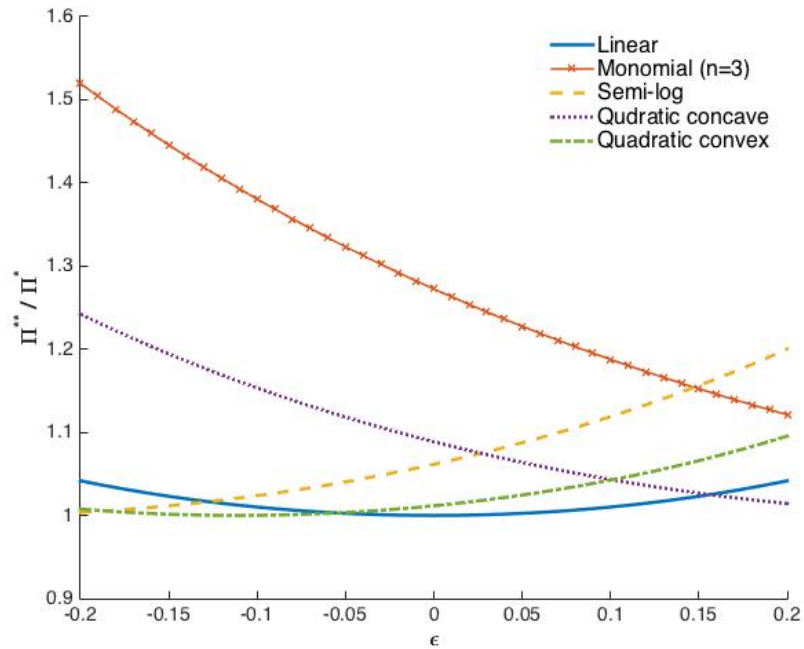


Figure 4-4: Profit ratios as a function of  $\epsilon$  for different demand curves

#### 4.4.1 Concave Demand Curves

Even if the true demand curve is concave,  $\Pi^{**}/\Pi^*$  might substantially exceed 1. In fact, if  $P_A(Q)$  is non-increasing and concave, one can show the following.

**Result 4.6.** *For any concave inverse demand curve, we have:*

$$1 \leq \Pi^{**}/\Pi^* \leq 2$$

$$1 \leq P^{**}/P^* \leq 2$$

In the worst case, the profit and price ratios will equal 2 if the true demand curve is a rectangle. For other concave functions,  $\Pi^{**}/\Pi^* < 2$ , but except for specific functional forms, we cannot say how much less.

We might expect that in some cases the inverse demand curve will not be concave and may even have a flat area (plateau), as in Figure 4-1. In this case,  $\Pi^{**}/\Pi^*$  will be sensitive to whether the plateau is below or above  $P^*$ . If the plateau is below  $P^*$

and very long,  $\Pi^{**}/\Pi^*$  can be arbitrarily large; by pricing at  $P^*$ , the firm is missing a large mass of consumers. But if the plateau is above  $P^*$ ,  $\Pi^{**}/\Pi^*$  will usually be close to 1. Thus if the firm believes there is such a plateau, it might set price below  $P^*$  in order to capture it.

In practice, a firm introducing a new product may know little or nothing about the shape of the demand curve. Indeed, that is the motivation for this work. The firm might have no reason to expect that demand is characterized by one of the commonly used functions we examined earlier, or any other particular function. If the firm uses our pricing rule — with no knowledge at all of the true demand curve, other than the maximum price  $P_m$  — how well can it expect to do?

#### 4.4.2 General Random Demand Curves

We address this question by randomly generating a set of “true” demand curves. For each randomly generated curve we compute (numerically) the profit-maximizing price and profit,  $P^{**}$  and  $\Pi^{**}$ , and we compare  $\Pi^{**}$  to the profit  $\Pi^*$  the firm would earn by using our pricing rule, i.e., by setting the price  $P^* = (P_m + c)/2$ . We generate 100,000 such demand curves, and then examine the resulting distribution of  $\Pi^{**}/\Pi^*$ . The only restriction we impose on the randomly generated demand curves is that they are non-increasing everywhere.

We generate each demand curve as follows. We take three numbers as known: the maximum price  $P_m$ , marginal cost  $c$ , and the maximum quantity  $Q_{max}$  that can be sold at a price of zero (i.e., the maximum potential size of the market). We present results for  $c/P_m = 0$  and 0.5. (We found that our results are insensitive to the value of  $Q_{max}$ .)

We divide the segment  $[0, Q_{max}]$  into  $S$  equally spaced intervals, and generate a piece-wise non-increasing demand curve by drawing random values for the different pieces. Since  $P(0) = P_m$  and  $P(Q_{max}) = 0$ ,  $S$  intervals imply  $S - 1$  breaking points between 0 and  $P_m$ . (One might interpret this partition of the market as representing customer segments, or simply an approximation to a continuous curve.) With this

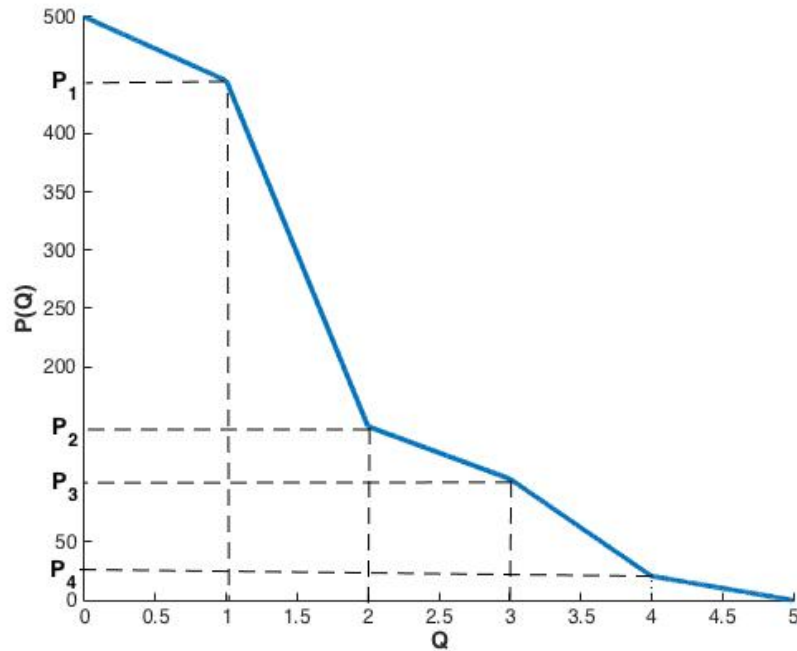


Figure 4-5: Randomly generated inverse demand curve with  $S = 5$  pieces

partition, we draw a value for the end of the first segment from a uniform distribution between 0 and  $P_m$ . Call this random value  $P_1$  (see one realization for  $P_1$  in Figure 4-5). Next, we draw a value for the end of the second segment, again from a uniform distribution but now between 0 and  $P_1$ . Call this random value  $P_2$ . We repeat this process until we have  $S - 1$  randomly generated non-increasing prices, i.e., a random demand curve that has  $S$  segments. Figure 4-5 shows an example of such a randomly generated demand curve that has 5 segments (for  $P_m = 500$  and  $Q_{max} = 5$ ). Given this demand curve, we calculate  $P^{**}$ ,  $\Pi^{**}$ , and the profit ratio  $\Pi^{**}/\Pi^*$ .

We generate 100,000 demand curves and compute 100,000 corresponding values for the profit ratio  $\Pi^{**}/\Pi^*$ . We calculate the mean value of  $\Pi^{**}/\Pi^*$ , as well as the 80% and 90% points (i.e., the value of  $\Pi^{**}/\Pi^*$ , such that 80% or 90% of the randomly generated ratios are below this number). The number of segments  $S$  can affect the resulting  $\Pi^{**}$ , so in Table 4.3 we show results for different values of  $S$  and for  $c/P_m$  equal to 0 and 0.5.

$c = 0$				$c = 0.5P_m$			
$S$	Mean	80%	90%	$S$	Mean	80%	90%
2	1.0672	1.1625	1.2442	2	1.0748	1.1255	1.3696
5	1.1332	1.2057	1.3926	5	1.0525	1.0645	1.2271
10	1.1351	1.2081	1.3979	10	1.0523	1.0647	1.2254
50	1.1379	1.2161	1.4071	50	1.0525	1.0621	1.2264
100	1.1344	1.2124	1.4045	100	1.0525	1.0628	1.2265

Table 4.3: Profit ratios for randomly generated demands

One can see that whatever the number of segments,  $S$ , the average profit ratio is less than 1.14 if  $c = 0$  and less than 1.08 if  $c = 0.5P_m$ . Also, 80% (90%) of the demand curves yield profit ratios less than 1.22 (1.41) if  $c = 0$  and less than 1.13 (1.37) if  $c = 0.5P_m$ . In Figure 4-6 we plot the histogram of the 100,000 profit ratios for  $S = 5$  and both  $c = 0$  and  $c = 0.5P_m$ . When  $c = 0$  ( $c = 0.5P_m$ ), more than 40% (75%) of the ratios are less than 1.01, and 54% (79%) are less than 1.05. Thus it is likely that our simple pricing rule will yield a profit close to what would result if the firm actually knew its demand curve.

### 4.4.3 Uncertain Maximum Price

What if the maximum price  $P_m$  is not known exactly? As in Section 4.3, we assume that the firm only has an estimate  $\hat{P}_m$  of  $P_m$ , given by eqn. (4.3.1). As before, we generate 10,000 random demand curves of  $S$  segments each for different values of  $S$ . However, for each random demand curve, we now assume that the firm must base its price on its estimate  $\hat{P}_m$  of  $P_m$ , with  $\epsilon$  in eqn. (4.3.1) uniformly distributed over  $[-0.2, 0.2]$ . For each random demand curve we draw independently 100 values of  $\epsilon$ , yielding a total of 1 million random  $\hat{P}_m$ -demand curve combinations. We compute

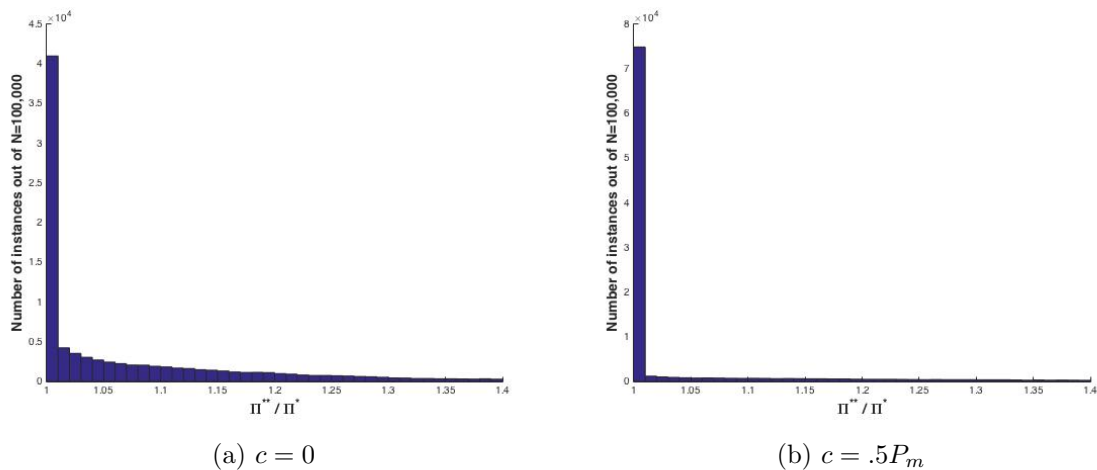


Figure 4-6: Histogram of profit ratios when  $S = 5$  for  $c = 0$  and  $c = 0.5P_m$

the expectation of the profit ratio  $\Pi^{**}/\Pi^*$  (over both the random variable  $\epsilon$  and the randomly generated demand curves). The results are shown in Table 4.4, along with the corresponding expected profit ratio when the maximum price is known exactly (i.e.,  $\epsilon = 0$ ).

Note from Table 4.4 that the expected profit ratio is always less than 1.15. We also calculated the 80% and 90% percentiles (as in Table 4.3) and the values are very close to the ones where the maximum price is known. This is consistent with the results in Section 4.3. Once again, our pricing rule is generally robust to miss-specification of the maximum price.

## 4.5 Conclusions

Setting price is one of the most basic economic decisions firms make. Introductory economics courses make this decision seem easy; just write down the demand curve and set marginal revenue equal to marginal cost. But of course firms rarely have precise knowledge of their demand curves. When introducing new products (or old products in new markets), firms may know little or nothing about demand, but must still set a price. Price experimentation is often not feasible, and the price a firm sets

$c = 0$			$c = 0.5P_m$		
$S$	Mean	$\epsilon = 0$	$S$	Mean	$\epsilon = 0$
2	1.0806	1.0672	2	1.1406	1.0748
5	1.145	1.1332	5	1.115	1.0525
10	1.149	1.1351	10	1.1159	1.0523
50	1.149	1.1379	50	1.115	1.0525
100	1.149	1.1344	100	1.116	1.0525

Table 4.4: Expected profit ratios when the maximum price is unknown

is often the one it sticks with for some time.

We have shown that under certain conditions the firm can use a simple pricing rule. The conditions are: (i) marginal cost  $c$  is known and constant, (ii) the firm need not predict its initial sales, and (iii) the firm can estimate the maximum price  $P_m$  it can charge and still expect to sell at least some units. These conditions often hold, especially for new products or services introduced by technology companies. The firm then simply sets a price of  $P^* = (P_m + c)/2$ .

How well can the firm expect to do if it follows this pricing rule? We have shown that if the true demand curve is one of several commonly used demand functions, or even if it is a more complex function, the firm can expect to do very well. In particular, it can expect to earn a profit  $\Pi^*$  that is reasonably close to the profit  $\Pi^{**}$  it could earn if it knew the true demand curve.

Some caveats are in order. Perhaps most important, our analysis has been entirely static. We have assumed that the true demand curve is fixed; it does not shift over time, perhaps in response to network externalities (which can be important for new products). We have also assumed that the firm sets and maintains a single price; it does not raise or lower price over time to inter-temporally price discriminate or to respond to changing market conditions, nor does it offer different prices to different

groups of customers. We have also ruled out learning about demand, either passively or via experimentation, which has been the focus of the earlier literature on optimal pricing with uncertain demand. To the extent that such dynamic considerations are important, our pricing rule can be viewed as a starting point. Managers often seek simple and robust rules for pricing; the rule we suggest is certainly simple, and we have seen that it is also effective.

## Bibliography

- [1] Elodie Adida and Georgia Perakis. A robust optimization approach to dynamic pricing and inventory control with no backorders. *Mathematical Programming*, 107(1-2):97–129, 2006.
- [2] Victor F Araman and René Caldentey. Revenue management with incomplete demand information. *Wiley Encyclopedia of Operations Research and Management Science*, 2011.
- [3] Yossi Aviv and Amit Pazgal. A partially observed markov decision process for dynamic pricing. *Management Science*, 51(9):1400–1416, 2005.
- [4] Michael O Ball and Maurice Queyranne. Toward robust revenue management: Competitive analysis of online booking. *Operations Research*, 57(4):950–963, 2009.
- [5] Dimitris Bertsimas and Georgia Perakis. *Dynamic pricing: A learning approach*. Springer, 2006.
- [6] Omar Besbes and Assaf Zeevi. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research*, 57(6):1407–1420, 2009.
- [7] Omar Besbes and Assaf Zeevi. On the (surprising) sufficiency of linear models

- for dynamic pricing with demand learning. *Columbia Business School Research Paper*, (12-5), 2013.
- [8] Gabriel Carroll. Robustness and linear contracts. *American Economic Review*, July 2014. forthcoming.
- [9] Chenghuan Sean Chu, Phillip Leslie, and Alan Sorensen. Bundle-size pricing as an approximation to mixed bundling. *American Economic Review*, 101:262–303, February 2011.
- [10] Robyn M Dawes. The robust beauty of improper linear models in decision making. *American psychologist*, 34(7):571, 1979.
- [11] Vivek F Farias and Benjamin Van Roy. Dynamic pricing with a prior on market response. *Operations Research*, 58(1):16–29, 2010.
- [12] Lars Peter Hansen and Thomas J. Sargent. *Robustness*. Princeton University Press, 2008.
- [13] Andrew EB Lim and J George Shanthikumar. Relative entropy, exponential utility, and robust dynamic pricing. *Operations Research*, 55(2):198–214, 2007.
- [14] Kyle Y Lin. Dynamic pricing with real-time demand learning. *European Journal of Operational Research*, 174(1):522–538, 2006.
- [15] Michael Rothschild. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9(2):185–202, 1974.
- [16] Oz Shy. *How to Price: A Guide to Pricing Techniques and Yield Management*. Cambridge University Press, 2006.
- [17] Aurélie Thiele. Multi-product pricing via robust optimisation. *Journal of Revenue & Pricing Management*, 8(1):67–80, 2009.



# Chapter 5

## The Impact of Demand

## Uncertainty on Consumer

## Subsidies for Green Technology

## Adoption

### 5.1 Introduction

Recent developments in green technologies have captured the interest of the public and private sectors. For example, electric vehicles (EV) historically predate gasoline vehicles, but have only received significant interest in the last decade (see [30] for an overview). In the height of the economic recession, the US government passed the American Recovery and Reinvestment Act of 2009 which granted a tax credit for consumers who purchased electric vehicles. Besides boosting the US economy, this particular tax incentive was aimed at fostering further research and scale economies in the nascent electric vehicle industry. In December 2010, the all-electric car, Nissan Leaf, and the plug-in hybrid General Motors' Chevy Volt were both introduced in the US market. After a slow first year, sales started to pick up and most major car

companies are now in the process of launching their own versions of electric vehicles.

More recently in 2012, Honda introduced the Fit EV model and observed low customer demand. After offering sizable leasing discounts, Honda quickly sold out in Southern California<sup>1</sup>. It is not uncommon to read about waitlists for Tesla's new Model S or the Fiat 500e, while other EVs are sitting unwanted in dealer parking lots. Both stories of supply shortages or oversupply have been commonly attributed to electric vehicle sales. At the root cause of both these problems is demand uncertainty. [53] studied the supply shortages and customer wait lists shortly after Toyota launched the hybrid electric Prius in 2002. When launching a new product, it is hard to know how many units customers will request. In addition, finding the correct price point is also not a trivial task, especially with the presence of a government subsidy. In fact, understanding demand uncertainty should be a first order consideration for both manufacturers and policy-makers alike.

For the most part, the subsidy design literature in green technologies has not studied demand uncertainty (see for example, [10], [7], [45] and [2]). In practice, demand uncertainty has also often been not considered. As suggested in private communication with several sponsors of the MIT Energy Initiative<sup>2</sup>, policy makers often ignore demand uncertainty when designing consumer subsidies for green technology adoption<sup>3</sup>. The purpose of this chapter is to study whether incorporating demand uncertainty in the design of subsidy programs for green technologies is important. In particular, we examine how governments should set subsidies when considering the manufacturing industry's response under demand uncertainty. We show that demand uncertainty plays a significant role in the system's welfare distribution and should not be overlooked.

Consider the following two examples of green technologies: electric vehicles and

---

<sup>1</sup>Electric vehicles in short supply - Los Angeles Times - 06/05/2013 - <http://articles.latimes.com/2013/jun/05/autos/la-fi-hy-autos-electric-cars-sold-out-20130605>

<sup>2</sup><http://mitei.mit.edu/about/external-advisory-board>

<sup>3</sup>Incentives for green technology adoption - Energy Futures, MIT Energy Initiative - <http://mitei.mit.edu/news/incentives-green-technology-adoption-getting-government-subsidies-right>

solar panels. By the end of 2013, more than 10GW of solar photovoltaic (PV) panels had been installed in the United States, producing an annual amount of electricity roughly equivalent to two Hoover Dams. While still an expensive generation technology, this large level of installation was only accomplished due to the support of local and federal subsidy programs, such as the SunShot Initiative. In 2011, the US energy secretary Steven Chu announced that the goal of the SunShot Initiative by 2020 is to reduce the total cost of PV systems by 75%, or an equivalent of \$1 a Watt ([29]), at which point solar technology will be competitive with traditional sources of electricity generation. Even before this federal initiative, many states have been actively promoting solar technology with consumer subsidies in the form of tax rebates or renewable energy credits.

Similarly, federal subsidies were also introduced to stimulate the adoption of electric vehicles through the Recovery Act. As we previously mentioned, General Motors and Nissan have recently introduced affordable electric vehicles in the US market. GM's Chevy Volt was awarded the most fuel-efficient compact car with a gasoline engine sold in the US, as rated by the United States Environmental Protection Agency ([31]). However, the price tag of the Chevy Volt is still considered high for its category. The cumulative sales of the Chevy Volt in the US since it was launched in December 2010 until September 2013 amount to 48,218. It is likely that the \$7,500 government subsidy offered to each buyer through federal tax credit played a significant role in the sales volume. The manufacturer's suggested retail price (MSRP) of GM's Chevy Volt in September 2013 was \$39,145 but the consumer was eligible for \$7,500 tax rebates so that the effective price reduced to \$31,645. The amount of consumer subsidies has remained constant since launch in December 2010 until the end of 2013. This seems to suggest that in order to isolate the impact of demand uncertainty without complicating the model, it is reasonable to consider a single period setting.

In this chapter, we address the following questions. How should governments design green subsidies when facing an uncertain consumer market? How does the uncertain demand and subsidy policy decision affect the supplier's price (MSRP)

and production quantities? Finally, what is the resulting effect on consumers? In practice, policy makers often ignore demand uncertainty and consider average values when designing consumer subsidies. This ignorance may be caused by the absence (or high cost) of reliable data, among other reasons. We are interested in understanding how the optimal subsidy levels, prices and production quantities as well as consumer surplus are affected when one explicitly considers demand uncertainty relative to the case when demand is approximated by its deterministic average value.

While the government designs subsidies to stimulate the adoption of new technologies, the manufacturing industry responds to these policies with the goal of maximizing its own profit. We model the supplier as a price-setting newsvendor that responds optimally to the government subsidy. More specifically, the supplier adjusts its production and price depending on the level of consumer subsidies offered by the government to the consumer. This study also helps us to expand the price-setting newsvendor model while accounting for the external influence of the government.

In our model, the government is assumed to have a given adoption target for the technology. This is motivated by several examples of policy targets for electric vehicles and solar panels. For instance, in the 2011 State of the Union, US President Barack Obama mentioned the following goal: “With more research and incentives, we can break our dependence on oil with bio-fuels and become the first country to have a million electric vehicles on the road by 2015” ([28]). Another example of such adoption target has been set for solar panels in the California Solar Incentive (CSI) program, which states that: “The CSI program has a total budget of \$2.167 billion between 2007 and 2016 and a goal to install approximately 1,940 MW of new solar generation capacity” ([25]). Hence, in our model, we optimize the subsidy level to achieve a given adoption target level while minimizing government expenditure. In Section 3, we discuss alternative models for the government (such as maximizing the total welfare) as well as consider the subsidy program budget, emission reductions and social welfare.

We then quantify the impact of demand uncertainty on government expenditures,

firm profit and consumer surplus. We further characterize who bears the cost of uncertainty depending on the structure of the demand model. Finally, we study the supply-chain coordination i.e., when the government owns the supplier, and show that subsidies coordinate the overall system. More precisely, we show that the price paid by consumers as well as the production level coincide in both the centralized (where the supplier is managed/owned by the government) and the decentralized models (where supplier and government act separately).

## Contributions

Given the recent growth of green technologies, supported by governmental subsidy programs, this chapter explores a timely problem in supply chain management. Understanding how demand uncertainty affects subsidy costs, as well as the economic surplus of suppliers and consumers, is an important part of designing sensible subsidy programs. The main contributions of this chapter are:

- *Demand uncertainty does not always benefit consumers: Nonlinearity plays a key role.*

As uncertainty increases, quantities produced increase whereas the price and the supplier's profit decrease. In general, demand uncertainty benefits consumers in terms of effective price and quantities. One might hence expect the aggregate consumer surplus to increase with uncertainty. In fact, we show this is not always true. We observe that the effect of uncertainty on consumer surplus depends on the demand form. For example, for linear demand, uncertainty increases the consumer surplus, whereas for iso-elastic demand the opposite result holds. Depending on the demand pattern, the possibility of not serving customers with high valuations can outweigh the benefit of reduced prices for the customers served.

- *By ignoring demand uncertainty, the government will under-subsidize and miss the desired adoption target.*

Through the case of the newly introduced Chevy Volt by General Motors in the US market, we measure by how much the government misses the adoption target by ignoring demand uncertainty. We show that when the supplier takes into account demand uncertainty information while the government considers only average information on demand, the resulting expected sales can be significantly below the desired target adoption level.

- *The cost of demand uncertainty is shared between the supplier and the government.*

We analyze who bears the cost of demand uncertainty between government and supplier, which we show depends on the profit margin of the product. In general, the government expenditure increases with the added inventory risk. For linear demand models, the cost of demand uncertainty shifts from the government to the supplier as the adoption target increases or the production cost decreases.

- *Consumer subsidies are a sufficient mechanism to coordinate the government and the supplier.*

We compare the optimal policies to the case where a central planner manages jointly the supplier and the government. We determine that the price paid by the consumers and the production levels coincide for both the decentralized and the centralized models. In other words, consumer subsidies coordinate the supply-chain in terms of price and quantities.

## 5.2 Literature Review

Our setting is related to the newsvendor problem which has been extensively studied in the literature (see, e.g., [63], [51], [61] and the references therein). An interesting extension that is even more related to this research is the price-setting newsvendor (see [49] and [62]). More recently, [42] identified a new measure of demand elasticity, the

elasticity of the lost sales rate, to generalize and complement assumptions commonly made in the price-setting newsvendor. [41] provides a good survey of the literature on inventory risk sharing in a supply chain with a newsvendor-like retailer, which is closer to our framework. Nevertheless, our problem involves an additional player (the government) that interacts with the supplier’s decisions and complicates the analysis and insights. Most previous works on the stochastic newsvendor problem treat the additive and multiplicative models separately (e.g., in [49]) or focus exclusively on one case, with often different conclusions regarding the price of demand uncertainty. In our problem however, we show that our conclusions hold for both demand models.

In the traditional newsvendor setting, the production cost is generally seen as the variable cost of producing an extra unit from raw material to finished good. In capital-intensive industries like electric vehicles, the per-unit cost of capacity investment in the manufacturing facility is usually much larger than the per-unit variable cost. For this reason, we define the production quantities of the supplier to be a capacity investment decision, similar to [15].

Another stream of research related to our work considers social welfare and government subsidies in the area of vaccines (see, e.g., [6], [47] and [58]). In [6], the authors study the impact of yield uncertainty, in a model that represents both supply and demand, on the inefficiency in the influenza vaccine supply chain. They show that the equilibrium demand can be greater than the socially optimal demand. In [58], the authors assume a single supplier with stochastic demand and consider how a donor can use sales and purchase subsidies to improve the availability of vaccines.

Among papers that study the design of subsidies for green technologies, [18] examines the social benefits of electric vehicle adoption in Sweden and report a pessimistic outlook for this technology in the context of net social welfare. [8] shows that adoption of electric vehicles has societal and environmental benefits, as long as the electricity grid is sufficiently clean. In [10], the authors develop a model for optimizing social welfare with solar subsidy policies in California. These two papers assume non-strategic industry players. While considering the manufacturer’s response, [7] studies the use of

a take-back subsidy and product recycling programs. In a similar way as the previous papers mentioned above, they optimize social welfare of the system assuming a known environmental impact of the product. Our work focuses on designing optimal policies to achieve a given adoption target level, which can be used to evaluate the welfare distribution in the system. In this work, we also incorporate the strategic response of the industry into the policy making decision. Also considering an adoption level objective, [45] studies the problem of optimizing subsidy policies for solar panels and present an empirical study of the German solar market. The paper shows evidence that the current feed-in-tariff system used in Germany might not be efficiently using the positive network externalities of early adopters. [2] also tackles the feed-in-tariff design problem, comparing strategies for welfare maximization and adoptions targets. Finally, [48] presents a price setting newsvendor model for the case of public interest goods. The authors compare, for the case of linear demand, different government intervention mechanisms and study under what conditions the system is coordinated in terms of welfare, prices and supply quantities. On the other hand, we investigate the impact of demand uncertainty on the various players of the system for non-linear demands and model explicitly the strategic response of the manufacturer to the subsidy policy.

Numerous papers in supply-chain management focus on linear demand functions. Examples include [3] and [32] and the references therein. These papers study supply chain contracts where the treatment mainly focuses on linear inverse demand curves. In this work, we show that the impact of demand uncertainty on the optimal policies differs for some classes of non-linear demand functions relative to linear models. In particular, we observe that the effect of demand uncertainty depends on whether demand is convex (rather than linear) with respect to the price. In addition, the demand non-linearity plays a key role on the consumer surplus.

As mentioned before, our paper also contributes to the literature on supply chain coordination (see [14] for a review). The typical supply chain setting deals with a supplier and a retailer, who act independently to maximize individual profits. Mech-

anisms such as rebates ([57]) or revenue-sharing ([16]) can coordinate the players to optimize the aggregate surplus in the supply chain. [44] examines how wholesale price, quantity flexibility or buybacks can incentivize information sharing when introducing a new product with uncertain demand. [46] studies how a supplier should share demand uncertainty risk with the retailer when there is a lead-time contract. In [35, 36, 37], the authors study different types of contracts in a Stackelberg framework using a price-setting newsvendor model. In particular, [37] analyzes the effect of price and order postponements in a decentralized newsvendor model with multiplicative demand, wherein the manufacturer possibly offers a buyback rate. In our setting, the government and the supplier are acting independently and could perhaps adversely affect one another. Instead, we show that the subsidy mechanism is sufficient to achieve a coordinated outcome. [21] and [47] have looked at supply chain coordination in government subsidies for vaccines. Nevertheless, as we discussed above the two supply chains are fairly different.

Without considering demand uncertainty, there is a significant amount of empirical work in the economics literature on the effectiveness of subsidy policies for hybrid and electric vehicles. For example, [27] shows that there is a strong relationship between gasoline prices and hybrid adoption. [20] shows that hybrid car rebates in Canada created a crowding out of other fuel efficient vehicles in the market. [34] argues that sales tax waivers are more effective than income tax credits for hybrid cars. The increase in hybrid car sales from 2000-2006 is mostly explained by social preferences and increasing gasoline prices. In [1], the authors show that the auto industry innovates more in clean technologies when fuel prices are higher. [40] determines that incentives are only effective when the amount is sufficiently large. For plug-in electric cars, [54] argues that financial incentives, charging infrastructure, and local presence of production facilities are strongly correlated with electric vehicle adoption rates across different countries.

Also in economics literature, one can find a vast amount of papers that consider welfare implications and regulations for a monopolist (see, e.g., [59]). There is also

a relevant stream of literature on market equilibrium models for new product introduction (see, e.g., [39]). However, most of these papers do not consider demand uncertainty.

The issue of how demand uncertainty creates a mismatch in supply and demand has been mostly researched in the operations management literature, therefore we mainly focus our literature survey in this area. There are some exceptions in economics, such as [53], who essentially argues that consumers captured most of the incentives for the Toyota Prius, while the firm did not appropriate any of that surplus despite a binding production constraint. [53] shows that there was a shortage of vehicles manufactured to meet demand when the Prius was launched. This reinforces our motivation for studying a newsvendor model in this context.

Also considering demand uncertainty, [33] shows that an export subsidy (as opposed to a tax) is the equilibrium government strategy for a duopoly where each firm is in a different country and is uncertain about demand in the other country. In a slightly different setting, [12] argues that subsidies can be used to protect workers from uncertain industry shocks, when there is limited labor mobility.

Some works on electricity peak-load pricing and capacity investments address the stochastic demand case (see [24] for a review on that topic). In this context, it is usually assumed that the supplier knows the willingness to pay of customers and can therefore decline the ones with the lowest valuations in the case of a stock-out. In our application however, one cannot impose such assumption and the demand model follows a general price dependent curve while the customers arrive randomly and are served according to a first-come-first-serve logic.

The remainder of the chapter is structured as follows. In Section 5.3, we describe the model. In Section 5.4, we consider both additive and multiplicative demand with pricing (price setter model), analyze special cases and finally study the effect of demand uncertainty on consumer surplus. In Section 5.5, we study the supply-chain coordination and Section 5.6 considers a different mechanism where the government subsidizes the manufacturer's cost. Finally, Section 5.7 presents some computational

results and our conclusions are reported in Section 5.8. The proofs of the different propositions and theorems are relegated to the Appendix.

### 5.3 Model

We model the problem as a two-stage Stackelberg game where the government is the leader and the supplier is the follower (see Figure 5-1). We assume a single time period model with a unique supplier and consider a full information setting. The government decides the subsidy level  $r$  per product and the supplier follows by setting the price  $p$  and production quantities  $q$  to maximize his/her profit. The subsidy  $r$  is offered from the government directly to the end consumer. We consider a general stochastic demand function that depends on the effective price paid by consumers,  $z = p - r$ , and on a random variable  $\epsilon$ , denoted by  $D(z, \epsilon)$ . Once demand is realized, the sales level is determined by the minimum of supply and demand, that is,  $\min(q, D(z, \epsilon))$ . It should be noted that this single period model is particularly suitable for policies with short time horizon, such as a one year. For policies with longer time horizons, a dynamic model with prices, quantities and subsidies changing over time would perhaps be more realistic. For the purpose of studying the impact of demand uncertainty, it is sufficient to look at a single period without the added complexity of time dynamics.

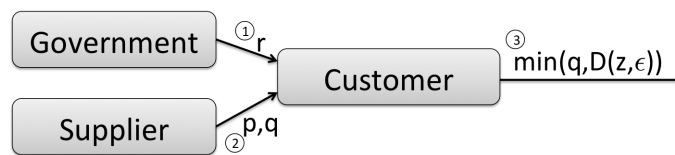


Figure 5-1: Order of events: 1. Subsidies; 2. Price and Quantity; 3. Sales

The selling price  $p$  can be viewed as the manufacturer’s suggested retail price (MSRP) that is, the price the manufacturer recommends for retail. Additionally, in industries where production lead time is long and incurs large fixed costs, we consider the production quantities to be equivalent to the capacity investment built in the manufacturing facility.

The goal of our model is to study the overall impact of demand uncertainty. In order to isolate this effect, we consider a single period monopolist model. These modeling assumptions are reasonable approximations for the Chevy Volt, which we use in our numerical analysis. Note that since the introduction of electric vehicles, the MSRP for the Chevy Volt and the subsidy level have remained fairly stable. Consumer subsidies were posted before the introduction of these products and have remained unchanged (\$7,500) since it was launched in December 2010. We assume the supplier is aware of the amount of subsidy offered to consumers before starting production. The supplier modeling choice is motivated by the fact that consumer subsidies for EVs started at a time where very few competitors were present in the market and the product offerings were significantly different. The Chevy Volt is an extended-range mid-priced vehicle, while the Nissan Leaf is a cheaper all-electric alternative and the Tesla Roadster is a luxury sports car. These products are also significantly different from traditional gasoline engine vehicles so that they can be viewed as price setting firms within their own niche markets.

Given a marginal unit cost,  $c$ , and consumer subsidy level,  $r$ , announced by the government, the supplier faces the following profit maximization problem.

$$\Pi = \max_{q,p} p \cdot \mathbb{E}[\min(q, D(z, \epsilon))] - c \cdot q \quad (5.3.1)$$

Denote  $\Pi$  as the optimal expected profit of the supplier. Note that the marginal cost  $c$  may incorporate both the manufacturing cost (such as material and labor) as well as the cost of building an additional unit of manufacturing capacity. Depending on the application setting, the cost of building capacity can be more significant than the per-unit manufacturing cost. If there is no demand information gained between the building of capacity and the production stage, then capacity is built according to the planned production. Therefore, we can assume both these costs to be combined in  $c$ . Furthermore, we can extend the model to incorporate a salvage value  $v$  for each unsold unit, such that  $v < c$ , or an underage cost  $u$  as a penalty for unmet

demand. These extensions do not qualitatively affect our results. They simply shift the newsvendor production quantile. To keep the exposition simple, we assume that salvage value and underage cost are both zero:  $v = u = 0$ .

We consider the general case for which the supplier decides upon both the price (MSRP) and the production quantities (i.e., the supplier is a price setter). An alternative case of interest is the one for which the price is exogenously given (i.e., the supplier is a price taker and decides only production quantity). As mentioned before, we consider the early stages of the EV market as a good application of the monopolist price setting model. It should be noted that a similar analysis can be done for the simpler price-taker model, which might be more appropriate in a different context.

We assume the government is introducing consumer subsidies,  $r$ , in order to stimulate sales to reach a given adoption target. We denote by  $\Gamma$  the target adoption level, which is assumed to be common knowledge. Conditional on achieving this target in expectation, the government wants to minimize the total cost of the subsidy program. Define  $Exp$  as the minimal expected subsidy expenditures, which is defined through the following optimization problem:

$$\begin{aligned}
 Exp = \min_r & \quad r \cdot \mathbb{E}[\min(q, D(z, \epsilon))] \\
 \text{s.t.} & \quad \mathbb{E}[\min(q, D(z, \epsilon))] \geq \Gamma \\
 & \quad r \geq 0
 \end{aligned} \tag{5.3.2}$$

In what follows, we discuss the modeling choices for the government in more detail.

**Government’s constraints** The adoption level constraint used in (5.3.2) is motivated by real policy-making practice. For example, President Obama stated the adoption target of 1 million of electric vehicles by 2015 (see [28]). More precisely, the government is interested in designing consumer subsidies so as to achieve the predetermined adoption target. An additional possibility is to incorporate a budget constraint for the government in addition to the adoption target. In various practical settings, the government may consider both requirements (see for example [25]).

Incorporating a budget constraint in our setting does not actually affect the optimal subsidy solution of the government problem (assuming the budget does not make the problem infeasible). In addition, one can show that there exists a one-to-one correspondence between the target adoption level and the minimum budget necessary to achieve this target. Hence, we will only solve the problem with a target adoption constraint, but the problem could be reformulated as a budget allocation problem with similar insights.

Given that actual sales are stochastic, the constraint used in our model meets the adoption target in expectation:

$$\mathbb{E}[\min(q, D(z, \epsilon))] \geq \Gamma. \quad (5.3.3)$$

Our results can be extended to the case where the government aims to achieve a target adoption level with some desired probability (chance constraint) instead of an expected value constraint. Such a modeling choice will be more suitable when the government is risk-averse and is given by:

$$\mathbb{P}(\min(q, D(z, \epsilon)) \geq \Gamma) \geq \Delta. \quad (5.3.4)$$

$\Delta$  represents the level of conservatism of the government. For example, when  $\Delta = 0.99$ , the government is more conservative than when  $\Delta = 0.9$ . We note that the insights we gain are similar for both classes of constraints (5.3.3) and (5.3.4) and therefore in the remainder of this chapter, we focus on the case of an expected value constraint.

Note that one can consider a constraint on greenhouse gas reduction instead of an adoption target. If the government has a desired target on emissions reduction, it can be translated to an adoption target in EV sales, for example. In particular, one can compute the decrease in carbon emissions between a gasoline car and an electric vehicle (see e.g., [5]). In other words, the value of  $\Gamma$  is directly tied to a value of a carbon emissions reduction target. With such a constraint on the expected sales

amount, our results remain valid. More generally, if we set a target on any increasing function of sales, the results also remain the same.

**Government's objective** Two common objectives for the government are to minimize expenditures or to maximize the welfare in the system. In the former, the government aims to minimize only its own expected expenditures, given by:

$$Exp = r \cdot \mathbb{E}[\min(q, D(z, \epsilon))]. \quad (5.3.5)$$

Welfare can be defined as the sum of the expected supplier's profit (denoted by  $\Pi$  and defined in (5.3.1)) and the consumer surplus (denoted by  $CS$ ) net the expected government expenditures:

$$W = \Pi + CS - Exp. \quad (5.3.6)$$

The consumer surplus is formally defined in Section 5.4.3 and aims to capture the consumer satisfaction. Interestingly, one can show that under some mild assumptions, both objectives are equivalent and yield the same optimal subsidy policy for the government. The result is summarized in the following Proposition 5.1.

**Proposition 5.1.** *Assume that the total welfare is a concave and unimodal function of the subsidy  $r$ . Then, there exists a threshold value  $\Gamma^*$  such that for any given value of the target level above this threshold, i.e.,  $\Gamma \geq \Gamma^*$ , both problems are equivalent.*

*Proof.* Since the welfare function is concave and unimodal, there exists a unique optimal unconstrained maximizer solution. If this unconstrained solution satisfies the adoption level target, the constrained problem is solved to optimality. However, if the target adoption level  $\Gamma$  is large enough, this solution is not feasible with respect to the adoption constraint. Since the expected sales increase with respect to  $r$ , one can see that the optimal solution of the constrained welfare maximization problem is obtained when the adoption level constraint is exactly met. Otherwise, by considering a larger subsidy level, one still satisfies the adoption constraint but does not increase

the welfare. Consequently, both problems are equivalent and yield the same optimal solution for which the adoption constraint is exactly met.  $\square$

In conclusion, if the value of the target level  $\Gamma$  is sufficiently large, both problems (minimizing expenditures in (5.3.5) and maximizing welfare in (5.3.6)) are equivalent. Note that the concavity and unimodality assumptions are satisfied for various demand models including the linear demand function. In particular, for linear demand, the threshold  $\Gamma^*$  can be characterized in closed form and is equal to twice the optimal production with zero subsidy and therefore satisfied in most reasonable settings. Furthermore, for smaller adoption target levels, [23] shows that even for multiple products in a competitive environment, the gaps between both settings (minimizing expenditures versus maximizing welfare) are small (if not zero) so that both problems yield solutions that are close to one another. For the remainder of the chapter, we assume the government objective is to minimize expenditures, while satisfying an expected adoption target, as in (5.3.2). This modeling choice was further motivated by private communications with sponsors of the MIT Energy Initiative.

Besides minimizing the subsidy cost, another objective for the government is often to maximize the positive environmental externalities of the green technology product. Assume there is a positive benefit, denoted by  $p_{CO_2}$ , for each ton of  $CO_2$  emission avoided by each unit sold of the green product. By introducing a monetary value to emissions, one can consider a combined government objective of minimizing the subsidy program cost, minus the benefit of emission reduction, i.e.,

$$r \cdot \mathbb{E}[\min(q, D(z, \epsilon))] - p_{CO_2} \cdot \mathbb{E}[\min(q, D(z, \epsilon))]. \quad (5.3.7)$$

Similarly to Proposition 5.1, we can show that if there is an adoption target larger than a certain threshold  $\bar{\Gamma}$ , then the objectives in (5.3.5) and (5.3.7) are equivalent and yield the same outcomes. In particular, if  $\Gamma \geq \bar{\Gamma}$ , the optimal subsidy policy will be defined by the tightness of the adoption target constraint. Alternatively, if the subsidy level  $\bar{r}$  required to reach  $\Gamma$  is significantly larger than the price of carbon,

$p_{CO_2}$ , the optimal subsidy is defined by the target constraint. Given a certain adoption target level  $\Gamma$ , there is a threshold level  $\bar{p}_{CO_2}$  such that for any price of carbon below this level,  $p_{CO_2} \leq \bar{p}_{CO_2}$ , the optimal subsidy policy is defined by the adoption target constraint. We next show that for an EV such as the Chevy Volt, a conservative estimate for the price of carbon emission mitigated for each EV sold is much lower than this threshold level.

As we mentioned, the positive externalities of EVs correspond to the reductions in CO<sub>2</sub> emissions throughout their lifetime, converted to US dollars. Following the analysis in [5], the emission rate per unit of energy amounts to 755 [Kg CO<sub>2</sub> × MWh<sup>-1</sup>]. By using the calculations in [23], an estimate for an EV gas emission reduction is about 50.3 [Ton CO<sub>2</sub>]. In order to convert this number to US dollars, we use the value assigned by the United States Environmental Protection Agency (EPA) to a Ton of CO<sub>2</sub>. The value for 2014 is 23.3 [\$ × (Ton CO<sub>2</sub>)<sup>-1</sup>] so that the monetary positive externality of an EV is equal to \$1,172. One can see that the positive externality for an EV is smaller than the consumer subsidies (equal to \$7,500) and therefore, it is sufficient to minimize expenditures, as described in (5.3.5).

With the formal definitions of the optimization problems faced by the supplier (5.3.1) and the government (5.3.2), in the next section we analyze of the optimal decisions of each party and the impact of demand uncertainty.

## 5.4 The Model

For products such as electric vehicles, where there are only a few suppliers in the market, it is reasonable to assume that the selling price (MSRP) of the product is an endogenous decision of the firm. In other words,  $p$  is a decision variable chosen by the supplier in addition to the production quantity  $q$ . In this case, the supplier's optimization problem can be viewed as a price setting newsvendor problem (see e.g., [49]). Note though that in our problem the solution also depends on the government subsidy. In particular, both  $q$  and  $p$  are decision variables that should be optimally

chosen by the supplier for each value of the subsidy  $r$  set by the government. To keep the analysis simple and to be consistent with the literature, we consider separately the cases of a stochastic demand with additive or multiplicative uncertainty. In each case, we first consider general demand functions and then specialize to linear and iso-elastic demand models that are common in the literature. Finally, we compare our results to the case where demand is approximated by a deterministic average value and draw conclusions about the cost of ignoring demand uncertainty.

In practice, companies very often ignore demand uncertainty and consider average values when taking decisions such as price and production quantities. As a result, we are interested in understanding how the optimal subsidy levels, prices and production quantities are affected when we explicitly consider demand uncertainty relative to the case when demand is just approximated by its deterministic average value. For example, the comparison may be useful to quantify the value of investing some large efforts in developing better demand forecasts.

We next present the analysis for both additive and multiplicative demand uncertainty.

### 5.4.1 Additive Noise

Define additive demand uncertainty as follows:

$$D(z, \epsilon) = y(z) + \epsilon. \tag{5.4.1}$$

Here,  $y(z) = \mathbb{E}[D(z, \epsilon)]$  is a function of the effective price  $z = p - r$  and represents the nominal deterministic part of demand and  $\epsilon$  is a random variable with cumulative distribution function (CDF)  $F_\epsilon$ .

*Assumption 8.* We impose the following conditions on demand:

- Demand depends only on the difference between  $p$  and  $r$  denoted by  $z$ .
- The deterministic part of the demand function  $y(z)$  is positive, twice differen-

tiable and a decreasing function of  $z$  and hence invertible.

- When  $p = c$  and  $r = 0$  the target level cannot be achieved, i.e.,  $y(c) < \Gamma$ .
- The noise  $\epsilon$  is a random variable with zero mean:  $\mathbb{E}[\epsilon] = 0$ .

We consider that the demand function represents the aggregate demand for all the consumers in the market during the entire horizon. As a result, the assumption that the target level cannot be achieved if the product is sold at cost and there are no subsidies translates to the fact that the total number of consumers will not reach the desired adoption target level without government subsidies. Under Assumption 8, we characterize the solution of problems (5.3.1) and (5.3.2) sequentially. First, we solve the optimal quantity  $q^*(p, r)$  and price  $p^*(r)$  offered by the supplier as a function of the subsidy  $r$ . By substituting the optimal solutions of the supplier problem, we can solve the government problem defined in (5.3.2). Note that problem (5.3.2) is not necessarily convex, even for very simple instances, because the government needs to account for the supplier's best response  $p^*(r)$  and  $q^*(p, r)$ . Nevertheless, one can still solve this using the tightness of the target adoption constraint. Because of the non-convexity of the problem, the tightness of the constraint cannot be trivially assumed. We formally prove the constraint is tight at optimality in Theorem 5.1. Using this result, we obtain the optimal subsidy of the stochastic problem (5.3.2), denoted by  $r_{sto}$ . The resulting optimal decisions of price and quantity are denoted by  $p_{sto} = p^*(r_{sto})$  and  $q_{sto} = q^*(p_{sto}, r_{sto})$ . From problems (5.3.1) and (5.3.2), the optimal profit of the supplier is denoted by  $\Pi_{sto}$  and government expenditures by  $Exp_{sto}$ .

We consider problems (5.3.1) and (5.3.2), where demand is equal to its expected value, that is:  $\mathbb{E}[D(z, \epsilon)] = y(z)$ . We denote this deterministic case with the subscript "det", with optimal values:  $r_{det}, p_{det}, q_{det}, z_{det}, \Pi_{det}, Exp_{det}$ . We next compare these metrics in the deterministic versus stochastic case.

**Theorem 5.1.** *Assume that the following condition is satisfied:*

$$2y'(z) + (p - c) \cdot y''(z) + \frac{c^2}{p^3} \cdot \frac{1}{f_\epsilon \left( F_\epsilon^{-1} \left( \frac{p-c}{p} \right) \right)} < 0. \quad (5.4.2)$$

*The following holds:*

1. *The optimal price of problem (5.3.1) as a function of  $r$  is the solution of the following non-linear equation:*

$$y(p - r) + \mathbb{E} \left[ \min \left( F_\epsilon^{-1} \left( \frac{p - c}{p} \right), \epsilon \right) \right] + y'(p - r) \cdot (p - c) = 0. \quad (5.4.3)$$

*In addition, using the solution from (5.4.3), one can compute the optimal production quantity:*

$$q^*(p, r) = y(p - r) + F_\epsilon^{-1} \left( \frac{p - c}{p} \right). \quad (5.4.4)$$

2. *The optimal solution of the government problem is obtained when the target adoption level is exactly met.*
  
3. *The optimal expressions follow the following relations:*

$$\begin{aligned} z_{sto} = y^{-1}(\Gamma - K(p_{sto})) &\leq z_{det} = y^{-1}(\Gamma) \\ q_{sto} = \Gamma + F_\epsilon^{-1} \left( \frac{p_{sto} - c}{p_{sto}} \right) - K(p_{sto}) &\geq q_{det} = \Gamma, \end{aligned}$$

*where  $K(p_{sto})$  is defined as:*

$$K(p_{sto}) = \mathbb{E} \left[ \min \left( F_\epsilon^{-1} \left( \frac{p_{sto} - c}{p_{sto}} \right), \epsilon \right) \right]. \quad (5.4.5)$$

If, in addition, the function  $y(z)$  is convex:

$$\begin{aligned} p_{sto} = c + \frac{\Gamma}{|y'(z_{sto})|} &\leq p_{det} = c + \frac{\Gamma}{|y'(z_{det})|} \\ \Pi_{sto} = \frac{\Gamma^2}{|y'(z_{sto})|} - c \cdot (q_{sto} - \Gamma) &\leq \Pi_{det} = \frac{\Gamma^2}{|y'(z_{det})|} \end{aligned}$$

**Remark 5.1.** Note that for a general function  $y(p - r)$ , one cannot derive a closed form solution of (5.4.3) for  $p^*(r)$ . Consequently, one cannot find a closed form expression for the optimal price  $p_{sto}$  for a general additive demand. This is consistent with the fact that there does not exist a closed form solution for the price-setting newsvendor. However, one can use (5.4.3) to characterize the optimal solution and even numerically compute the optimal price by using a binary search method (see more details in the Appendix).

Similarly, one cannot generally express  $K(p_{sto})$  in closed form. Instead,  $K(p_{sto})$  represents a measure of the magnitude of the noise that depends on the price  $p_{sto}$  and the noise distribution.  $K(p_{sto})$  is mainly used to draw insights on the impact of demand uncertainty on the optimal decision variables.

Assumption (5.4.2) guarantees the uniqueness of the optimal price as a function of the subsidies, as it implies the strict concavity of the profit function with respect to  $p$ . In case this condition does not hold, problem (5.3.1) is still numerically tractable (see [49]). For the remainder of this chapter, we will assume condition (5.4.2) is satisfied. For the case of linear demand we discuss relation (5.4.8) which is a sufficient condition that is satisfied in many reasonable settings.

Note that the optimal ordering quantity in (5.4.4) is expressed as the expected demand plus the optimal newsvendor quantile  $(p - c)/p$  related to the demand uncertainty. The government can ensure that the expected sales achieve the desired target adoption level  $\Gamma$  by controlling the effective price  $z$ . When demand is stochastic, in order to achieve an expected sales of  $\Gamma$ , the government must encourage the supplier to produce a higher quantity than  $\Gamma$  to compensate for the demand scenarios where

stock-outs occur. The additional production level is captured by  $K(p_{sto})$ .

The optimal price  $p$  is characterized by the optimality condition written in (5.4.3) that depends on the cost and on the price elasticity evaluated at that optimal price  $p$ , denoted by  $E_d(p)$ . This can be rewritten so that the marginal cost equals the marginal revenue, i.e.,  $c = p\left(1 - 1/E_d(p)\right)$ . Even without knowing a closed form expression for the optimal price, we can still show that the optimal price decreases in the presence of demand uncertainty and so does the firm profit.

**Remark 5.2.** *The results of Theorem 5.1 can be generalized to describe how the optimal variables (i.e.,  $z, q, p$  and  $\Pi$ ) change as demand uncertainty increases. Instead of comparing the stochastic case to the deterministic case (i.e., where there is no demand uncertainty), one can instead consider how the optimal variables vary in terms of the magnitude of the noise (for more details, see the proof of Theorems 5.1 in the Appendix). In particular, the quantity that captures the effect of demand uncertainty is  $K(p_{sto})$ .*

*Since the noise  $\epsilon$  has zero-mean, the quantity  $K(p_{sto})$  in (5.4.5) is always non-positive. In addition, when there is no noise (i.e.,  $\epsilon = 0$  with probability 1),  $K(p_{sto}) = 0$  and the deterministic scenario is obtained as a special case. For any intermediate case,  $K(p_{sto})$  is negative and non-increasing with respect to the magnitude of the noise. For example, if the noise  $\epsilon$  is uniformly distributed, the inverse CDF function can be written as a linear function of the standard deviation  $\sigma$  as follows:*

$$F_\epsilon^{-1}\left(\frac{p-c}{p}\right) = \sigma\sqrt{3} \cdot \left(2 \cdot \frac{p-c}{p} - 1\right). \quad (5.4.6)$$

*Therefore,  $K(p_{sto})$  scales monotonically with the standard deviation for uniform demand uncertainty. In other words, all the comparisons of the optimal variables (e.g., effective price, production quantities etc) are monotonic functions of the standard deviation of the noise. This result is true for a large class of common distributions that can be parameterized by the standard deviation, such as uniform, normal, and exponential. As a result, one can extend our insights in a continuous fashion with*

respect to the magnitude of the noise. For example, the inequality of the effective price is given by:  $z_{sto} = y^{-1}(\Gamma - K(p_{sto}))$ . This equation is non-increasing with respect to the magnitude of  $K(p_{sto})$  and is maximized when there is no noise (deterministic demand) so that  $z_{det} = z_{sto} = y^{-1}(\Gamma)$ . In general, as the magnitude of the noise increases, the gaps between the optimal decision variables increase (see plots of optimal decisions as functions of the standard deviation of demand uncertainty in Figure 5-4 of Section 5.7). For more general demand distributions, the relationship with the standard deviation is not as simple. The quantile  $F_{\epsilon}^{-1}\left(\frac{p-c}{p}\right)$  may not move monotonically with the standard deviation, for example with non-unimodal distributions.

**Remark 5.3.** *The solution of the optimal quantity  $q$  and the effective price  $z$  provide another interesting insight. Theorem 5.1 states that when demand is uncertain, the consumers are better off in terms of effective price and production quantities (this is true for any decreasing demand function). Furthermore, the selling price and the profit of the supplier are lower in the presence of uncertainty, assuming demand is convex. These results imply the consumers are in general better off when demand is uncertain. Nevertheless, as we will show in Section 5.4.3, this is not always the case when we use the aggregate consumer surplus as a metric.*

By focusing on a few demand functions, we can provide additional insights. We will first consider the linear demand case, which is the most common in the literature. The simplicity of this demand form enables us to derive closed-form solutions and a deeper analysis of the impact of demand uncertainty. Note that the insights can be quite different for non-linear demand functions. The results presented in Theorem 5.1 justify the need for considering non-linear functions as well. For this reason, we later consider the iso-elastic demand case and compare it to the linear case.

The impact of demand uncertainty on the subsidy level  $r$  and on the government expenditure is harder to observe for a general demand form. In order to explore this further, we focus on the cases of linear and iso-elastic demands. For both cases, we show that the subsidy increases with the added inventory risk captured by  $K(p_{sto})$ .

## Linear Demand

In what follows, we quantify the effect of demand uncertainty on the subsidy level and the expected government expenditures. We can obtain such results for specific demand models, among them the linear demand model. Define the linear demand function as:

$$D(z, \epsilon) = \bar{d} - \alpha \cdot z + \epsilon, \quad (5.4.7)$$

where  $\bar{d}$  and  $\alpha$  are given positive parameters that represent the maximal market share and the price elasticity respectively. Note that for this model, a sufficient condition for assumption (5.4.2) to hold is given by:

$$\alpha > \frac{1}{2c \cdot \inf_x f_\epsilon(x)}.$$

For example, if the additive noise is uniformly distributed, i.e.,  $\epsilon \sim U[-a_2, a_2]$ ;  $a_2 > 0$ , (note that since the noise is uniform with zero mean, it has to be symmetric) we obtain:

$$\alpha > \frac{a_2}{c}. \quad (5.4.8)$$

One can see that by fixing the cost  $c$ , condition in (5.4.8) is satisfied if the price elasticity  $\alpha$  is large relative to the standard deviation of the noise. Next, we derive closed form expressions for the optimal price, production quantities, subsidies, profit and expenditures for both deterministic and stochastic demand models and compare the two settings.

**Theorem 5.2.** *The closed form expressions and comparisons for the linear demand*

model in (5.4.7) are given by:

$$\begin{aligned}
p_{sto} &= c + \frac{\Gamma}{\alpha} = p_{det} \\
q_{sto} &= \Gamma + F_{\epsilon}^{-1}\left(\frac{p_{sto} - c}{p_{sto}}\right) - K(p_{sto}) \geq q_{det} = \Gamma \\
r_{sto} &= \frac{2\Gamma}{\alpha} + c - \frac{\bar{d}}{\alpha} - \frac{1}{\alpha} \cdot K(p_{sto}) \geq r_{det} = \frac{2\Gamma}{\alpha} + c - \frac{\bar{d}}{\alpha} \\
\Pi_{sto} &= \frac{\Gamma^2}{\alpha} - c \cdot (q_{sto} - \Gamma) \leq \Pi_{det} = \frac{\Gamma^2}{\alpha} \\
Exp_{sto} &= \Gamma \cdot r_{sto} \geq Exp_{det} = \Gamma \cdot r_{det}
\end{aligned}$$

We note that the results of Theorem 5.2 can be presented in a more general continuous fashion as explained in Remark 5.2. Surprisingly, the optimal price is the same for both the deterministic and stochastic models. In other words, the optimal selling price is not affected by demand uncertainty for linear demand. On the other hand, with increased quantities, the expected profit of the supplier is lower under demand uncertainty. At the same time, the optimal subsidy level and expenditures increase with uncertainty. Therefore, both supplier and government are worse off when demand is uncertain. Corollary 5.1 and the following discussion provide further intuition in how this cost of demand uncertainty is shared between the supplier and the government.

**Corollary 5.1.**

1.  $q_{sto} - q_{det}$  decreases in  $c$  and increases in  $\Gamma$ .
2.  $r_{sto} - r_{det}$  increases in  $c$  and decreases in  $\Gamma$ .
3. Assume that  $\epsilon$  has support  $[a_1, a_2]$ . Then, the optimal subsidy for the stochastic and deterministic demands relate as follows:

$$r_{det} \leq r_{sto} \leq r_{det} + \frac{|a_1|}{\alpha}.$$

Corollary 5.1 can be better understood in terms of the optimal service level for

stochastic demand, denoted by  $\rho = \frac{p_{sto}-c}{p_{sto}}$ . Note that  $\rho$  is an endogenous decision of the supplier, which is a function of the optimal price  $p_{sto}$ . For linear demand, the optimal service level can be simplified as:  $\rho = \frac{\Gamma}{c\alpha+\Gamma}$ . This service level is decreasing in the cost  $c$  but increasing with respect to the target adoption  $\Gamma$ .

On one hand, when the optimal price is significantly higher than the production cost, i.e.,  $p_{sto} \gg c$ , the high profit margin encourages the supplier to satisfy a larger share of demand by increasing its production. In this case, the supplier has incentives to overproduce and bear more of the inventory risk. The government may then set low subsidies, in fact the same as in the deterministic case, which guarantee that the average demand meets the target. On the other hand, when  $p_{sto}$  is close to  $c$  (low profit margin), the supplier has no incentives to bear any risk and produces quantities to match the lowest possible demand realization. In this case, the government will bear all the inventory risk by increasing the value of the subsidies.

Note that as production cost  $c$  increases, the required subsidy is larger for both stochastic and deterministic demands, meaning the average subsidy expenditure is higher. At the same time, the service level  $\rho$  decreases and, from Corollary 5.1, the gap between  $r_{sto}$  and  $r_{det}$  increases. The supplier's cost increase amplifies the cost of demand uncertainty for the government.

A similar reasoning can be applied to the target adoption level. As  $\Gamma$  increases, the overall cost of the subsidy program increases, as expected. Interestingly, the service level  $\rho$  also increases. From Corollary 5.1, the production gap between  $q_{sto}$  and  $q_{det}$  widens, while the subsidy gap between  $r_{sto}$  and  $r_{det}$  shrinks. Effectively, the burden of demand uncertainty is transferred from the government to the supplier as  $\Gamma$  increases. This means that a higher target adoption will induce the product to be more profitable. This will make the supplier take on more of the inventory risk and consequently switching who bears the cost of demand uncertainty.

Corollary 5.1.3 shows that the subsidy decision is bounded by the worst case demand realization normalized by the price sensitivity. In other words, it provides a guarantee on the gap between the subsidies for stochastic and deterministic demands.

In conclusion, by studying the special case of a linear demand model, we obtain the following additional insights: (i) The optimal price does not depend on demand uncertainty. (ii) The optimal subsidy set by the government increases with demand uncertainty. Consequently, the introduction of demand uncertainty decreases the effective price paid by consumers. In addition, the government will spend more when demand is uncertain. (iii) The cost of demand uncertainty is shared by the government and the supplier and depends on the profit margin (equivalently, service level) of the product. As expected, lower/higher margins mean the supplier takes less/more inventory risk. Therefore, increasing the adoption target or decreasing the manufacturing cost will shift the cost of demand uncertainty from the government to the supplier.

## 5.4.2 Multiplicative Noise

In this section, we consider a demand with a multiplicative noise (see for example, [37]). The nominal deterministic part is assumed to be a function of the effective price, denoted by  $y(z)$ :

$$D(z, \epsilon) = y(z) \cdot \epsilon \tag{5.4.9}$$

*Assumption 9.* • Demand depends only on the difference between  $p$  and  $r$  denoted by  $z$ .

- The deterministic part of the demand function  $y(z)$  is positive, twice differentiable and a decreasing function of  $z$  and hence invertible.
- When  $p = c$  and  $r = 0$  the target level cannot be achieved, i.e.,  $y(c) < \Gamma$ .
- The noise  $\epsilon$  is a positive and finite random variable with mean equal to one:  $\mathbb{E}[\epsilon] = 1$ .

One can show that the results of Theorem 5.1 hold for both additive and multiplicative demand models. The proof for multiplicative noise follows a similar method-

ology and is not repeated due to space limitations. We next consider the iso-elastic demand case to derive additional insights on the optimal subsidy.

Iso-elastic demand models are very popular in various application areas. In particular, a large number of references in economics consider such models (see, e.g., [55] and [50]) as well as revenue management (see [56]). Iso-elastic demand is also sometimes called the log-log model and the main property is that elasticities are constant for any given combination of price and quantities. In addition, it does not require to know a finite upper limit on price. For more details, see for example [38]. Various papers in oligopoly competition consider iso-elastic demand (see, e.g., [43], [9] and [52]). In [52], the authors study the dynamics of two competing firms in a market in terms of Cournot's duopoly theory. In [43], the authors consider (among others) an iso-elastic demand in a multi-echelon inventory/pricing setting and show that the results might differ depending on the demand shape. Another application that uses iso-elastic demands relates to commodity pricing (see, e.g., [26]). Finally, practitioners and researchers have used iso-elastic demand models for products in retail such as groceries, fashion (see, e.g., [50], [17] and [4]) and gasoline (e.g., [11]).

## **Iso-Elastic Demand**

Define the iso-elastic demand as:

$$y(z) = \bar{d} \cdot z^{-\alpha} \quad (\alpha > 1). \quad (5.4.10)$$

The iso-elastic model considered in the literature usually assumes that  $\alpha > 1$  in order to satisfy the Increasing Price Elasticity (IPE) property (see, e.g., [62]). Note that the function  $y(z)$  is convex with respect to  $z$  for any value  $\alpha > 1$ . Therefore, the results from Theorem 5.1 hold. Using this particular demand structure, we obtain the following additional results on the optimal subsidy.

**Proposition 5.2.** *For the iso-elastic demand model in (5.4.10), we have:*

$$r_{sto} \geq r_{det}.$$

We note that the result of Proposition 5.2 can be presented in a continuous fashion, as explained in Remark 5.2. Note also that this allows us to recover the same results as the linear additive demand model regarding the impact of demand uncertainty on the subsidies. These two cases show that the subsidy increases with demand uncertainty.

### 5.4.3 Consumer Surplus

In this section, we study the effect of demand uncertainty on consumers using consumer surplus as a metric. For that purpose, we compare the aggregate level of consumer surplus under stochastic and deterministic demand models. The consumer surplus is an economic measure of consumer satisfaction calculated by analyzing the difference between what consumers are willing to pay and the market price. For a general deterministic price demand curve, the consumer surplus is denoted by  $CS_{det}$  and can be computed as the area under the demand curve above the market price (see, e.g., [60]):

$$CS_{det} = \int_0^{q_{det}} \left( D^{-1}(q) - z_{det} \right) dq = \int_{z_{det}}^{z_{max}} D(z) dz. \quad (5.4.11)$$

We note that in our case, the market price is equal to the effective price paid by consumers  $z = p - r$ . Denote,  $D^{-1}(q)$  as the effective price that will generate demand exactly equal to  $q$ . Note that  $z_{det}$  and  $q_{det}$  represent the optimal effective price and production, whereas  $z_{max}$  corresponds to the value of the effective price that yields zero demand. The consumer surplus represents the surplus induced by consumers that are willing to pay more than the posted price.

When demand is uncertain however, defining the consumer surplus (denoted by  $CS_{sto}$ ) is somewhat more subtle due to the possibility of a stock-out. Several papers

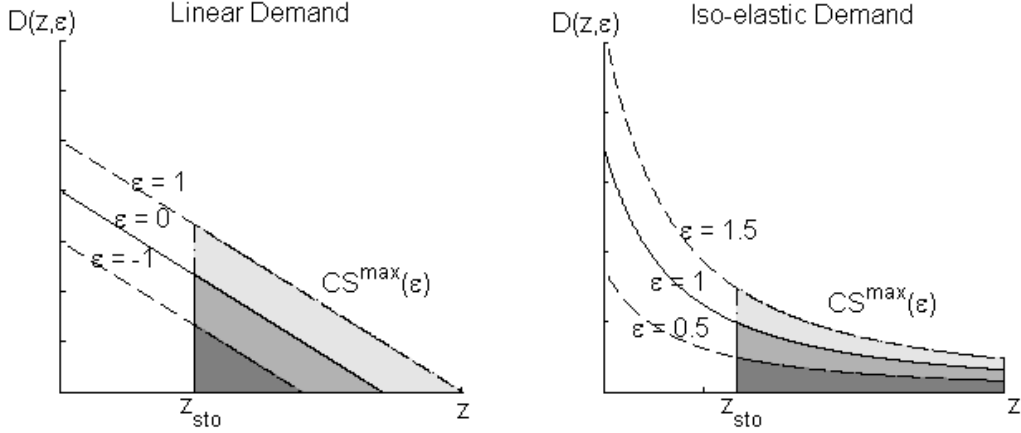


Figure 5-2: Consumer surplus for stochastic demand, before demand rationing

on peak load pricing and capacity investments by a power utility under stochastic demand address partially this modeling issue (see [19], [24] and [13]). Nevertheless, the models developed in this literature are not applicable to the price setting newsvendor. More specifically, in [13] the authors assume that the utility power facility has access to the willingness to pay of the customers so that it can decline the ones with the lowest valuations. This assumption is not justifiable in our setting where a “first-come-first-serve” logic with random arrivals is more suitable. In [48], the authors study a price setting newsvendor model for public goods and consider the consumer surplus for linear additive stochastic demand.

For general stochastic demand functions, the consumer surplus  $CS_{sto}(\epsilon)$  is defined for each realization of demand uncertainty  $\epsilon$ . If there was no supply constraint, considering the effective price and the realized demand, the total amount of potential consumer surplus is defined as:

$$CS^{max}(\epsilon) = \int_{z_{sto}}^{z_{max}(\epsilon)} D(z, \epsilon) dz$$

Figure 5-2 displays the area under the demand curves (linear and iso-elastic) that defines the maximum consumer surplus  $CS^{max}(\epsilon)$ , for a given demand realization  $\epsilon$ . Note that the actual consumer surplus will be a fraction of this maximum surplus,

based on the fraction of customers that are actually served. Since customers are assumed to arrive in a first-come-first-serve manner, irrespective of their willingness to pay, under certain demand realizations, some proportion of these customers will not be served due to stock-outs. The proportion of served customers under one of these demand realizations is given by the ratio of actual sales over potential demand:  $\frac{\min(D(z_{sto}, \epsilon), q_{sto})}{D(z_{sto}, \epsilon)}$ . Therefore, the consumer surplus can be defined as the total available surplus times the proportion of that surplus that is actually served.

$$CS_{sto}(\epsilon) = CS^{max}(\epsilon) \cdot \frac{\min(D(z_{sto}, \epsilon), q_{sto})}{D(z_{sto}, \epsilon)}. \quad (5.4.12)$$

We note that in this case, the consumer surplus is a random variable that depends on the demand through the noise  $\epsilon$ . Note that we are interested in comparing  $CS_{det}$  to the expected consumer surplus  $\mathbb{E}_\epsilon[CS_{sto}(\epsilon)]$ . For stochastic demand, (5.4.12) has a similar interpretation as its deterministic counterpart. Nevertheless, we also incorporate the possibility that a consumer who wants to buy the product does not find it available. As we will show, the effect of demand uncertainty on consumer surplus depends on the structure of the nominal demand function. In particular, we provide the results for the two special cases we have considered in the previous section and show that the effect is opposite. For the linear demand function in (5.4.7), we have:

$$CS_{det} = \int_0^{q_{det}} (D^{-1}(q) - z_{det}) dq = \frac{q_{det}^2}{2\alpha} = \frac{\Gamma^2}{2\alpha}. \quad (5.4.13)$$

For iso-elastic demand from (5.4.10), we obtain:

$$CS_{det} = \int_0^{q_{det}} (D^{-1}(q) - z_{det}) dq = \frac{\bar{d}}{\alpha - 1} \cdot \left(\frac{\bar{d}}{\bar{\Gamma}}\right)^{\frac{1-\alpha}{\alpha}} \quad (\alpha > 1). \quad (5.4.14)$$

One can then show the following results regarding the effect of demand uncertainty on the consumer surplus for these two demand functions.

**Proposition 5.3.** *For the linear demand model in (5.4.7), we have:*

$$\mathbb{E}[CS_{sto}] \geq CS_{det}. \quad (5.4.15)$$

*For the iso-elastic demand model in (5.4.10) with  $\alpha > 1$ , we have:*

$$\mathbb{E}[CS_{sto}] \leq CS_{det}. \quad (5.4.16)$$

The consumer surplus result in Proposition 5.3 is perhaps one of the most counter-intuitive findings of this chapter. Proposition 5.3 shows that under linear demand, the expected consumer surplus is larger when considering demand uncertainty, whereas it is lower for the iso-elastic model. We already have shown in Theorem 5.1 that the effective price is lower and that the production quantities are larger when considering demand uncertainty relative to the deterministic model. In addition, this result was valid for both models (i.e., additive and multiplicative noises for linear and non-linear demand). As a result, demand uncertainty benefits overall the consumers in terms of effective price and available quantities. With this in mind, one could expect consumer surplus to increase with uncertainty. However, when comparing the consumer surplus using equation (5.4.12) for stochastic demand, we obtain that for the iso-elastic demand, consumers are in aggregate worse-off when demand is uncertain. On one hand, demand uncertainty benefits the consumers since it lowers the effective price and increase the quantities. On the other hand, demand uncertainty introduces a stock-out probability because some of the consumers may not be able to find the product available. These two factors (effective price and stock-out probability) affect the consumer surplus in opposite ways.

For iso-elastic demand, the second factor is dominant and therefore the consumer surplus is lower when demand is uncertain. In particular, the iso-elastic demand admits some consumers that are willing to pay a very large price. If these consumers experience a stock-out, it will reduce drastically the aggregate consumer surplus. In fact, the gap between the stochastic and deterministic consumer surplus widens

when  $K(p_{sto})$  is smaller. This happens when the profit margin is low, meaning there is more inventory risk for the supplier. For linear demand, the dominant factor is not the stock-out probability and consequently, the consumer surplus is larger when demand is uncertain. We note that this result is related to the structure of the nominal demand rather than the noise effect. For example, if we were to consider a linear demand with a multiplicative noise, we will have the same result as for the linear demand with additive noise.

Next, we compare and contrast our findings on production quantity, price and profit as well as consumer surplus against what is already known in the literature about the classical price setting newsvendor problem. This way, we can investigate the impact of incorporating the government as an additional player in the system. In the classical price setting newsvendor, there does not exist a closed form expression for the optimal price and production even for simple demand forms. However, one can still compare the outcomes between stochastic and deterministic scenarios. We compare the results of Theorems 5.1 and 5.2 to the classical price setting newsvendor (i.e., without the government). The optimal price, quantity and profit can be found in a similar way. First, one can show that the optimal price follows the same relation as in our work, i.e.,  $p_{sto} \leq p_{det}$ . Note that in the classical model,  $p$  is equivalent to the effective price paid by consumers and therefore, similar insights apply (see Theorem 5.1). However, the relation for the optimal quantity differs. More precisely, the inequality on quantity depends on the critical newsvendor quantile being larger or smaller than 1. For symmetric additive noises, if the profit margin is below 0.5 (this is usually the case for the EV industry), the supplier will not take the over stock risk and the optimal quantity decreases with respect to the magnitude of the noise. As a result, the optimal quantity relation will be opposite than the one we obtain in this work, where the government is an additional player in the supply chain. In addition, the results on optimal profits agree with our findings (again, assuming the profit margin is below 0.5) so that the expected profits for stochastic demand are lower relative to the case where demand is deterministic, as expected. In conclusion, the

effect of demand uncertainty for the classical price setting newsvendor (assuming the profit margin is below 0.5) states that quantity, price and profit are all lower when demand is stochastic. When comparing to Theorem 5.1, we first observe that the results do not depend on the profit margin. In addition, the optimal quantity follows the opposite relation, whereas the price and profit follow the same one. Therefore, in our setting, the government is bearing some uncertainty risk together with the supplier and incentivizes the supplier to over produce in order to make sure that the adoption target is achieved on expectation. Finally, one can do a similar analysis for the expected consumer surplus. However, the analysis is not straightforward and depends on the demand function, the structure of the noise (additive or multiplicative) and the capacity rationing rule (see [22]). Indeed, when the government is present in the supply chain, he/she can help increasing the production and consequently, inducing larger consumer surplus in expectation. Note that the profit is still lower, as the stochastic scenario remains more risky for the supplier.

## 5.5 Supply-chain Coordination

In this section, we examine how the results change in the case where the system is centrally managed. In this case, one can imagine that the government and the supplier take coordinated decisions together. The central planner needs to decide the price, the subsidy and the production quantities simultaneously. This situation may arise when the firm is owned by the government. We study the centrally managed problem as a benchmark to compare to the decentralized case developed in the previous sections. In particular, we are interested in understanding if the decentralization will have an adverse impact on either party and more importantly if it will hurt the consumers. We show in this section that this is not the case. In fact, the decentralized problem achieves the same outcome as the centralized problem and hence, government subsidies act as a coordinating mechanism as far as consumer are concerned. Supply chain coordination has been extensively studied in the literature. In

particular, some of the supply-chain contracting literature (see, e.g., [14]) discusses mechanisms that can be used to coordinate operational decisions such as price and production quantities.

Define the central planner's combined optimization problem to maximize the firm's profits minus government expenditures as follows:

$$\begin{aligned} \max_{q,z} \quad & z \cdot \mathbb{E}[\min(q, D(z, \epsilon))] - c \cdot q \\ \text{s.t.} \quad & \mathbb{E}[\min(q, D(z, \epsilon))] \geq \Gamma \end{aligned} \tag{5.5.1}$$

Note that in this case, we impose the additional constraint  $p \geq c$  so that the selling price has to be at least larger than the cost. Indeed, for the centralized version, it is not clear that this constraint is automatically satisfied by the optimal solution as it was in the decentralized setting. Our goal is to show how the centralized solutions for  $q, p$  and  $r$  compare to their decentralized counterparts from Section 5.4. We consider both deterministic and stochastic demand models and focus on additive uncertainty under Assumption 8.

**Theorem 5.3.** *The optimal effective price  $z = p - r$  and production level  $q$  are the same in both the decentralized and centralized models. Therefore, consumer subsidies are a sufficient mechanism to coordinate the government and the supplier.*

Note that for problem (5.5.1), one can only solve for the effective price and not  $p$  and  $r$  separately. In particular, there are multiple optimal solutions for the centralized case and the decentralized solution happens to be one them. If the government and the supplier collude into a single entity, this does not affect the consumers in terms of effective price and production quantities. Therefore, the consumers are not affected by the coordination. This result might be surprising as one could think that the coordination will add additional information and power to the central planner as well as mitigate some of the competition effects between the supplier and the government. However, in the original decentralized problem, the government acts as a quantity

coordinator in the sense that the optimal solutions in both cases are obtained by the tightness of the target adoption constraint.

## 5.6 Subsidizing the manufacturer's cost

In this section, we consider a different incentive mechanism where the government offers subsidies directly to the manufacturer (as opposed to the end consumers). In particular, our goals are (i) to study if the impact of demand uncertainty and most of our insights are preserved if the government were to use a cost subsidy mechanism and (ii) to compare the outcomes of both mechanisms. Offering subsidies directly to the manufacturer can be implemented by partially sharing the cost of production or in form of loans or free capital to the supplier. An example of this type of subsidy was the \$249 million federal grant provided to battery maker A123 under the American Recovery and Reinvestment Act to increase manufacturing of batteries for electric and hybrid vehicles<sup>4</sup>. This grant was later criticized after the company declared bankruptcy, along with another failed subsidy program to solar panel manufacturer Solyndra<sup>5</sup>. Knowing that this type of subsidy mechanism is also used in practice, in addition to consumer subsidies, we hope to do a similar analysis to observe the impact of demand uncertainty. Below, we formalize the model for this setting and provide the results. We then summarize our findings as well as compare both settings.

The government still seeks to encourage green technology adoption. Instead of offering rebates to the end consumers, the government provides a subsidy, denoted by  $s \geq 0$ , directly to the manufacturer. Note that this mechanism does not have a direct impact on the demand function that depends only on the selling price  $p$  and not explicitly on  $s$ . As before, the government leads the game by solving the following

---

<sup>4</sup>U.S. Li-ion battery production ramping up - 9/16/2010 - <http://articles.sae.org/8863/>

<sup>5</sup><http://www.bostonglobe.com/business/2012/10/16/bankruptcy-latest-troubled-obama-clean-energy-deals-from-solyndra-beacon-power/loeshQsK4xkudsgb9c77uN/story.html>

optimization problem:

$$\begin{aligned}
\min_s \quad & s \cdot q(p, s) \\
\text{s.t.} \quad & \mathbb{E}[\min(q, D(p, \epsilon))] \geq \Gamma \\
& s \geq 0
\end{aligned} \tag{5.6.1}$$

Note that in this case, the government subsidizes the total produced units instead of the total expected sold units as before.

Given a subsidy level  $s$  announced by the government, the supplier faces the following profit maximization problem. Note that  $c$  denotes the cost of building an additional unit of manufacturing capacity, as before.

$$\Pi = \max_{q,p} p \cdot \mathbb{E}[\min(q, D(p, \epsilon))] - (c - s) \cdot q \tag{5.6.2}$$

For simplicity, we assume a linear and additive demand but one can extend the results for non-linear models as well as for multiplicative uncertainty. However, to keep the analysis simple, we present the results for the linear case, given by:

$$D(p, \epsilon) = \bar{d} - \alpha \cdot p + \epsilon, \tag{5.6.3}$$

First, we study the impact of demand uncertainty on the decision variables for the cost subsidy mechanism (denoted by CSM). Second, we compare the outcomes for both mechanisms and elaborate on the differences. The results on the impact of demand uncertainty are summarized in the following Theorem.

**Theorem 5.4.** *Assume a linear demand as in (5.6.3). The comparisons for the cost*

subsidy mechanism are given by:

$$\begin{aligned}
p_{sto} &= \frac{\bar{d} - \Gamma}{\alpha} + \frac{1}{\alpha} \cdot K'_\epsilon \leq p_{det} = \frac{\bar{d} - \Gamma}{\alpha} \\
q_{sto} &= \Gamma + F_\epsilon^{-1}\left(\frac{p_{sto} - c + s_{sto}}{p_{sto}}\right) - K'_\epsilon \geq q_{det} = \Gamma \\
s_{sto} &= \frac{2\Gamma}{\alpha} + c - \frac{\bar{d}}{\alpha} - \frac{1}{\alpha} \cdot K'_\epsilon \geq s_{det} = \frac{2\Gamma}{\alpha} + c - \frac{\bar{d}}{\alpha} \\
\Pi_{sto} &= \frac{\Gamma^2}{\alpha} - (c - s_{sto}) \cdot (q_{sto} - \Gamma) \leq \Pi_{det} = \frac{\Gamma^2}{\alpha} \\
Exp_{sto} &= q_{sto} \cdot s_{sto} \geq Exp_{det} = \Gamma \cdot s_{det}
\end{aligned}$$

Where,  $K'_\epsilon = \mathbb{E}[\min(F_\epsilon^{-1}\left(\frac{p_{sto} - c + s_{sto}}{p_{sto}}\right), \epsilon)]$ . Note that  $s_{sto}$  is not given in a closed form expression as both sides depend on  $s_{sto}$  and one needs to solve a non-linear fixed point equation. The proof of Theorem 5.4 is not reported due to space limitations (it is based on a similar methodology as Theorem 5.2). One can see that the impact of demand uncertainty on all the decision variables is the same as in the consumer rebates mechanism (see Theorem 5.2). Next, we compare the outcomes of both mechanisms under deterministic and stochastic demands. The results are summarized below:

- The amount of subsidy (per unit) paid by the government follows the following relation:

$$s_{sto} \leq r_{sto}; \quad s_{det} = r_{det}.$$

- The price paid by the consumers follows the following relation:

$$p_{sto}^{CSM} \geq p_{sto} - r_{sto}; \quad p_{det}^{CSM} = p_{det} - r_{det}.$$

- The production quantities follows the following relation:

$$q_{sto}^{CSM} \geq q_{sto}; \quad q_{det}^{CSM} = q_{det}.$$

- For the profit of the supplier and the government expenditures, one cannot find the relation analytically. Instead, we study and compare the expressions computationally (see discussion below).

Note that for the cost subsidy mechanism the price paid by consumers is equal to  $p$  (since there is no subsidy to consumers), whereas for the consumer subsidy mechanism, it is captured by  $p - r$ . We observe that the government can save money on the per unit subsidy but it does not mean that the overall expenditures are lower as more units are potentially subsidized. In addition, the consumers are paying a larger price to compensate this government saving per unit. As a result, the consumers are worse off in terms of price but better off in terms of available quantities. In addition, when the demand is deterministic, all the outcomes are the same for both mechanisms. However, in a stochastic setting, the type of mechanisms plays a key role in the risk sharing between the supplier and the government induced by the demand uncertainty.

We next vary the different model parameters in order to compare the outcomes for both mechanisms computationally. We obtain the following results. As the variance of the noise increases, the expected government expenditures and the expected profit of the supplier under the subsidy mechanism are higher when comparing to the consumer subsidy mechanism. Consequently, although the optimal subsidy (per unit) is lower, the overall subsidy program is actually more costly to the government. Indeed, the government is subsidizing all the produced units instead of the sold ones and therefore is bearing some of the overstock risk from the demand uncertainty. Since, the supplier is sharing this overstock risk with the government, he can achieve higher profits on expectation despite the fact that the subsidy is smaller by charging a higher price to the consumers.

## 5.7 Computational Results

In this section, we present some numerical examples that provide further insights into the results derived in Section 5.4. The data used in these experiments is inspired by

the sales data of the first two years of General Motors' Chevy Volt (between December 2010 and June December 2012). The total aggregate sales was roughly equal to 35,000<sup>6</sup> electric vehicles, the listed price (MSRP) was \$40,280 and the government subsidy was set to \$7,500. In addition, we assume a 10% profit margin<sup>7</sup> so that the per-unit cost of building manufacturing capacity is \$36,000. Note that we tested the robustness of all our results and plots in this section with respect to  $c$  by considering varying the profit margin from 0.5% to 20% and obtained very similar insights. For simplicity, we present here the results using a linear demand with an additive Gaussian noise. We observe that our results along with the analysis are robust with respect to the distribution of demand uncertainty. In fact, we obtain in our computational experiments the same insights for several demand distributions (including non-symmetric ones). As discussed in Section 5.3, the government can either minimize expenditures or maximize the total welfare. In particular, the two objectives are equivalent and give rise to the same optimal subsidy policies for any target level  $\Gamma$  above a certain threshold. In this case, this threshold is equal to 860 and the condition is therefore easily satisfied.

Throughout these experiments, we compute the optimal decisions for both the deterministic and the stochastic demand models by using the optimal expressions derived in Section 5.4.1. We first consider a fixed relatively large standard deviation  $\sigma = 42,000$  (when demand is close to the sales, this is equivalent to a coefficient of variation of 1.2) and plot the optimal subsidy, production level, supplier's profit and government expenditures as a function of the target level  $\Gamma$  for both the deterministic and stochastic models. The plots are reported in Figure 5-3. We have derived in Section 5.4.1 a set of inequalities regarding the relations of the optimal variables for deterministic and stochastic demand models. The plots allow us to quantify the magnitude of these differences and study the impact of demand uncertainty on the optimal policies. One can see from Figure 5-3 that the optimal production levels are

---

<sup>6</sup><http://www.hybridcars.com/december-2012-dashboard>

<sup>7</sup><http://www.nada.org/NR/rdonlyres/C1C58F5A-BE0E-4E1A-9B56-1C3025B5B452/0/NADADATA2012Final.pdf>

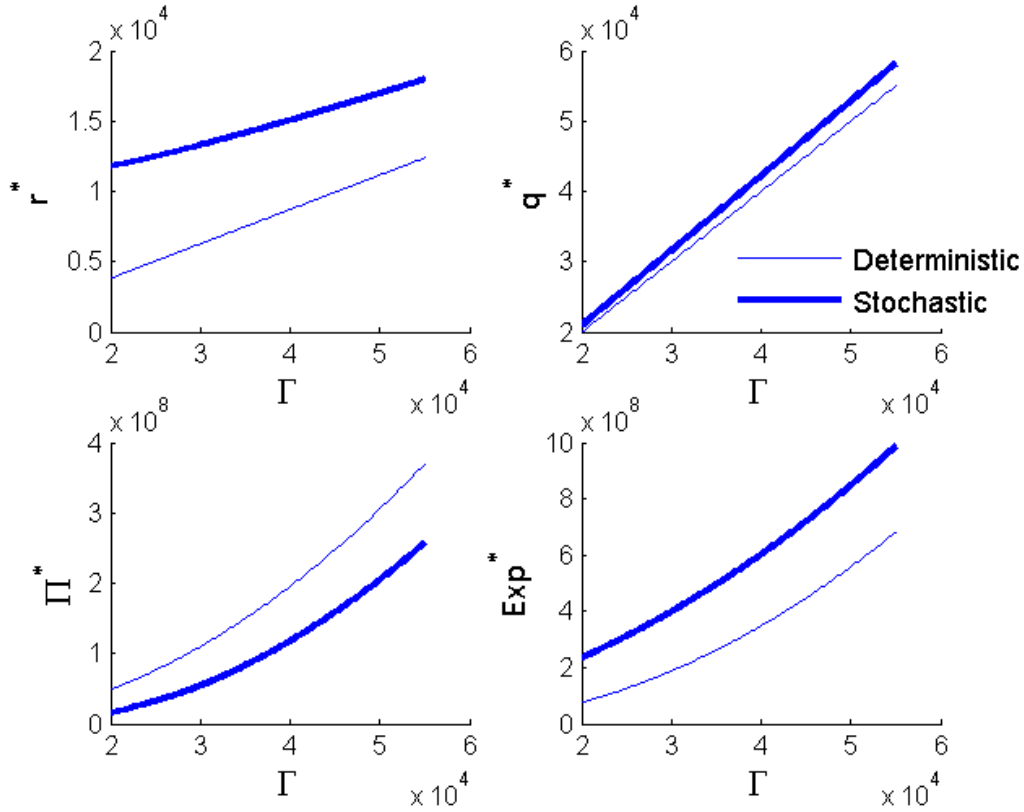


Figure 5-3: Optimal values as a function of the target level

not strongly affected by demand uncertainty (even for large values of  $\sigma$ ) when the target level  $\Gamma$  is set close to the expected sales value of 35,000. However, the optimal value of the subsidy is almost multiplied by a factor of 2 when demand uncertainty is taken into account. In other words, when the government and the supplier consider a richer environment that accounts for demand uncertainty, the optimal subsidy nearly doubles.

One can see that the optimal production quantities for deterministic and stochastic cases differ very little, while the subsidy and profit show significantly higher discrepancy. By looking at the closed form expressions for linear demand in Theorem 5.1, one can see that the difference in optimal quantity is equal to  $F_{\epsilon}^{-1}\left(\frac{p_{sto}-c}{p_{sto}}\right) - K(p_{sto})$ , whereas the difference in optimal subsidy is proportional to  $K(p_{sto})$ . Note that in our case, the profit margin is relatively small (order of 0.1) and therefore the quantile

value,  $F_{\epsilon}^{-1}\left(\frac{p_{sto}-c}{p_{sto}}\right)$ , is likely to be negative. In particular, in our example,  $c = 36,000$  and  $\Gamma$  ranges from 20,000 to 55,000. As a result, since we assume a symmetric noise distribution, the quantile is always negative. Consequently, the difference in quantities is clearly smaller than the difference in subsidy. One interesting interpretation relies on the fact that the cost of uncertainty in production quantity is shared between the supplier and the government. Indeed, the government wants to incentivize the supplier to increase production in order to reach the adoption target, and therefore is willing to share some of the uncertainty risk so that  $q_{sto}$  is not far from  $q_{det}$ . Finally, the profit discrepancy is larger than the quantity discrepancy as it is equal to the same difference scaled by the cost  $c$  (see Theorem 5.1). In our example,  $c = 36,000$  and therefore, we can see a more significant difference.

This raises the following interesting question. What happens if the government ignores demand uncertainty and decides to under-subsidize by using the optimal value from the deterministic model? It is clear that in this case, since the real demand is uncertain, the expected sales will not attain the desired expected adoption. We address this question in the remaining of this section. We first plot the subsidies and the supplier's profit as a function of the standard deviation of the noise that represents a measure of the demand uncertainty magnitude.

More precisely, we plot in Figure 5-4 the relative differences in subsidies (i.e.,  $\frac{r_{sto}-r_{det}}{r_{det}}$ ) as well as the supplier's profit as a function of the target level  $\Gamma$  (or equivalently, the expected sales) for different standard deviations of the additive noise varying from 35 to 12,500. For  $\Gamma = 35,000$ , this is equivalent to a coefficient of variation varying between 0.001 and 0.357. As expected, one can see from Figure 5-4 that as the standard deviation of demand increases, the optimal subsidy is larger whereas the supplier's profit is lower. As a result, demand uncertainty benefits consumers at the expense of hurting both the government and the supplier.

Finally, we analyze by how much the government will miss the actual target level (now  $\Gamma$  is fixed and equal to 35,000) by using the optimal policy assuming demand is deterministic,  $r_{det}$ , instead of using  $r_{sto}$ . Recall that  $r_{sto} \geq r_{det}$ . In other words, the

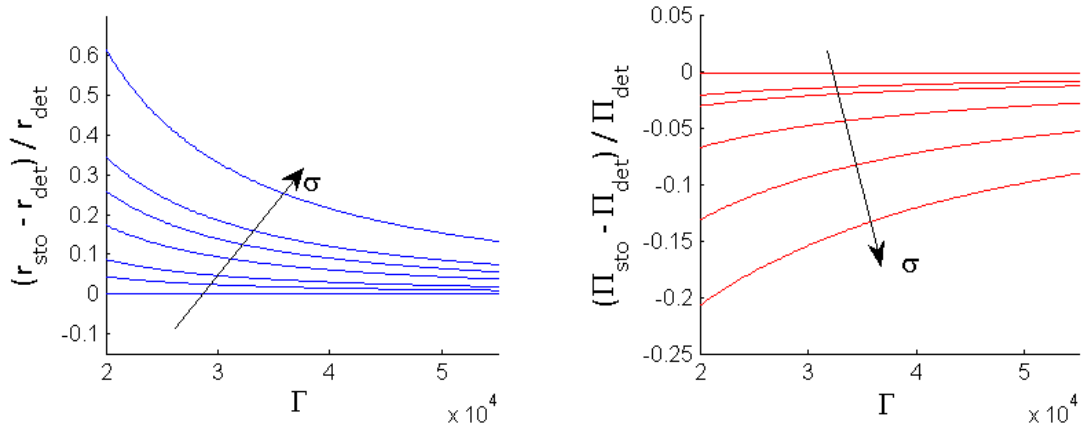


Figure 5-4: Relative normalized differences in subsidies (left) and supplier's profit (right)

government assumes a simple average deterministic demand model whereas in reality demand is uncertain. In particular, this allows us to quantify the value of using a more sophisticated model that takes into account demand uncertainty instead of simply ignoring it. Note that this analysis is different from the previous comparisons in this chapter, where we compared the optimal decisions as a function of demand uncertainty. Here, we assume that demand is uncertain with some given distribution but the government decides to ignore the uncertainty. To this extent, we consider two possible cases according to the modeling assumption of the supplier. First, we assume that the supplier is non-sophisticated, in the sense that he uses an average demand approximation model as well (i.e., no information on demand distribution is used). In this case, both the supplier and the government assume an average deterministic demand but in reality demand is random. Second, the supplier is more sophisticated. Namely, the supplier optimizes (over both  $p$  and  $q$ ) by using a stochastic demand model together with the distribution information. The results are presented in Figure 5-5, where we vary the value of the coefficient of variation of the noise from 0 to 0.7. When the government and the supplier are both non-sophisticated, the government can potentially save money (by under-subsidizing) and still gets close to the target in expectation when demand uncertainty is not very large. As expected, when the

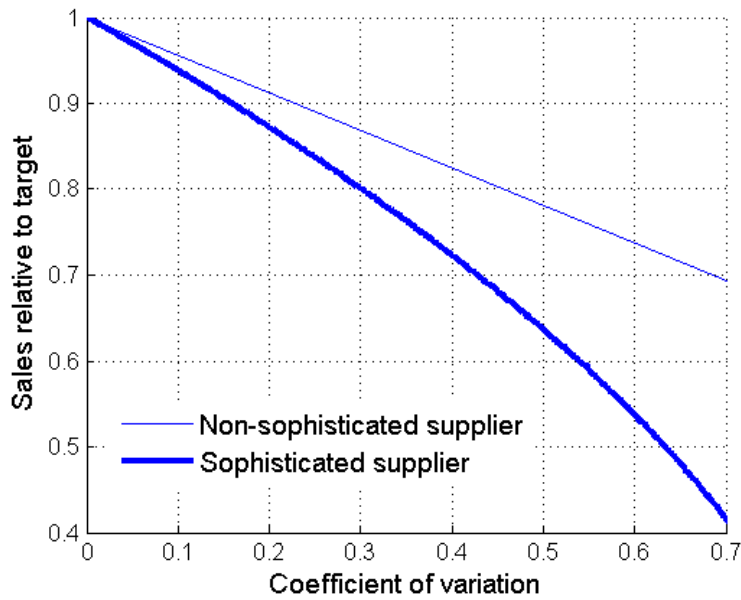


Figure 5-5: Expected sales relative to target

supplier has more information on demand distribution (as it is usually the case), the expected sales are farther from the target and the government could miss the target level significantly. If in addition, demand uncertainty is large (i.e., coefficient of variation larger than 1), the government misses the target in both cases.

One can formalize the previous comparison analytically by quantifying the gap by which the government misses the target by ignoring demand uncertainty. In particular, let us consider the additive demand model given in (5.4.1).

**Proposition 5.4.** *Consider that the government ignores demand uncertainty when designing consumer subsidies.*

1. *The government misses the adoption target (in expectation) regardless of whether the supplier is sophisticated or not.*
2. *The exact gaps are given by:*

- For non-sophisticated supplier:

$$\mathbb{E}[\min(q, D(z, \epsilon))] = \Gamma + \mathbb{E}[\min(0, \epsilon)] \leq \Gamma. \quad (5.7.1)$$

- For sophisticated supplier, assuming a linear demand model:

$$\mathbb{E}[\min(q, D(z, \epsilon))] = \Gamma + \frac{1}{2} \cdot \mathbb{E}\left[\min\left(F_\epsilon^{-1}\left(\frac{p^* - c}{p^*}\right), \epsilon\right)\right] \leq \Gamma. \quad (5.7.2)$$

In addition, the optimal price of the sophisticated supplier,  $p^*$ , is lower than the deterministic case, i.e.,  $p^* \leq p_{det}$ .

3. For low (high) profit margins, the gap is larger (smaller) when the supplier is sophisticated (non-sophisticated).

Note that all the results of Proposition 5.4 (with the exception of equation (5.7.2)) are valid for a general demand model. Nevertheless, when the supplier is sophisticated, one needs to assume a specific model (in our case, linear) in order to compute the expected sales. As expected, the previous analysis suggests that the target adoption will be missed by a higher margin as demand becomes more uncertain. When comparing the non-sophisticated and sophisticated cases, one can see that in the former the government misses the target by  $\mathbb{E}[\min(0, \epsilon)]$ , whereas in the latter by  $\frac{1}{2} \cdot \mathbb{E}\left[\min\left(F_\epsilon^{-1}\left(\frac{p^* - c}{p^*}\right), \epsilon\right)\right]$ . Consequently, this difference depends not only on the distribution of the noise but also on the profit margin. Since the EV industry has rather low profit margins, the gap may be much larger when the supplier is sophisticated. Indeed, the sophisticated supplier decreases the price (relative to  $p_{det}$ ). In addition, he reduces the production quantities as he is not willing to bear significant over-stock risk due to the low profit margin. As a result, the expected sales are lower and therefore the government misses the adoption target level. In conclusion, this analysis suggests that policy makers should take into account demand uncertainty when designing consumer subsidies. Indeed, by ignoring demand uncertainty, one can significantly miss the desired adoption target.

## 5.8 Conclusions

We propose a model to analyze the interaction between the government and the supplier when designing consumer subsidy policies. Subsidies are often introduced at the early adoption stages of green technologies to help them become economically viable faster. Given the high level of uncertainty in these early stages, we hope to have shed some light on how demand uncertainty affects consumer subsidy policies, as well as price and production quantity decisions from manufacturers and the end consumers.

In practice, policy makers often ignore demand uncertainty and consider only deterministic forecasts of adoption when designing subsidies. We demonstrate that uncertainty will significantly change how these programs should be designed. In particular, we show by how much the government misses the adoption target by ignoring demand volatility. Among some of our main insights, we show that the shape of the demand curve will determine who bears demand uncertainty risk. When demand is uncertain, quantities produced will be higher and the effective price for consumers will be lower. For convex demand functions prices will be lower, leading to lower industry profits.

Focusing on the linear demand model, we can derive further insights. For instance, to compensate for uncertain demand, quantities produced and subsidy levels are shifted by a function of the service level, i.e., the profitability of the product. For highly profitable products, the supplier will absorb most of the demand risk. When profit margins are smaller, the government will need to increase the subsidy amount and pay a larger share for the risk.

When evaluating the uncertainty impact on consumers, we must consider the trade-off between lower effective prices and the probability of a stock-out (unserved demand). We again show that the shape of the demand curve plays an important role. For linear demand, consumers will ultimately benefit from demand uncertainty. This is not the case, for instance, with an iso-elastic demand model, where the possibility

of not serving customers with high valuations will out-weight the benefits of decreased prices.

We also compare the optimal policies to the case where a central planner manages jointly the supplier and the government and tries to optimize the entire system simultaneously. We show that the optimal effective price and production level coincide in both the decentralized and centralized models. Consequently, the subsidy mechanism is sufficient to coordinate the government and the supplier and the collusion does not hurt consumers in terms of price and quantities.

There are interesting directions for future research. Our model focuses on a single period, which is more applicable for policies with short time horizons. In cases where the policy horizon is long, it would be interesting to understand the effect of time dynamics on the actions of both government and supplier. It is not obvious how the frequency of policy adjustment will affect the outcome of the subsidy program, when considering a strategic response of the supplier. Additionally, introducing competition between suppliers could also be an interesting direction to extend the current model.

## Bibliography

- [1] P. Aghion, A. Dechezleprêtre, D. Hemous, R. Martin, and J. Van Reenen. Carbon taxes, path dependency and directed technical change: evidence from the auto industry. Technical report, National Bureau of Economic Research, 2012.
- [2] S. Alizamir, F. de Véricourt, and P. Sun. Efficient feed-in-tariff policies for renewable energy technologies. 2013. Working paper.
- [3] K. Anand, R. Anupindi, and Y. Bassok. Strategic inventories in vertical contracts. *Management Science*, 54(10):1792–1804, 2008.
- [4] T. Andreyeva, M.W. Long, and K.D. Brownell. The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food. *American journal of public health*, 100(2):216, 2010.

- [5] J.I. Arar. New directions: the electric car and carbon emissions in the us. *Atmospheric environment*, 44(5):733–734, 2010.
- [6] K. Arifoglu, S. Deo, and S.M.R. Iravani. Consumption externality and yield uncertainty in the influenza vaccine supply chain: Interventions in demand and supply sides. *Management Science*, 58(6):1072–1091, 2012.
- [7] A. Atasu, L.N. Van Wassenhove, and M. Sarvary. Efficient take-back legislation. *Production and Operations Management*, 18(3):243–258, 2009.
- [8] B. Avci, Girotra K., and S. Netessine. Electric vehicles with a battery switching station: Adoption and environmental impact. 2013. INSEAD Working Paper No. 2013/46/TOM.
- [9] R. Beard. N-firm oligopoly with general iso-elastic demand. *Bulletin of Economic Research*, 2013.
- [10] A. Benthem, K. Gillingham, and J. Sweeney. Learning-by-doing and the optimal solar policy in california. *The Energy Journal*, 29(3):131151, 2008.
- [11] J. Bentzen. An empirical analysis of gasoline demand in denmark using cointegration techniques. *Energy Economics*, 16(2):139–143, 1994.
- [12] R.W. Boadway and D.E. Wildasin. Optimal tax-subsidy policies for industrial adjustment to uncertain shocks. *Oxford Economic Papers*, pages 105–134, 1990.
- [13] G. Jr Brown and M. B. Johnson. Public utility pricing and output under risk. *AER*, 59(1):119–128, 1969.
- [14] G.P. Cachon. Supply chain coordination with contracts, s. graves and f. de kok, eds. *Handbooks in Operations Research and Management Science: Supply chain Management*, 11:229 – 339, 2003.

- [15] G.P. Cachon and M.A. Lariviere. Capacity choice and allocation: strategic behavior and supply chain performance. *Management Science*, 45(8):1091–1108, 1999.
- [16] G.P. Cachon and M.A. Lariviere. Supply chain coordination with revenue-sharing contracts: strengths and limitations. *Management science*, 51(1):30–44, 2005.
- [17] O. Capps. Utilizing scanner data to estimate retail demand functions for meat products. *American Journal of Agricultural Economics*, 71(3):750–760, 1989.
- [18] F. Carlsson and O. Johansson-Stenman. Costs and benefits of electric vehicles. *Journal of Transport Economics and Policy*, 37(1):138, 2003.
- [19] D. Carlton. Peak load pricing with stochastic demand. *American Economic Review*, 76(4):637–658, 1986.
- [20] A. Chandra, S. Gulati, and M. Kandlikar. Green drivers or free riders? an analysis of tax rebates for hybrid vehicles. *Journal of Environmental Economics and Management*, 60(2):78–93, 2010.
- [21] S.E. Chick, H. Mamani, and D. Simchi-Levi. Supply chain coordination and influenza vaccination. *Operations Research*, 56(6):1493–1506, 2008.
- [22] M. C. Cohen, G. Perakis, and C. Thraves. Consumer surplus when demand is stochastic. 2013. Working paper.
- [23] M. C. Cohen, G. Perakis, and C. Thraves. The effects of competition in green technology adoption. 2013. Working paper.
- [24] M. A. Crew, C. S. Fernando, and P. R. Kleindorfer. The theory of peak-load pricing: A survey. *Journal of Regulatory Economics*, 8(3):215–248, 1995.
- [25] CSI. California solar incentive program. 2007.
- [26] A. Deaton and G. Laroque. On the behaviour of commodity prices. *The Review of Economic Studies*, 59(1):1–23, 1992.

- [27] D. Diamond. The impact of government incentives for hybrid-electric vehicles: Evidence from us states. *Energy Policy*, 37(3):972–983, 2009.
- [28] DOE. One million electric vehicles by 2015 - status report. 2011.
- [29] DOE. Sunshot vision study. 2012.
- [30] U. Eberle and R. Von Helmut. Sustainable transportation based on electric vehicle concepts: A brief overview. *Energy and Environmental Science*, 3:689–699, 2010.
- [31] EPA. Model year 2012 green vehicle guide. 2012.
- [32] F. Erhun, P. Keskinocak, and S. Tayur. Dynamic procurement, quantity discounts, and supply chain efficiency. In *Supply Chain Coordination under Uncertainty*, pages 219–233. Springer, 2011.
- [33] H. Fujimoto and E. Park. Optimal export subsidy when demand is uncertain. *Economics Letters*, 55(3):383–390, 1997.
- [34] K.S. Gallagher and E. Muehlegger. Giving green to get green? Incentives and consumer adoption of hybrid vehicle technology. *Journal of Environmental Economics and Management*, 61(1):1–15, 2011.
- [35] D. Granot and S. Yin. On the effectiveness of returns policies in the price-dependent newsvendor model. *Naval Research Logistics*, 52(8):765–779, 2005.
- [36] D. Granot and S. Yin. On sequential commitment in the price-dependent newsvendor model. *European Journal of Operational Research*, 177(2):939–968, 2007.
- [37] D. Granot and S. Yin. Price and order postponement in a decentralized newsvendor model with multiplicative and price-dependent demand. *Operations Research*, 56(1):121–139, 2008.

- [38] J. Huang, M. Leng, and M. Parlar. Demand functions in decision modeling: A comprehensive survey and research directions. *Decision Sciences*, 44(3):557–609, 2013.
- [39] X. Huang and G. Sošić. Analysis of industry equilibria in models with sustaining and disruptive technology. *European Journal of Operational Research*, 207(1):238–248, 2010.
- [40] A. Jenn, I.L. Azevedo, and P. Ferreira. The impact of federal incentives on the adoption of hybrid electric vehicles in the united states. *Energy Economics*, 40:936–942, 2013.
- [41] M. Kaya and Ö. Özer. *Pricing in Business to Business Contracts: Sharing Risk, Profit and Information*. The Oxford Handbook of Pricing Management, Oxford University Press, 2012.
- [42] A. Kocabiykoğlu and I. Popescu. An elasticity approach to the newsvendor with price-sensitive demand. *Operations Research*, 59(2):301–312, 2011.
- [43] A.H.L. Lau and H. Lau. Effects of a demand-curve’s shape on the optimal solutions of a multi-echelon inventory/pricing model. *European Journal of Operational Research*, 147(3):530–548, 2003.
- [44] H. Liu and Ö. Özer. Channel incentives in sharing new product demand information and robust contracts. *European Journal of Operational Research*, 207(3):1341–1349, 2010.
- [45] R. Lobel and G. Perakis. Consumer choice model for forecasting demand and designing incentives for solar technology. 2013. Working paper.
- [46] H. Lutze and Ö. Özer. Promised lead-time contracts under asymmetric information. *Operations Research*, 56(4):898–915, 2008.

- [47] H. Mamani, E. Adida, and D. Dey. Vaccine market coordination using subsidy. *IEE Transactions on healthcase Systems Engineering*, 2(1):78–96, 2011.
- [48] A. Ovchinnikov and G. Raz. A newsvendor model with pricing for public interest goods. 2013. Working paper.
- [49] N. C. Petruzzi and M. Dada. Pricing and the newsvendor problem: A review with extensions. *Operations Research*, 47(2):183–194, 1999.
- [50] R.S. Pindyck and D.L. Rubinfeld. *Microeconomics*, 8th. 2001.
- [51] E. L. Porteus. Stochastic inventory theory. *Handbooks in OR & MS*, 2:605–652, 1990.
- [52] T. Puu. Chaos in duopoly pricing. *Chaos, Solitons & Fractals*, 1(6):573–581, 1991.
- [53] J.M. Sallee. The surprising incidence of tax credits for the toyota prius. *American Economic Journal: Economic Policy*, pages 189–219, 2011.
- [54] W. Sierzchula, S. Bakker, K. Maat, and B. van Wee. The influence of financial incentives and other socio-economic factors on electric vehicle adoption. *Energy Policy*, 68:183–194, 2014.
- [55] C.P. Simon and L. Blume. *Mathematics for economists*, volume 7. Norton New York, 1994.
- [56] K.T. Talluri and G.J. Van Ryzin. *The theory and practice of revenue management*, volume 68. springer, 2006.
- [57] T.A. Taylor. Supply chain coordination under channel rebates with sales effort effects. *Management science*, 48(8):992–1007, 2002.
- [58] T.A. Taylor and W. Xiao. Subsidizing the distribution channel: Donor funding to improve the availability of malaria drugs. 2013. Working paper.

- [59] K. E. Train. *Optimal regulation*. 1991. MIT press, MA, USA.
- [60] X. Vives. *Oligopoly Pricing: Old Ideas and New Tools*, volume 1 of *MIT Press Books*. The MIT Press, 2001.
- [61] W. L. Winston. *Operations Research: Applications and Algorithms*. Duxbury, 3rd edition, 1994.
- [62] L. Yao, Y. Chen, and H. Yan. The newsvendor problem with pricing: Extensions. *IJMSEM*, 1:3–16, 2006.
- [63] P. H. Zipkin. *Foundations of inventory management*. McGraw-Hill, 2000.



# Chapter 6

## Conclusion

This thesis started with the motivation that finding the right pricing strategy has always been an interesting topic for retailers and researchers. In the last years, many tools and opportunities have emerged that can potentially help retailers to improve their pricing decisions. The ability to develop mathematical models that capture consumer behavior and the power of optimization are two examples of tools that one can use to achieve this goal. In addition, an unprecedented amount of data is becoming available to retailers. In particular, retailers can collect very valuable information on their customers such as transaction history, personal features and social interactions. It is no doubt that pricing decisions are present everywhere but interestingly, new applications are arising. This thesis investigates four different such applications: promotions for supermarket retailers, social networks, new products and subsidies for green technology. This clearly only scratches the surface, as one can think about numerous other applications that are equally important.

Some of the results in this thesis may be directly used for inferring better pricing decisions (e.g., the promotion optimization model and tool developed in Chapter 2), whereas others are based on stylized models that can allow to draw useful qualitative insights. One of the main rewarding feelings of this thesis was to realize that we could identify important real-world problems, develop analytical results based on operations research tools and ultimately reach conclusions on how to improve current practices.



# Appendix A

## Proofs for Chapter 2

### Proof of Lemma 2.1

1. Since the proof may not be easy to follow, we present it together with a concrete example to illustrate the different steps. Let  $T = 6$ ,  $q^0 = 7$ ,  $A = \{(1, 1), (3, 3)\}$ ,  $B = \{(3, 3)\}$  and  $(t', k') = (5, 5)$ . We denote by  $POP_t(\mathbf{p}_A)$  the profits at time  $t$  for the price vector  $\mathbf{p}_A$ . In addition, we further assume that:  $\delta_5 = g_4(1) = 0.8$ ,  $\delta_6 = g_5(1) = 0.9$ . We next define the following quantities:

$$\begin{aligned} a_t &= POP_t(\mathbf{p}_A) &&= POP_t(1, 7, 3, 7, 7, 7) \\ a'_t &= POP_t(\mathbf{p}_{A \cup (t', k')}) &&= POP_t(1, 7, 3, 7, 5, 7) \\ b_t &= POP_t(\mathbf{p}_B) &&= POP_t(7, 7, 3, 7, 7, 7) \\ b'_t &= POP_t(\mathbf{p}_{B \cup (t', k')}) &&= POP_t(7, 7, 3, 7, 5, 7). \end{aligned}$$

For each time  $t$ , we define the following coefficient:

$$\delta_t = \frac{g_1((\mathbf{p}_A)_{t-1}) \cdot g_2((\mathbf{p}_A)_{t-2}) \cdots g_{t-1}((\mathbf{p}_A)_1)}{g_1((\mathbf{p}_B)_{t-1}) \cdot g_2((\mathbf{p}_B)_{t-2}) \cdots g_{t-1}((\mathbf{p}_B)_1)}.$$

$\delta_t$  represents the multiplicative reduction in demand at time  $t$  from the promotions present in the set  $A$  but not in  $B$ . Observe that from Assumption 1, we

have  $o \leq \delta_{t'} \leq \delta_{t'+1} \leq \dots \leq \delta_T \leq 1$ . In addition, we have:  $a_t = \delta_t b_t$ ,  $a'_t = \delta_t b'_t$ . Observe also that condition (2.6.5) is equivalent to:

$$\sum_{t=1}^T a'_t - \sum_{t=1}^T a_t \geq 0. \quad (\text{A.0.1})$$

Note that  $a_t = a'_t$  for all  $t < t'$ . In the example, we have  $a_1 = a'_1, \dots, a_4 = a'_4$  as the prices in periods 1-4 are the same. Therefore, (A.0.1) becomes:  $\sum_{t=t'}^T a'_t \geq \sum_{t=t'}^T a_t$ . In the example, we obtain:  $a'_5 + a'_6 \geq a_5 + a_6$ . Note that  $a'_t \leq a_t$  for any  $t > t'$ . In the example,  $a'_t$  has a promotion at  $t = 5$ . However, there is no promotion in  $a_t$  at  $t = 5$  and therefore, the objective at  $t = 6$  for  $a'_t$  is lower than the one in  $a_t$ , i.e.,  $a'_6 \leq a_6$ . This implies that:

$$a'_{t'} - a_{t'} \geq \sum_{t=t'+1}^T (a_t - a'_t) \geq 0.$$

In the example, this translates to  $a'_5 - a_5 \geq a_6 - a'_6 \geq 0$ . We next multiply the left hand side by  $1/\delta_{t'}$  and the terms in the right hand side by  $1/\delta_t$  (recall that  $1/\delta_{t'} \geq 1/\delta_t$  for  $t > t'$ ). Therefore, we obtain:

$$b'_{t'} - b_{t'} = \frac{a'_{t'} - a_{t'}}{\delta_{t'}} \geq \sum_{t=t'+1}^T \left( \frac{a_t - a'_t}{\delta_t} \right) = \sum_{t=t'+1}^T (b_t - b'_t) \geq 0.$$

In the example, this translates to:  $b'_5 - b_5 = \frac{a'_5 - a_5}{0.8} \geq \frac{a_6 - a'_6}{0.9} = b_6 - b'_6 \geq 0$ . Recall that our goal is to show equation (2.6.6), or alternatively:  $\sum_{t=1}^T a'_t - \sum_{t=1}^T a_t \leq \sum_{t=1}^T b'_t - \sum_{t=1}^T b_t$ . Note that this is equivalent to:  $\sum_{t=t'}^T (a'_t - a_t) = \sum_{t=t'}^T \delta_t (b'_t - b_t) \leq \sum_{t=t'}^T (b'_t - b_t)$ . By rearranging the terms, we obtain:

$$\sum_{t=t'+1}^T (1 - \delta_t)(b_t - b'_t) \leq (1 - \delta_{t'})(b'_{t'} - b_{t'}).$$

In the example, this would be:  $0.1(b_6 - b'_6) \leq 0.2(b'_5 - b_5)$ . Finally, note that the above inequality is true because of the following:

$$\sum_{t=t'+1}^T (1 - \delta_t)(b_t - b'_t) \leq \sum_{t=t'+1}^T (1 - \delta_{t'}) (b_t - b'_t) \leq (1 - \delta_{t'})(b'_{t'} - b_{t'}).$$

In the example, this is clear because:  $0.1(b_6 - b'_6) \leq 0.2(b_6 - b'_6) \leq 0.2(b'_5 - b_5)$ .

2. We first introduce the following notation. Let  $\gamma^{POP}$  be an optimal solution to the POP and  $\{(t_1, k_1), \dots, (t_n, k_n)\}$  the set of promotions in  $\gamma^{POP}$ . For any subset  $B \subset \{1, 2, \dots, n\}$ , we define:  $\gamma(B) = \gamma(\{(t_i, k_i) : i \in B\})$ . For example, let the price ladder be  $\{q^0 = 5, q^1 = 4\}$  and  $\gamma^{POP} = \gamma(\{(1, 1), (3, 1), (5, 1)\})$ . Then,  $\gamma(\{1, 3\}) = \gamma(\{(1, 1), (5, 1)\})$ .

Note that one can write the following telescoping sum:

$$POP(\gamma^{POP}) = POP(\gamma\{1\}) + \sum_{m=1}^{n-1} \left[ POP(\gamma\{1, \dots, m+1\}) - POP(\gamma\{1, \dots, m\}) \right].$$

Based on Proposition A.1 for each  $m = 1, 2, \dots, n-1$ :  $POP(\gamma\{1, \dots, m+1\}) - POP(\gamma\{1, \dots, m\}) \geq 0$ . By applying the submodularity property from Lemma 2.1 part 1, we obtain:  $0 \leq POP(\gamma\{1, \dots, m+1\}) - POP(\gamma\{1, \dots, m\}) \leq POP(\gamma\{m+1\}) - POP(\gamma^0)$ . Therefore, we have:

$$\begin{aligned} POP(\gamma^{POP}) &= POP(\gamma\{1\}) + \sum_{m=1}^{n-1} \left[ POP(\gamma\{1, \dots, m+1\}) - POP(\gamma\{1, \dots, m\}) \right] \\ &\leq POP(\gamma^0) + \sum_{m=1}^n \left[ POP(\gamma\{m\}) - POP(\gamma^0) \right] = LP(\gamma^{POP}). \end{aligned}$$

**Proposition A.1.** *Let  $n \geq 2$  be an integer and  $\gamma^{POP}$  an optimal solution to the POP with  $n$  promotions. Then,  $POP(\gamma\{1, \dots, m+1\}) - POP(\gamma\{1, \dots, m\}) \geq 0$  for  $m = 1, 2, \dots, n-1$ .*

The proof proceeds by induction on the number of promotions. We first show that the claim is true for the base case i.e.,  $n = 2$ . By the optimality of  $\gamma^{POP} = \gamma\{1, 2\}$ ,

we have:

$$0 \leq POP(\gamma\{1, 2\}) - POP(\gamma\{1, 2\}).$$

Next, we assume that the claim is true for  $n$  and show its correctness for  $n + 1$ . Let POP' denote the POP problem with the additional constraint that promotion  $(t_1, k_1)$  is used, i.e.,  $p_{t_1} = q^{k_1}$ . One can see that the set of promotions  $\{(t_2, k_2), \dots, (t_{n+1}, k_{n+1})\}$  is an optimal solution to POP' with  $n$  promotions. Therefore, by using the induction hypothesis, we have:

$$\begin{array}{rcl} POP'(\gamma\{2, \dots, n, n+1\}) & - & POP'(\gamma\{2, \dots, n\}) \geq 0 \\ \vdots & & \vdots \\ POP'(\gamma\{2, 3\}) & - & POP'(\gamma\{2\}) \geq 0 \end{array}$$

Equivalently, in terms of the POP:

$$\begin{array}{rcl} POP(\gamma\{1, \dots, n, n+1\}) & - & POP(\gamma\{1, \dots, n\}) \geq 0 \\ \vdots & & \vdots \\ POP(\gamma\{1, 2, 3\}) & - & POP(\gamma\{1, 2\}) \geq 0 \end{array}$$

Therefore, it remains to show that:  $POP(\gamma\{1, 2\}) - POP(\gamma\{1\}) \geq 0$ . We next prove the following chain of inequalities:

$$\begin{aligned} POP(\gamma\{1, 2\}) - POP(\gamma\{1\}) &\geq POP(\gamma\{1, 2, 3\}) - POP(\gamma\{1, 3\}) \\ &\geq POP(\gamma\{1, 2, 3, 4\}) - POP(\gamma\{1, 3, 4\}) \\ &\vdots \\ &\geq POP(\gamma\{1, \dots, n, n+1\}) - \\ &POP(\gamma\{1, 3, 4, \dots, n, n+1\}). \end{aligned} \tag{A.0.2}$$

By using the induction hypothesis together with the submodularity property from Lemma 2.1 part 1, we obtain for each  $m = 2, 3, \dots, n - 1$ :

$$\begin{aligned} POP(\gamma\{1, \dots, m, m + 1\}) - POP(\gamma\{1, \dots, m\}) &\leq \\ &POP(\gamma\{1, 3, 4, \dots, m, m + 1\}) - POP(\gamma\{1, 3, 4, \dots, m\}). \end{aligned}$$

Finally, from the optimality of  $\gamma\{1, \dots, n, n + 1\}$ , we have:  $POP(\gamma\{1, \dots, n, n + 1\}) - POP(\gamma\{1, \dots, n\}) \geq 0$ . By rearranging the terms in the above equations, one can derive the chain of inequalities in (A.0.2) and this concludes the proof.

## Proof of Proposition 2.1

(a) When  $L = 1$ , only a single promotion is allowed and therefore the IP approximation is equivalent to the POP. Indeed, the IP approximation evaluates the POP objective through the sum of unilateral price changes.

(b) In the second case, demand at time  $t$  is assumed to depend only on the current price  $p_t$  and not on past prices. Consequently, the objective function is separable in terms of time (note that the periods are still tied together through some of the constraints). In this case too, the IP approximation is exact since each price change affects only the profit at the time it was made.

(c) We next show that the IP approximation is exact for the case where  $S \geq 1$  and the demand at time  $t$  depends on the current and last period prices only.

Note that in this case, promotions affect only current and next period demands, but not demand in periods  $t + 2, t + 3, \dots$ . We consider a price vector with two promotions at times  $t$  and  $u$  (i.e.,  $p_t = q^i$  and  $p_u = q^j$ ) and no promotion at all the remaining times, denoted by  $\mathbf{p}\{p_t = q^i, p_u = q^j\}$ . From the feasibility with respect to the separating constraints, we know that  $t$  and  $u$  are separated by at least one time period. We need to show that the profits from doing both promotions is equal to the

sum of the incremental profits from doing each promotion separately, that is:

$$\begin{aligned} POP(\mathbf{p}\{p_t = q^i, p_u = q^j\}) - POP(\mathbf{p}^0) = \\ POP(\mathbf{p}\{p_t = q^i\}) - POP(\mathbf{p}^0) + POP(\mathbf{p}\{p_u = q^j\}) - POP(\mathbf{p}^0). \end{aligned} \quad (\text{A.0.3})$$

(d) One can extend the previous argument to generalize the proof for the case where the number of separating periods is larger or equal than the memory. Indeed, if  $S \geq M$ , the IP approximation is not neglecting correlations between different promotions and hence optimal.

## Proof of Proposition 2.2

We prove the result by expressing the LP relaxation of the IP in Linear Programming standard form, and then showing that the constraint matrix is totally unimodular.

We collect the decision variables  $\gamma_t^k$ , into a vector of size  $(K + 1)T$  as follows:

$$\boldsymbol{\gamma} = [\gamma_1^0, \dots, \gamma_1^K, \gamma_2^0, \dots, \gamma_2^K, \dots, \gamma_T^0, \dots, \gamma_T^K]^T.$$

Similarly, we denote by  $\mathbf{b}$  the vectorization of the objective coefficients  $b_t^k$  defined in (2.5.1). By relaxing the integrality constraints, the IP problem can be written in the following standard LP form:

$$\begin{aligned} \max_{\boldsymbol{\gamma}} \quad & \mathbf{b}^T \boldsymbol{\gamma} \\ \text{s.t.} \quad & A\boldsymbol{\gamma} \leq \mathbf{u} \end{aligned} \quad (\text{A.0.4})$$

$$0 \leq \boldsymbol{\gamma} \leq 1$$

where  $\mathbf{1}_K^T$  is a vector of ones with length  $K$ , and the matrix  $A$  and the vector  $\mathbf{u}$  are





totally unimodular and consequently every basic feasible solution of (2.5.2) is integral.

## Proof of Theorem 2.1

Note that the lower bound follows directly from the feasibility of  $\gamma^{LP}$  for the POP.

We next prove the upper bound by showing the following chain of inequalities:

$$\underline{R} \cdot LP(\gamma^{LP}) \stackrel{(i)}{\leq} POP(\gamma^{LP}) \stackrel{(ii)}{\leq} POP(\gamma^{POP}) \stackrel{(iii)}{\leq} LP(\gamma^{POP}) \stackrel{(iv)}{\leq} LP(\gamma^{LP}).$$

Inequality (i) follows from Proposition 2.3 below. Inequality (ii) follows from the optimality of  $\gamma^{POP}$  and inequality (iii) follows from part 2 of Lemma 2.1 below.

Finally, inequality (iv) follows from the optimality of  $\gamma^{LP}$ . Therefore, we obtain:

$$\underline{R} = \underline{R} \cdot \frac{POP(\gamma^{POP})}{POP(\gamma^{POP})} \leq \underline{R} \cdot \frac{LP(\gamma^{LP})}{POP(\gamma^{POP})} \leq \frac{POP(\gamma^{LP})}{POP(\gamma^{POP})} \leq \frac{POP(\gamma^{POP})}{POP(\gamma^{POP})} = 1.$$

## Proof of Proposition 2.3

We denote the set of promotions in the price vector  $\mathbf{p}$  by:  $\mathbf{p} = \mathbf{p}\{(t_1, q^{t_1}), \dots, (t_N, q^{t_N})\}$ , where  $N$  is the number of promotions. The price vector  $\mathbf{p}^n = \mathbf{p}\{(t_n, q^{t_n})\}$  for each  $n = 1, \dots, N$  denotes the single promotion price at time  $t_n$  (no promotion at the remaining periods). By convention, let us denote  $n = 0$  to be the regular price only vector  $\mathbf{p}^0 = (q^0, \dots, q^0)$ . We denote the cumulative POP objective in periods  $[u, v)$  when using  $\mathbf{p}^n$  by:  $x_{[u,v]}^n = POP(\mathbf{p}\{(t_n, q^{t_n})\})_{[u,v]} = \sum_{t=u}^{v-1} \mathbf{p}_t\{(t_n, q^{t_n})\} d_t(\mathbf{p}_t\{(t_n, q^{t_n})\})$ . Note that the LP objective can be written as:  $LP(\mathbf{p}) = x_{[1,T]}^0 + \sum_{n=1}^N (x_{[1,T]}^n - x_{[1,T]}^0)$ . Since  $\mathbf{p}^n$  and  $\mathbf{p}^0$  do not promote before  $t_n$ , we have  $x_{[1,t_n]}^n = x_{[1,t_n]}^0$ . In addition, since  $\mathbf{p}^n$  promotes at  $t = t_n$  and  $\mathbf{p}^0$  does not, the vector  $\mathbf{p}^n$  yields a lower objective for the periods after  $t_n$ , i.e.,  $x_{[t_{n+1},T]}^n \leq x_{[t_{n+1},T]}^0$ . Therefore, we obtain for each  $n = 1, \dots, N$ :  $x_{[1,T]}^n - x_{[1,T]}^0 = x_{[1,t_n]}^n + x_{[t_n,t_{n+1}]}^n + x_{[t_{n+1},T]}^n - x_{[1,t_n]}^0 - x_{[t_n,t_{n+1}]}^0 - x_{[t_{n+1},T]}^0 \leq x_{[t_n,t_{n+1}]}^n - x_{[t_n,t_{n+1}]}^0$ . Therefore:  $LP(\mathbf{p}) \leq UB = x_{[1,T]}^0 + \sum_{n=1}^N (x_{[t_n,t_{n+1}]}^n - x_{[t_n,t_{n+1}]}^0) = x_{[1,t_1]}^0 + \sum_{n=1}^N x_{[t_n,t_{n+1}]}^n$ . Let  $UB_t$  denote the value of  $UB$  at time  $t$ . Specifically, if  $t \in [t_n, t_{n+1})$ , then  $UB_t = x_t^n$ . We can write for any feasible price vector  $\mathbf{p}$ :  $POP(\mathbf{p}) = \sum_{t=1}^T a_t UB_t$ , where  $a_t$  is

the decrease in demand at time  $t$  due to the past promotions in  $\mathbf{p}$ . In particular, if  $t_n < t \leq t_{n+1}$ , then:  $a_t = g_{t-t_1}(q^{t_1})g_{t-t_2}(q^{t_2}) \cdots g_{t-t_n}(q^{t_n})$ . Since  $0 \leq \underline{R} \leq a_t \leq 1$ , we obtain:  $\underline{R} \cdot LP(\mathbf{p}) \leq \underline{R} \cdot UB \leq POP(\mathbf{p})$ .

## Proof of Proposition 2.4

### 1. Lower bound

In the case when  $S \geq M$ , we know from Proposition 2.1 that the LP approximation is exact. Therefore, the result holds in this case.

We next consider that  $S < M$  and construct an instance of the POP as well as a price vector  $\mathbf{p}^*$ . We then show that this price vector  $\mathbf{p}^*$  is optimal for both the POP and the LP approximation.

Let  $T = L(M + 1)$  and let us define the following price vector:

$$\mathbf{p}^* = \left( q^K, \underbrace{q^0, \dots, q^0}_{M \text{ times}}, q^K, \underbrace{q^0, \dots, q^0}_{M \text{ times}}, \dots, q^K, \underbrace{q^0, \dots, q^0}_{M \text{ times}} \right).$$

Let  $\mathbb{U} = \{1, (M + 1) + 1, 2(M + 1) + 1, \dots, (L - 1)(M + 1) + 1\}$  denote the set of promotion periods in  $\mathbf{p}^*$ . We choose the demand functions  $f_t$  to be:

$$f_t(p_t) = \begin{cases} Z/q^K & \text{if } t \in \mathbb{U} \text{ and } p_t = q^K, \\ 1/q^0 & \text{otherwise,} \end{cases}$$

where:

$$Y = 1 + \sum_{m=1}^M (1 - g_m(q^K)),$$

$$Z = (M + 2)Y.$$

We define all the costs to be zero, i.e.,  $c_t = 0, \forall t = 1, \dots, T$ . We prove the proposition by the following steps:

**Step 1:** We show that  $\mathbf{p}^*$  is an optimal LP solution.

**Step 2:** We show that there exists an optimal POP solution with promotions only during periods  $t \in \mathbb{U}$ .

**Step 3:** We show that if  $\mathbf{p}$  promotes only during periods  $t \in \mathbb{U}$ , then  $POP(\mathbf{p}) \leq POP(\mathbf{p}^*)$ .

By combining steps 2 and 3, we conclude that  $\mathbf{p}^*$  is an optimal POP solution. Consequently,  $POP(\mathbf{p}^{POP}) = POP(\mathbf{p}^{LP})$ , implying that the lower bound is tight.

**Proof of Step 1.** By definition, we have:  $POP(\mathbf{p}\{(t, K)\}) = POP(\mathbf{p}^0) + Z - Y$  for  $t \in \mathbb{U}$ . Therefore the LP coefficients as defined in (2.5.1) are given by:

$$b_t^k = \begin{cases} Z - Y & \text{if } t \in \mathbb{U}, k = K, \\ \leq 0 & \text{otherwise.} \end{cases}$$

Any LP optimal solution selects at most  $L$  of  $\gamma_t^k$ , for  $k = 1, \dots, K$  to be 1. Consequently, the optimal LP objective is bounded above by  $T + L(Z - Y)$ . In fact, the following  $\gamma^{LP}$  achieves this bound and is therefore optimal:

$$(\gamma^{LP})_t^k = \begin{cases} 1 & \text{if } t \in \mathbb{U}, k = K \\ 1 & \text{if } t \notin \mathbb{U}, k = 0 \\ 0 & \text{otherwise} \end{cases}$$

We then conclude that  $\mathbf{p}^{LP} = \mathbf{p}^*$  is an optimal LP solution.

**Proof of Step 2.** Consider any feasible price vector  $\mathbf{p}$  and let  $\mathbb{A}$  be the set of promotions in  $\mathbf{p}$ . We next show that  $POP(\mathbf{p}) \leq POP(\mathbf{p}^*)$  so that  $\mathbf{p}^*$  is an optimal POP solution. If  $\mathbf{p}$  uses the promotion  $p_t = q^k$  during a period  $t \notin \mathbb{U}$ , then we can consider the reduced set of promotions  $\mathbb{B} = \mathbb{A} \setminus \{(t, k)\}$ . Note that

the promotion  $(t, k)$  does not increase the profit at time  $t$ . Indeed, decreasing the price  $p_t$  will not increase the profit at time  $t$  since  $f_t(p_t) = 1/q^0$  for all  $p_t$ , and potentially will reduce the profit in future periods  $t + 1, \dots, t + M$ . Thus, removing the promotion  $(t, k)$  increases the total profit, that is  $POP(\gamma(\mathbb{A})) \leq POP(\gamma(\mathbb{B}))$ . By applying this procedure repeatedly, one can reach a price vector with only promotions in periods  $t \in \mathbb{U}$  that achieves a profit at least equal to  $POP(\mathbf{p})$ . In other words, there exists an optimal POP solution with promotions only during periods  $t \in \mathbb{U}$ .

**Proof of Step 3.** Let  $\mathbf{p}$  be a price vector that only contains promotions during periods  $t \in \mathbb{U}$ . Let  $n$  be the number of periods  $t$  in  $\mathbf{p}$  such that  $p_t = q^K$  ( $n \leq L$  because  $\mathbb{U}$  is composed of  $L$  periods). Note that all the successive promotions in  $\mathbb{U}$  are separated by at least  $M$  periods so that each pair of promotions of  $\mathbf{p}$  does not interact. Therefore, the profit of  $\mathbf{p}$  is given by:

$$POP(\mathbf{p}) = POP(\mathbf{p}^0) + n(Z - Y) \leq POP(\mathbf{p}^0) + L(Z - Y).$$

From the definition of  $\mathbf{p}^*$ , we have that  $POP(\mathbf{p}^*) = POP(\mathbf{p}^0) + L(Z - Y)$ . Indeed, each promotion  $(t, K)$  of  $\mathbf{p}^*$  results in an increase in profit of  $Z - Y$ , and each pair of promotions of  $\mathbf{p}^*$  is separated by at least  $M$  periods so that there is no interaction between promotions. Consequently,  $\mathbf{p}^*$  is an optimal POP solution and the lower bound is tight.

## 2. Upper bound

Let us denote the bound with  $n$  promotions by:

$$\underline{R}_n = \prod_{i=1}^{n-1} g_{i(S+1)}(q^K), \tag{A.0.5}$$

when  $\underline{R}_0 = 1$  by convention. We can also define the following limit:

$$\underline{R}_\infty = \lim_{n \rightarrow \infty} \underline{R}_n.$$

Note that  $g_m(q^K) \leq 1$  so that  $R_n$  is non-increasing with respect to  $n$ . Note also that  $g_m(q^K) = 1$  for  $m > M$  so that  $\underline{R}_{M+1} = \underline{R}_{M+2} = \dots = \underline{R}_\infty$ , i.e., the sequence  $\underline{R}_n$  converges.

In the case when  $S \geq M$ , we know from Proposition 2.1 that the LP approximation is exact. We also know from (A.0.5) that  $\underline{R}_n = 1$  for all  $n$ . Therefore, the result holds in this case.

We next consider that  $S < M$  and define the following sequence of problems:

$$POP^n = POP(\langle q^k \rangle_{k=0}^K, \langle f_t^n \rangle_{t=1}^{T_n}, \langle c_t \rangle_{t=1}^{T_n}, \langle g_m \rangle_{m=1}^M, L_n, S),$$

where  $\langle q^k \rangle_{k=0}^K, \langle g_m \rangle_{m=1}^M, S$  are given parameters and the costs  $c_t = 0$ . In addition,  $L_n = n$ , and  $T_n = n(M + 1)$ . We choose the functions  $f_t^n$  to be equal:

$$f_t^n(p_t) = \begin{cases} Z/q^K & \text{if } 1 \leq t \leq LM + 1 \text{ and } p_t = q^K, \\ 1/q^0 & \text{otherwise.} \end{cases}$$

where,

$$Y = 1 + \sum_{m=1}^M (1 - g_m(q^K)),$$

$$Z = 100Yn.$$

We prove the proposition by the following steps:

**Step 1:** We show that the following price vector is an optimal LP solution:

$$\mathbf{p}^{LP} = \left( q^K, \underbrace{q^0, \dots, q^0}_{S \text{ times}}, q^K, \underbrace{q^0, \dots, q^0}_{S \text{ times}}, \dots, q^K, \underbrace{q^0, \dots, q^0}_{T-(L-1)(S+1)-1 \text{ times}} \right).$$

**Step 2:** We show that:

$$POP^n(\mathbf{p}_n^{LP}) \leq T - L + Z(\underline{R}_1 + \dots + \underline{R}_n).$$

**Step 3:** We show the following lower bound for the optimal profit:  $POP^n(\mathbf{p}_n^{POP}) \geq nZ$ .

**Step 4:** We finally prove the convergence of the following limit, implying the desired result:

$$\lim_{n \rightarrow \infty} \frac{POP_n(\mathbf{p}_n^{POP})}{POP_n(\mathbf{p}_n^{LP})} = \frac{1}{\underline{R}_\infty}.$$

**Proof of Step 1.** Based on the above definitions, we have:  $POP(\mathbf{p}\{(t, K)\}) = POP(\mathbf{p}^0) + Z - Y$  for  $1 \leq t \leq LM + 1$ . Therefore, the LP coefficients are given by:

$$b_t^k = \begin{cases} Z - Y & \text{if } 1 \leq t \leq LM + 1, k = K, \\ \leq 0 & \text{otherwise.} \end{cases}$$

Let  $\mathbb{U} = \{1, S + 1, 2S + 1, \dots, LS + 1\}$  denote the set of promotion periods in  $\mathbf{p}^{LP}$ .

Any LP optimal solution selects at most  $L$  of  $\gamma_t^k$ , for  $k = 1, \dots, K$  to be 1. Consequently, the optimal LP objective is bounded above by  $T + L(Z - Y)$ . In fact, the following  $\gamma^{LP}$  achieves this bound and is therefore optimal:

$$(\gamma^{LP})_t^k = \begin{cases} 1 & \text{if } t \in \mathbb{U}, k = K \\ 1 & \text{if } t \notin \mathbb{U}, k = 0 \\ 0 & \text{otherwise} \end{cases}$$

Therefore, we conclude that the price vector  $\mathbf{p}^{LP}$  is an optimal LP solution.

**Proof of Step 2.** One can see that the profit induced by the  $i$ -th promotion of  $\mathbf{p}_L^{LP}$  (at time  $t = (i - 1)S + 1$ ) is  $\underline{R}_i Z$  due to the effect of the promotions  $1, 2, \dots, (i - 1)$ . In addition, the profit from each non-promotion period is bounded above by 1. We obtain:

$$POP_n(\gamma_n^{LP}) \leq T - L + Z(\underline{R}_1 + \underline{R}_2 + \dots + \underline{R}_n).$$

**Proof of Step 3.** Consider the following price vector:

$$\mathbf{p} = \left( q^K, \underbrace{q^0, \dots, q^0}_{M \text{ times}}, q^K, \underbrace{q^0, \dots, q^0}_{M \text{ times}}, \dots, q^K, \underbrace{q^0, \dots, q^0}_{M \text{ times}} \right).$$

Note that  $\mathbf{p}$  is feasible for  $POP_n$ . Note that all the successive promotions are separated by at least  $M$  periods so that each pair of promotions of  $\mathbf{p}$  does not interact. Therefore, the profit induced by the  $i$ -th promotion in  $\mathbf{p}$  (at time  $t = (i - 1)M + 1$ ) is  $Z$ . As a result, we obtain the following lower bound for the POP profit of  $\mathbf{p}$ :

$$POP_n(\mathbf{p}) \geq nZ.$$

This also provides us a lower bound for the optimal POP profit:

$$POP_n(\mathbf{p}_n^{POP}) \geq POP_n(\mathbf{p}) \geq nZ.$$

**Proof of Step 4.** We show that  $\frac{1}{\underline{R}_\infty}$  is both a lower and upper bound of the limit. First, using Theorem 2.1 for  $POP^n$ , we have:

$$\frac{POP^n(\mathbf{p}_n^{POP})}{POP^n(\mathbf{p}_n^{LP})} \leq \frac{1}{\underline{R}_n}.$$

By taking the limit when  $n \rightarrow \infty$  on both sides:

$$\lim_{n \rightarrow \infty} \frac{POP^n(\mathbf{p}_n^{POP})}{POP^n(\mathbf{p}_n^{LP})} \leq \lim_{n \rightarrow \infty} \frac{1}{\underline{R}_n} = \frac{1}{\underline{R}_\infty}.$$

By using Steps 2 and 3, we obtain:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{POP^n(\mathbf{p}_n^{POP})}{POP^n(\mathbf{p}_n^{LP})} &\geq \lim_{n \rightarrow \infty} \frac{n \cdot 100nY}{nM + 100nY(\underline{R}_1 + \underline{R}_2 + \cdots + \underline{R}_n)} \\ &= \lim_{n \rightarrow \infty} \frac{1}{\frac{M}{100nY} + \frac{\underline{R}_1 + \underline{R}_2 + \cdots + \underline{R}_n}{n}} = \frac{1}{\underline{R}_\infty}. \end{aligned}$$

In the last equality, we have used the fact that if  $\langle a_n \rangle_{n=1}^\infty$  converges to a finite limit  $a$ , then  $\langle \sum_{i=1}^n a_i/n \rangle_{n=1}^\infty$  also converges to  $a$ .

## Proof of Proposition 2.5

- We next show that all the coefficients of orders 3 and higher in the  $App(N)$  formulation are equal to zero. It then allows us to conclude that  $App(2) = App(N)$ . First, let us show that all the coefficients of order 3 (i.e., the sets with 3 items on promotion) are zero. We then proceed by induction.

Recall that the coefficient for items  $i$ ,  $j$  and  $k$  at time  $t$  is given by (we drop the time index for clarity):

$$\begin{aligned} b_{ijk} &= POP(i, j, k) - POP(i, j) - POP(i, k) - POP(j, k) \\ &\quad + POP(i) + POP(j) + POP(k) - POP(0). \end{aligned} \tag{A.0.6}$$

Here,  $POP(i, j, k)$  denotes the total profits of the  $N$  items throughout the planning horizon of  $T$  periods, when only items  $i$ ,  $j$  and  $k$  are on promotion at time  $t$  (similarly,  $POP(j)$  denotes the total profits when only item  $j$  is on promotion at time  $t$ ). First, observe that all the terms in equation (A.0.6) at times different than  $t$  are zero. Observe also that the coefficient in equation (A.0.6) affects only the items  $i$ ,  $j$  and  $k$  and is zero otherwise. As a result, we

remain with only three contributions (for items  $i$ ,  $j$  and  $k$  at time  $t$ ). From symmetry in the indices, it is sufficient to consider item  $i$  and show that its contribution is zero. For item  $i$  at time  $t$ , equation (A.0.6) yields:

$$\begin{aligned}
[b_{ijk}]_{i,t} &= (p_i^t - c_i^t) \left[ h_t^i(p_i^t, p^0, \dots, p^0) + H_t^{ji}(p_t^j) + H_t^{ki}(p_t^k) + \sum_{\ell \neq i, \ell \neq j, \ell \neq k} H_t^{\ell i}(p^0) \right] \\
&\quad - (p_i^t - c_i^t) \left[ h_t^i(p_i^t, p^0, \dots, p^0) + H_t^{ji}(p_t^j) + \sum_{\ell \neq i, \ell \neq j} H_t^{\ell i}(p^0) \right] \\
&\quad - (p_i^t - c_i^t) \left[ h_t^i(p_i^t, p^0, \dots, p^0) + H_t^{ki}(p_t^k) + \sum_{\ell \neq i, \ell \neq k} H_t^{\ell i}(p^0) \right] \\
&\quad - (p^0 - c_i^t) \left[ h_t^i(p^0, \dots, p^0) + H_t^{ji}(p_t^j) + H_t^{ki}(p_t^k) + \sum_{\ell \neq i, \ell \neq j, \ell \neq k} H_t^{\ell i}(p^0) \right] \\
&\quad + (p_i^t - c_i^t) \left[ h_t^i(p_i^t, p^0, \dots, p^0) + \sum_{\ell \neq i} H_t^{\ell i}(p^0) \right] \\
&\quad + (p^0 - c_i^t) \left[ h_t^i(p^0, \dots, p^0) + H_t^{ji}(p_t^j) + \sum_{\ell \neq i, \ell \neq j} H_t^{\ell i}(p^0) \right] \\
&\quad + (p^0 - c_i^t) \left[ h_t^i(p^0, \dots, p^0) + H_t^{ki}(p_t^k) + \sum_{\ell \neq i, \ell \neq k} H_t^{\ell i}(p^0) \right] \\
&\quad - (p^0 - c_i^t) \left[ h_t^i(p^0, \dots, p^0) + \sum_{\ell \neq i} H_t^{\ell i}(p^0) \right] \\
&= (p^0 - c_i^t)[0] + (p_i^t - c_i^t)[0] = 0.
\end{aligned}$$

The last step follows from canceling terms and is not reported for conciseness. Therefore, all the coefficients with three items simultaneously on promotion are equal to zero. We assume by induction that all the coefficients with  $S = 4, 5, \dots, K - 1$  items are zero for some given  $K - 1 < N$ . We next show that the claim is true for  $K$ .

Note that we have several coefficients that include  $K$  items (one such coefficient for any subset of the  $N$  items with size  $K$ ). For example, the coefficient for items  $1, 2, \dots, K$  at time  $t$  is given by (we drop the time index for clarity):

$$b_{12\dots K} = POP(1, 2, \dots, K) - [\text{All the } K - 1] + [\text{All the } K - 2] - \dots POP(0),$$

where the term [All the  $K - 1$ ] refers to all the contributions of having a total of  $K - 1$  out of the  $K$  items on promotion. Alternatively, one can express the coefficient  $b_{12\dots K}$  as a function of the smaller order coefficients as follows:

$$b_{12\dots K} = POP(1, 2, \dots, K) - POP(0) - \sum[\text{All } bs \text{ with } K - 1 \text{ items}] \\ - \sum[\text{All the } bs \text{ with } K - 2 \text{ items}] - \dots - \sum_{\text{pairs}} b_{ij} - \sum_{j=1}^K b_j.$$

By using the induction hypothesis, we obtain:

$$b_{12\dots K} = POP(1, 2, \dots, K) - POP(0) - \sum_{\text{pairs}} b_{ij} - \sum_{j=1}^K b_j. \quad (\text{A.0.7})$$

Here,  $POP(1, 2, \dots, K)$  denotes the total profits of the  $N$  items throughout the planning horizon of  $T$  periods, when only items  $1, 2, \dots, K$  are on promotion at time  $t$ . We next look at the different types of contributions of the coefficient in equation (A.0.7). First, observe that all the terms at times different than  $t$  are zero. Observe also that the coefficient in equation (A.0.7) affects only the items  $1, 2, \dots, K$  and is zero otherwise. From symmetry in the indices, it is sufficient to consider item  $i$  and show that its contribution is zero. For item  $i$  at time  $t$ , we have:

$$\begin{aligned}
& [b_{12\dots K}]_{i,t} \\
&= (p_i^t - c_i^t) \left[ h_t^i(p_i^t, p^0, \dots, p^0) + \sum_{\ell=1,2,\dots,K,\ell \neq i} H_t^{\ell i}(p_i^\ell) + \sum_{\ell \neq 1,2,\dots,K,\ell \neq i} H_t^{\ell i}(p^0) \right] \\
&- (p^0 - c_i^t) \left[ h_t^i(p^0, \dots, p^0) + \sum_{\ell \neq i} H_t^{\ell i}(p^0) \right] \\
&- \sum_{\text{pairs } (i,n); i=1,2,\dots,K,n \neq i} \left\{ (p_i^t - c_i^t) \left[ h_t^i(p_i^t, p^0, \dots, p^0) + H_t^{ni}(p_t^n) \right] \right. \\
&+ \sum_{j \neq i, j \neq n} H_t^{ji}(p^0) - h_t^i(p_i^t, p^0, \dots, p^0) - \sum_{j \neq i} H_t^{ji}(p^0) \left. \right] \\
&- (p^0 - c_i^t) \left[ h_t^i(p^0, \dots, p^0) + H_t^{ni}(p_t^n) \right. \\
&+ \sum_{j \neq i, j \neq n} H_t^{ji}(p^0) - h_t^i(p^0, \dots, p^0) - \sum_{j \neq i} H_t^{ji}(p^0) \left. \right] \left. \right\} \\
&- (p_i^t - c_i^t) \left[ h_t^i(p_i^t, p^0, \dots, p^0) + \sum_{\ell \neq i} H_t^{\ell i}(p^0) \right] \\
&+ (p^0 - c_i^t) \left[ h_t^i(p^0, \dots, p^0) + \sum_{\ell \neq i} H_t^{\ell i}(p^0) \right] \\
&- \sum_{j=1,\dots,K,j \neq i} (p^0 - c_i^t) \left[ H_t^{ji}(p_j^t) - H_t^{ji}(p^0) \right] \\
&= (p^0 - c_i^t)[0] + (p_i^t - c_i^t)[0] = 0.
\end{aligned}$$

The last step follows from canceling terms and is not reported for conciseness. Note that all the pairs of items where item  $i$  is not included have a contribution of zero.

- The second part can be proved in the exact same way as in Proposition 2.1 and is not repeated.

## Proof of Proposition 2.6

We separate the proof into two parts.

**Lemma A.1.** *Consider a demand function with additively separable cross-effects,*

i.e.,

$$d_t^i(p_t^i, p_{t-1}^i, \dots, p_{t-M_i}^i, \mathbf{p}_t^{-i}) = h_t^i(p_t^i, p_{t-1}^i, \dots, p_{t-M_i}^i) + \sum_{j \neq i} H_t^{ji}(p_t^j). \quad (\text{A.0.8})$$

Then, the cross coefficients for  $\text{App}(2)$ ,  $b_{ij}^t$ ,  $\forall i, j > i$  and  $\forall t$  are non-negative.

*Proof.* Recall that the pairwise coefficient for items  $i$  and  $j$  at time  $t$  is given by (we drop the time index for clarity):

$$b_{ij} = \text{POP}(i, j) - \text{POP}(i) - \text{POP}(j) + \text{POP}(0). \quad (\text{A.0.9})$$

As before,  $\text{POP}(i, j)$  denotes the total profits of the  $N$  items throughout the planning horizon of  $T$  periods, when only items  $i$  and  $j$  are on promotion at time  $t$ . First, observe that all the terms at times different than  $t$  are zero. Observe also that the coefficient in equation (A.0.9) has three different types of contribution (for items  $i$ ,  $j$  and  $k \neq i, j$  at time  $t$ ). We next show that each one of the three contributions is non-negative.

1. For item  $i$  at time  $t$ , we have:

$$\begin{aligned} [b_{ij}]_{i,t} &= (p_i^t - c_i^t) \left[ h_t^i(p_i^t, p^0, \dots, p^0) + H_t^{ji}(p_t^j) + \sum_{\ell \neq i, \ell \neq j} H_t^{\ell i}(p^0) \right] \\ &\quad - (p_i^t - c_i^t) \left[ h_t^i(p_i^t, p^0, \dots, p^0) + \sum_{\ell \neq i} H_t^{\ell i}(p^0) \right] \\ &\quad - (p^0 - c_i^t) \left[ h_t^i(p^0, \dots, p^0) + H_t^{ji}(p_t^j) + \sum_{\ell \neq i, \ell \neq j} H_t^{\ell i}(p^0) \right] \\ &\quad + (p^0 - c_i^t) \left[ h_t^i(p^0, \dots, p^0) + \sum_{\ell \neq i} H_t^{\ell i}(p^0) \right] \\ &= (p^0 - c_i^t) \left[ H_t^{ji}(p^0) - H_t^{ji}(p_t^j) \right] + (p_i^t - c_i^t) \left[ H_t^{ji}(p_t^j) - H_t^{ji}(p^0) \right] \\ &= (p^0 - p_i^t) \left[ H_t^{ji}(p^0) - H_t^{ji}(p_t^j) \right] \end{aligned}$$

We know that  $p^0 \geq p_i^t$  (recall that  $p^0$  is the regular price and  $p_i^t$  is a promotion

price). Using the fact that products  $i$  and  $j$  are substitutable, the function  $H_t^{ji}(\cdot)$  is non-decreasing and therefore:  $H_t^{ji}(p^0) - H_t^{ji}(p_t^j) \geq 0$ . Therefore, the contribution above is non-negative.

2. For item  $j$  at time  $t$ : The exact same argument follows by symmetry.
3. For any item  $k \neq i$  and  $k \neq j$ , at time  $t$  we have:

$$\begin{aligned}
[b_{ij}]_{k,t} &= (p^0 - c_k^t) \left[ h_t^k(p^0) + H_t^{ik}(p_t^i) + H_t^{jk}(p_t^j) + \sum_{\ell \neq i, \ell \neq j, \ell \neq k} H_t^{lk}(p^0) \right] \\
&\quad - (p^0 - c_k^t) \left[ h_t^k(p^0) + H_t^{ik}(p_t^i) + \sum_{\ell \neq i, \ell \neq k} H_t^{lk}(p^0) \right] \\
&\quad - (p^0 - c_k^t) \left[ h_t^k(p^0) + H_t^{jk}(p_t^j) + \sum_{\ell \neq j, \ell \neq k} H_t^{lk}(p^0) \right] \\
&\quad + (p^0 - c_k^t) \left[ h_t^k(p^0) + \sum_{\ell \neq k} H_t^{lk}(p^0) \right] \\
&= -(p^0 - c_k^t) \left[ H_t^{jk}(p^0) - H_t^{jk}(p_t^j) \right] = 0
\end{aligned}$$

□

Next, we consider the following optimization problem, called the Unconstrained Binary Quadratic Program (UBQP).

$$\begin{aligned}
(UBQP) \max_x \quad & \sum_{i=1}^N b_i x_i + \sum_i \sum_{j>i} b_{ij} x_i x_j & (A.0.10) \\
\text{s.t.} \quad & x_i \in \{0, 1\}; \quad \forall i
\end{aligned}$$

**Lemma A.2.** *If all the cross-coefficients  $b_{ij}$  are non-negative, one can solve (UBQP) by an LP.*

*Proof.* First, observe that one can reformulate the problem (UBQP) as an IP by defining a new variable for each pair of items  $i, j > i$ , denoted by  $x_{ij}$  and adding the

following consistency constraints:

$$x_{ij} \leq x_i; \quad x_{ij} \leq x_j; \quad x_{ij} \geq 0; \quad x_{ij} \geq x_i + x_j - 1. \quad (\text{A.0.11})$$

In this case, the new variable  $x_{ij}$  is equal to the product  $x_i x_j$ . Note that the two problems are equivalent. Note also that the variables  $x_{ij}$  are continuous but the variables  $x_i$  are binary. In addition, the LP relaxation is not tight in general, as the constraint matrix (composed of the consistency constraints) is not totally unimodular.

We next show that when all the cross-coefficients  $b_{ij}$  are non-negative, the formulation is integral. Assume by contradiction that we are given an optimal solution  $x^*$  with at least one fractional component. First, since  $b_{ij} \geq 0$  and using the fact that the objective is linear, we always have:  $x_{ij}^* = \min(x_i^*, x_j^*)$ . Second, remove all the components  $x_i^*$  that are equal to 0, as  $x_{ij}^* = 0; \forall j$ . Now, consider the elements such that  $x_k = \min x_i^*$  over all the positive values  $x_i^* > 0$ . If  $x_k^* = 1$ , we are done. Therefore, we consider that  $0 < x_k^* < 1$ . Note that we can have several elements that attain this minimal positive value (ties).

Assume there exist  $V$  such elements that attain the minimal (positive) value,  $x_{k_1}^*, x_{k_2}^*, \dots, x_{k_V}^*$ . We next construct an integer feasible solution with at least the same objective as  $x^*$ . We do so by looking at the minimal positive elements of  $x^*$  and show that we can set them to zero. We then repeat the procedure with the second minimal positive elements and so on, until we reach a solution with only integer values. At each step, we show that the objective function cannot decrease.

Consider for each  $i = 1, \dots, V$ :  $\tilde{x}_{k_i} = x_{k_i}^* - \epsilon$  and  $\tilde{x}_j = x_j^*, \forall j \neq k_i$ . In other words, the elements that correspond to the minimal indices  $k_1, \dots, k_V$  are each decreased by  $\epsilon$  and the remaining elements remain the same. The variation in the objective function from  $\tilde{x}$  relative to  $x^*$  is given by:  $-\sum_{i=1}^V b_{k_i} \epsilon - \sum_{i=1}^V \sum_j b_{k_{ij}} \epsilon = -\epsilon \sum_{i=1}^V (b_{k_i} + \sum_j b_{k_{ij}})$ . If  $\sum_{i=1}^V (b_{k_i} + \sum_j b_{k_{ij}}) > 0$ , by taking a small negative value for  $\epsilon$ , one can obtain a strictly better solution and this is a contradiction to the fact that  $x^*$  is optimal. Therefore, we have:  $\sum_{i=1}^V (b_{k_i} + \sum_j b_{k_{ij}}) \leq 0$ , so that the solution  $\tilde{x}$  yields at least the

same objective as  $x^*$  for a small  $\epsilon > 0$ . In addition, one can take all the  $\tilde{x}_{k_i}$  to be zero (i.e.,  $\epsilon = x_{k_1}^*$ ) without decreasing the objective. At this stage, we set the minimal positive elements to zero without decreasing the value of the objective. We then repeat the same procedure for the second minimal positive elements and so on, until we have constructed a solution with integer values, with at least the same objective as the initial optimal solution we started with.  $\square$

By combining the results from both lemmas, we can conclude the proof of the Proposition.

## **Proof of Theorem 2.2**

The proof is in the same spirit as in Theorem 2.1 by taking the minimum across all the items  $j = 1, 2, \dots, N$ .



# Appendix B

## Proofs for Chapter 3

### Proof of Theorem 3.1

1. We first prove the existence of a PNE for any given vector of prices. We know that a best response dynamics only terminates at a PNE, if it terminates (it can cycle whether or not a PNE exists). We use a constructive argument to show that the best response dynamics terminates and hence it must be a PNE. We begin at a state in which no agent buys. At each step, we pick an agent who performs a best response action and change his action. In particular, in our setting an agent decides whether to buy or not given other agents' current actions. We assume that when the agent switches from one action to another, he gains positive utility. We note that any agent who buys has no incentive to switch back to the no-buying action. This follows from the generalized non-negativity assumption in (3.2.2), where the influence of any agent (in the future steps) is non-negative. Therefore, in each step one agent has to switch from no-buy to buy until no one switches and the procedure terminates. Hence, a PNE exists and this concludes the proof.

Note that instead of (3.2.2), the weaker assumption (3.2.4) where the total value of any agent is non-negative suffices. This is because if another agent (in a future step) buys the item, this agent who is currently buying can lose some

value under sub-modular valuations (for example) but it never goes below zero because of (3.2.4). This means that this agent does not gain strictly by not buying (as the no-buying utility is 0) and hence never switches back.

2. If there are no ties in any of the PNEs, one can take  $\epsilon = 0$ . Consider the case where there are ties for some of the agents in one or more PNEs. In this case, one can choose  $\epsilon > 0$  to be very small such that (i) the agents that were buying in any PNE, are still buying (in particular, their utility strictly increase and they become better off); (ii) the agents that were not buying (i.e., have negative utility) continue not to buy and (iii) the agents who were indifferent (i.e., exactly at zero utility), become strictly better off as they perceive a positive utility when the price is reduced by  $\epsilon$ . Consequently, all the ties are eliminated for all the PNEs.
  
3. In the case of a unique PNE, this is by definition a Pareto optimal PNE for the agents. We next consider a setting where there are multiple equilibria. Consider any agent whose actions differ in the different PNEs. As there are two actions, one of the actions is to buy in one of the PNEs. Note that when the agent buys, he derives a (strictly) positive utility with the perturbed prices, and zero utility otherwise. As a result, this agent prefers to buy in the Pareto optimal solution. By using the generalized non-negativity assumption, this agent can only positively (non-negative to be precise) impact the valuation of other agents and increase their utility. This implies that all the other agents also prefer the buying decision of the primary agent. Similarly, one can argue that all agents who buy in one of the equilibria, will buy (and prefer) the Pareto solution. The only agents who remain are the ones that do not buy in any equilibria. Note that they will not buy in the Pareto optimal solution either. This Pareto optimal solution is also a PNE because the buyers have no incentive to deviate to not buying as they perceive a strictly positive utility and the non-buyers should not deviate as well because of their negative utility from buying.

### Proof of Proposition 3.1

Consider the continuous relaxation of problem Z that replaces the binary constraint  $\alpha_i \in \{0, 1\}$  by  $0 \leq \alpha_i \leq 1 \forall i \in \mathcal{I}$ . Let  $V^* = (p_i^*, y_i^*, \alpha_i^*) \forall i \in \mathcal{I}$  be an optimal solution for the relaxed problem with the corresponding objective  $\Pi^*$ . Assume that the latter optimal solution has at least one fractional component i.e.,  $\exists j \in \mathcal{I}$  s.t.  $0 < \alpha_j^* < 1$ . Now, consider two alternative feasible solutions, denoted by  $\bar{V}$  and  $\underline{V}$  to the relaxed problem Z given by:

$$\bar{V} = (\bar{p}_i, \bar{y}_i, \bar{\alpha}_i) = \begin{cases} (p_i^*, y_i^*, 1) & \text{if } i = j \\ (p_i^* + (1 - \alpha_j^*)W_{ji}, y_i^*, \alpha_i^*) & \forall i \in N_j \\ (p_i^*, y_i^*, \alpha_i^*) & \text{otherwise} \end{cases}$$

$$\underline{V} = (\underline{p}_i, \underline{y}_i, \underline{\alpha}_i) = \begin{cases} (p_i^*, y_i^*, 0) & \text{if } i = j \\ (p_i^* - \alpha_j^*W_{ji}, y_i^*, \alpha_i^*) & \forall i \in N_j \\ (p_i^*, y_i^*, \alpha_i^*) & \text{otherwise} \end{cases}$$

where  $N_j$  denotes the set of neighbors of agent  $j$  (excluding  $j$ ) and  $W_{ji}$  is given by:

$$W_{ji} = \sum_{\substack{S \ni j, \\ S \subset \mathcal{I} \setminus \{i\}}} g_{S,i} \prod_{k \in S \setminus \{j\}} \alpha_k^*$$

which is the total influence of agent  $j$  on  $i$ . We observe that both solutions are feasible to the problem for the three following reasons. First, since  $0 < \alpha_j^* < 1$  it implies that  $(\sum_{S \subset \mathcal{I} \setminus \{j\}} g_{S,j} \prod_{k \in S} \alpha_k^* - p_j^*) = 0$ . Otherwise, it cannot be a best response for agent

$j$  and cannot satisfy the equilibrium constraints. Therefore, changing  $\alpha_j^*$  to 1 or 0 does not affect the best response of agent  $j$ . Second, we have modified the prices of the neighbors of agent  $j$  exactly by the change in the level of influence from agent  $j$  on them. As a result, it yields the same profit for agent  $i \in N_j$  and their purchasing decisions remain the same. Third, since the agents in  $\mathcal{I} \setminus (\{j\} \cup N_j)$  are unaffected by the change in  $\alpha_j^*$  or  $p_i^* \forall i \in N_j$ , the solution remains feasible for them as well.

Let us denote the objective corresponding to these new solutions by  $\bar{\Pi}$  and  $\underline{\Pi}$  respectively. We observe that  $\Pi^* - \bar{\Pi} = -(1 - \alpha_j^*) \left[ (p_j^* - c) + \sum_{i \in N_j} \alpha_i^* W_{ji} \right]$  and  $\Pi^* - \underline{\Pi} = \alpha_j^* \left[ (p_j^* - c) + \sum_{i \in N_j} \alpha_i^* W_{ji} \right]$ . Since  $\Pi^*$  is the optimal value of the objective and  $0 < \alpha_j^* < 1$ , it should be the case that  $\left[ (p_j^* - c) + \sum_{i \in N_j} \alpha_i^* W_{ji} \right] = 0$ . Because if not, one of the solutions we constructed is strictly better than the optimal solution and this is a contradiction. Consequently, one can see that both  $\bar{V}$  and  $\underline{V}$  are optimal solutions as well since they are feasible and yield the same objective as  $V^*$ . In the process, we have reduced the number of fractional components by one. One can now repeat the same procedure for each fractional value  $\alpha_j^*$  to derive a constructive way of identifying a feasible integral solution to the original problem with an objective function that is as good as the initial fractional solution. Note that since the number of agents is finite, this step is repeated at most  $N$  times. One can conclude that the continuous relaxation of problem Z is tight, meaning that for any feasible fractional solution, one can find an integral solution with at least the same objective, if not better.

## Proof of Theorem 3.2

Before proving the main Theorem, we state and prove Lemma B.1 that identifies the optimal values of  $z_i$  and  $p_i$  given (discrete or fractional)  $\alpha_S \forall S \subset \mathcal{I}$ .

**Lemma B.1.** *For a given (discrete or fractional)  $\alpha_S \forall S \subset \mathcal{I}$ , the revenue maximizing*

solution (and hence profit maximizing as well because the  $\alpha_i$ s are fixed) is given by:

$$z_i = \sum_{S \subset I \setminus \{i\}} g_{S,i} \alpha_{S \cup i}, \quad p_i = z_i + (1 - \alpha_i) p^{\max} \quad \forall i \in \mathcal{I}.$$

To show this, consider the feasibility constraints for each agent, when  $\alpha_i$ s are fixed.

Eliminating  $y_i$  reduces them to:

$$\begin{aligned} \max \{0, p_i - (1 - \alpha_i) p^{\max}\} &\leq z_i \\ &\leq \min \left\{ p_i, \alpha_i p^{\max}, \sum_{S \subset I \setminus \{i\}} g_{S,i} \alpha_{S \cup i}, \sum_{S \subset I \setminus \{i\}} g_{S,i} [\alpha_{S \cup i} - \alpha_S] + p_i \right\}. \end{aligned} \quad (\text{B.0.1})$$

We know that  $\sum_{S \subset I \setminus \{i\}} g_{S,i}^+ \leq p^{\max}$  and  $\sum_{S \subset I \setminus \{i\}} g_{S,i} \alpha_{S \cup i} \leq \alpha_i \sum_{S \subset I \setminus \{i\}} g_{S,i}^+$  and therefore,  $\sum_{S \subset I \setminus \{i\}} g_{S,i} \alpha_{S \cup i} \leq \alpha_i p^{\max}$ . Note also that the objective aims to maximize  $z_i$ . Because  $p_i$  is also decision variable and increasing  $p_i$  also increases the feasible region for  $z_i$  (see (B.0.1)), it is easy to see that  $p_i = z_i + (1 - \alpha_i) p^{\max}$ . Putting all together we obtain: ( $0 \leq z_i$ )

$$z_i \leq \min \left\{ z_i + (1 - \alpha_i) p^{\max}, \sum_{S \subset I \setminus \{i\}} g_{S,i} \alpha_{S \cup i}, \sum_{S \subset I \setminus \{i\}} g_{S,i} [\alpha_{S \cup i} - \alpha_S] + z_i + (1 - \alpha_i) p^{\max} \right\}. \quad (\text{B.0.2})$$

Observe that  $\sum_{S \subset I \setminus \{i\}} g_{S,i} [\alpha_S - \alpha_{S \cup i}] \leq (1 - \alpha_i) p^{\max}$ . Therefore, maximizing for  $z_i$  results in  $z_i = \sum_{S \subset I \setminus \{i\}} g_{S,i} \alpha_{S \cup i}$  thus proving our claim.

Consider solving the relaxed version Z-MIP where the binary constraints for each  $\alpha_i \forall i \in \mathcal{I}$  are replaced by the constraint:  $0 \leq \alpha_i \leq 1$ . Let  $V^* = (\boldsymbol{\alpha}^*, \mathbf{p}^*, \mathbf{y}^*, \mathbf{z}^*)$  be a fractional optimal solution to the relaxed problem with a corresponding objective  $\Pi^*$ . We construct a new solution with all the  $\alpha_i$ s being integer, show its feasibility to the problem Z-MIP (hence relaxed Z-MIP as well) with an objective that is at least as good as  $V^*$ . We denote the solution we construct by  $\tilde{V} = (\tilde{\boldsymbol{\alpha}}, \tilde{\mathbf{p}}, \tilde{\mathbf{y}}, \tilde{\mathbf{z}})$  and its corresponding objective by  $\tilde{\Pi}$ . We next construct this new solution.

For any agent  $i$ ,  $\tilde{\alpha}_i = \lceil \alpha_i^* \rceil$  where  $\lceil \cdot \rceil$  refers to the ceiling function that maps a real number to the smallest following integer. If  $\tilde{\alpha}_i = 1$ ,  $\tilde{z}_i = \tilde{p}_i = \sum_{S \subset T \setminus \{i\}} g_{S,i}$  and  $\tilde{y}_i = 0$ . Otherwise,  $\tilde{z}_i = \tilde{y}_i = 0$  and  $\tilde{p}_i = p^{\max}$ . Let  $T$  be the subset of agents who buy i.e.,  $\tilde{\alpha}_i = 1$  which also refers to those with  $\alpha_i^* > 0$ . Then,  $\tilde{\alpha}_S = 1$  for any  $S \subset T$  and 0 otherwise.

From Lemma B.1, we know that

$$\Pi^* = \sum_{i \in T} \sum_{S \subset T \setminus \{i\}} g_{S,i} \alpha_{S \cup i}^* - c \sum_{i \in T} \alpha_i^* \quad (\text{B.0.3})$$

$$\tilde{\Pi} = \sum_{i \in T} \sum_{S \subset T \setminus \{i\}} g_{S,i} - c|T| \quad (\text{B.0.4})$$

If  $\sum_{S \subset T} g_{S,i} \geq c \forall i \in T$ ; since  $\alpha_{S \cup \{i\}} \leq \alpha_i$  and  $\sum_{S \subset T \setminus \{i\}} g_{S,i} \geq 0$ , we obtain  $\tilde{\Pi} - \Pi^* \geq 0$  and it completes the proof. However, this condition may be hard to check as it depends on the optimized set  $T$ .

In this case, we introduce the following sequence of iterative steps to show that  $\tilde{\Pi} \geq \Pi^*$ . We next reorder the agents, rewrite the objective function and argue that a certain property holds.

- Order the agents in the set  $T = \{k_1, k_2, \dots, k_{|T|}\}$  such that  $\alpha_{k_1}^* \geq \alpha_{k_2}^* \geq \dots \geq \alpha_{k_{|T|}}^*$ .
- Create the nested sets of agents:  $T_1 \subset T_2 \subset \dots \subset T_{|T|} = T$ , where  $k_1 = T_1, T_1 \cup k_2 = T_2, \dots, T_m \cup k_{m+1} = T_{m+1}, \dots, T_{|T|} = T$ .
- Rewrite  $\Pi^*$  as follows:

$$\Pi^* = \sum_{m=1}^{|T|} \left[ \sum_{\substack{S \ni k_m \\ S \subset T_m}} \sum_{i \in S} g_{S \setminus \{i\}, i} \alpha_S^* - c \alpha_{k_m}^* \right]. \quad (\text{B.0.5})$$

We build the above objective by considering the marginal terms that one obtains adding agent by agent starting from  $k_1$  all the way to  $k_{|T|}$ . For every agent that we add, say  $k_m$ , related value terms and cost terms are included. The value term corresponds to all influence terms related to all sets  $S$  that consists of  $k_m$ . In particular, they consist of terms where  $k_m$  is the influencer (i.e.,  $i \neq k_m$ ) and terms where  $k_m$  is influenced (i.e.,  $i = k_m$ ). We write this in the form the influence of  $S \setminus \{i\}$  on  $i$ .

We know  $\Pi^* \geq 0$  because a no-buy for everyone results in 0 profit and is a feasible solution. Now we argue that any cumulative sum starting from the last term in (B.0.6) has to be non-negative. That is,

$$\Pi_{[l,|T|]}^* = \sum_{m=l}^{|T|} \left[ \sum_{\substack{S \ni k_m \\ S \subset T_m}} \sum_{i \in S} g_{S \setminus i, i} \alpha_S^* - c \alpha_{k_m}^* \right] \geq 0 \quad \forall l = 1, \dots, |T|. \quad (\text{B.0.6})$$

Otherwise, it would have been beneficial to set the  $\alpha_{k_m}$ s corresponding to these agents ( $\alpha_{k_l}, \dots, \alpha_{k_{|T|}}$ ) to 0 and restrict  $T$  to just  $T_{l-1}$  (where  $T_0 = \emptyset$ ).

The sequence of iterative steps are as follows. For simplicity of notation we choose a decreasing iterate,  $l$ .

1. Let  $l = |T|$ ,  $V^l = V^*$  and  $\Pi^l = \Pi^*$ .
2. If  $V^l$  is fractional i.e.,  $\exists i \in \mathcal{I}$  s.t.  $0 < \alpha_i^l < 1$  then proceed to step 3. Otherwise, set  $\tilde{V} = V^l$ ,  $\tilde{\Pi} = \Pi^l$  and the procedure terminates.
3. For all sets  $S$  such that any agent in  $\{k_l, \dots, k_{|T|}\}$  is in  $S$ , update as follows:  $\alpha_S^{l-1} = \alpha_{k_{l-1}}^l$  where  $\alpha_{k_0} = 1$ . All other sets  $S$ , there is no change i.e.,  $\alpha_S^{l-1} = \alpha_S^l$ . Observe that  $\alpha_{k_{l-1}}^{l-1} = \alpha_{k_{l-1}}^l$ . We denote the objective value corresponding to this solution by  $\Pi^{l-1}$ .
4. Proceed back to step 2 after setting  $l := l - 1$ .

Therefore, the algorithm terminates in at most  $|T|$  steps and the final solution is such that all the  $\alpha_{S}$  are integer as  $\alpha_{k_0} = 1$ .

First observe that the updated solution in every stage is feasible to the relaxed Z-MIP. This is because all the  $\alpha_i$ s were ordered and the lowest  $\alpha_i$  value for some agent  $i$  is increased along with all the sets that agent  $i$  is part of. Therefore, constraints (3.4.3) are automatically satisfied. Constraints (3.4.4) are also satisfied because  $\alpha_i$  is set equal to  $\alpha_S$  whenever  $i \in S$ .

Next, we show that for any  $l$ ,  $\Pi_{[s,|T|]}^{l-1} \geq \Pi_{[s,|T|]}^l \geq 0 \forall s \in \{1, \dots, T\}$ . Observe that  $\Pi_{[1, l-1]}^{l-1} = \Pi_{[1, l-1]}^l$ . Therefore it suffices to show that  $\Pi_{[s, |T|]}^{l-1} \geq \Pi_{[s, |T|]}^l \geq 0 \forall s \geq l$ . We show these by induction on  $l$ . We know when  $l = |T|$ ,  $\Pi_{[s, |T|]}^l \geq 0 \forall s \geq l$  from (B.0.6). Assume it

holds for some  $l$ . Consider any  $s > l$ :

$$0 \leq \Pi_{[s,|T|]}^l = \sum_{m=s}^{|T|} \left[ \sum_{\substack{S \ni k_m \\ S \subset T_m}} \sum_{i \in S} g_{S \setminus i, i} - c \right] \alpha_{k_l}^l \leq \sum_{m=s}^{|T|} \left[ \sum_{\substack{S \ni k_m \\ S \subset T_m}} \sum_{i \in S} g_{S \setminus i, i} - c \right] \alpha_{k_{l-1}}^l \quad (\text{B.0.7})$$

$$= \sum_{m=s}^{|T|} \left[ \sum_{\substack{S \ni k_m \\ S \subset T_m}} \sum_{i \in S} g_{S \setminus i, i} - c \right] \alpha_{k_{l-1}}^{l-1} = \Pi_{[s,|T|]}^{l-1} \quad (\text{B.0.8})$$

Now consider the case when  $s = l$ :

$$0 \leq \Pi_{[l,|T|]}^l = \left[ \sum_{\substack{S \ni k_l \\ S \subset T_l}} \sum_{i \in S} g_{S \setminus i, i} \alpha_S^l - c \alpha_{k_l}^l \right] + \sum_{m=l+1}^{|T|} \left[ \sum_{\substack{S \ni k_m \\ S \subset T_m}} \sum_{i \in S} g_{S \setminus i, i} - c \right] \alpha_{k_l}^l \quad (\text{B.0.9})$$

$$\leq \sum_{m=l}^{|T|} \left[ \sum_{\substack{S \ni k_m \\ S \subset T_m}} \sum_{i \in S} g_{S \setminus i, i} - c \right] \alpha_{k_l}^l \quad (\text{B.0.10})$$

$$\leq \sum_{m=l}^{|T|} \left[ \sum_{\substack{S \ni k_m \\ S \subset T_m}} \sum_{i \in S} g_{S \setminus i, i} - c \right] \alpha_{k_{l-1}}^l = \Pi_{[l,|T|]}^{l-1} \quad (\text{B.0.11})$$

The second inequality holds because  $\alpha_S^l \leq \alpha_{k_l}^l$  and  $\sum_{\substack{S \ni k_l \\ S \subset T_l}} \sum_{i \in S} g_{S \setminus i, i} \geq 0$  by the generalized non-negativity condition (3.2.2) (and (3.2.4) which is itself derived from (3.2.2)). Using the above result when  $s = 0$  for every  $l$ , completes the proof that the objective is increasing in every iterative step starting from  $\Pi^*$  and finally converging to  $\tilde{\Pi}$ . At termination we have an integer solution  $\tilde{V}$  which is feasible to the relaxed Z-MIP and therefore, also feasible to Z-MIP with an objective  $\tilde{\Pi} \geq \Pi^*$ .

## Proof of correctness of Algorithm 1

First, we note that after each iteration of the procedure, at least one agent is removed from the network. Therefore, the algorithm clearly terminates in finite time, more precisely, at most after  $N$  iterations. Let us denote by  $I_T(\leq N)$  the total number of iterations and by  $N^{(t)}$  the number of agents in the network at iteration  $t = 1, 2, \dots, I_T$ .

Next, we show that the only candidates for the optimal uniform price are  $p_{min}^{(t)} \forall t \in \{1, \dots, I_T\}$ . First, observe that the uniform optimal price cannot be smaller than  $p_{min}^{(1)}$ . Indeed, for any price  $p \leq p_{min}^{(1)}$ , all the agents that bought in the discriminative case will still buy at this smaller price. But a lower price than  $p_{min}^{(1)}$  will result in a lower profit (per buyer) for the seller. It is possible though that some new agents buy the item at the lower price that can result in an overall higher profit. However, one can see that the new lower price in a uniform pricing scheme certainly will not be less than  $c$ . Therefore, it suffices to consider prices that are at least  $c$  but lower than  $p_{min}^{(1)}$ , if any. If this is the case it would have been profitable to offer this price (which is higher than  $c$ ) to those agents in the discriminative pricing scheme as well. But because it was not optimal to offer a lower price than  $p_{min}^{(1)}$  to the non-buyers, it is not profitable to decrease the uniform price lower than  $p_{min}^{(1)}$ . As a result, we conclude that the optimal uniform price cannot be smaller than  $p_{min}^{(1)}$ . We now consider the case where the uniform price is larger than  $p_{min}^{(1)}$ . In this case, we lose the buyers with  $p_i \leq p_{min}^{(1)}$  from the discriminative pricing scheme. Otherwise, in the discriminative case one would offer a higher price. We can therefore remove those agents from the network. Now applying the same argument, it is the case that the uniform optimal price cannot be equal to a value that is strictly between  $p_{min}^{(1)}$  and  $p_{min}^{(2)}$ . By repeating this procedure, we conclude that the optimal uniform price has to be equal to one of the  $p_{min}^{(t)}$  prices. In order to select the best uniform price among these  $I_T$  candidates, we just need to evaluate the corresponding profits (denoted by  $\Pi^{(t)} \forall t = 1, 2, \dots, I_T$ ) and choose the one that yields the maximal profits. One can do so by using the following relation:

$$\Pi^{(t)} = (p_{min}^{(t)} - c) \sum_{i=1}^{N^{(t)}} \alpha_i^{(t)}, \quad (\text{B.0.12})$$

where  $N^{(t)}$  is the remaining number of agents in the network at iteration  $t$ .

## Proof of Proposition 3.2

1. The probability that all the agents buy at the price  $p^*$  is  $\prod_i \mathbb{P}(\sum_{m \in T_i} g_{mi} \geq Lg^*)$ . Note that there are exactly  $L$  elements in every  $T_i$  which represents the neighbors of

agent  $i$ . When,  $L, N \rightarrow \infty$ , from the strong law of large numbers we have:

$$\mathbb{P} \left( \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{m \in T_i} g_{mi} = \mathbb{E}(g) \right) = 1.$$

Therefore, if  $g^* \leq \mathbb{E}(g)$ , all the agents buy almost surely and one can also show that no one buys if  $g^* > \mathbb{E}(g)$ .

The expected profit for any  $g^* \leq \mathbb{E}(g)$  (when  $L$  and  $N$  are very large) is given by:

$$\mathbb{E}[\Pi](g^*) = N(\theta + Lg^* - c).$$

This is maximized when  $g^* = \mathbb{E}[g]$  as long as  $\theta + L\mathbb{E}[g] - c \geq 0$  (otherwise, it is 0).

2. Consider the problem with  $L$  neighbors for any given  $N$  with the corresponding optimal price  $p^*(L)$ . By using  $p^*(L)$  for the problem with  $L + 1$  neighbors, for any realization of the influence factors, the profit is at least as large as the corresponding profit of the problem with  $L$  neighbors. By setting the price  $p = p^*(L) + g_1$  for the problem with  $L + 1$ , one can strictly improve the profit for any given realization. This follows from the fact that each agent has an additional neighbor that influences by at least  $g_1$  (worst case realization). As a result, we have:  $p^*(L + 1) > p^*(L)$ .
3. This part is easy to deduce from the fixed point equation in (3.6.7) when the RHS is decreasing in  $g^*$ . This is ensured by Assumption 7.

### Proof of Proposition 3.3

Consider the continuous relaxation of problem  $Zi$  that replaces the binary constraint  $\alpha_i \in \{0, 1\}$  by  $0 \leq \alpha_i \leq 1 \forall i \in \mathcal{I}$ . We consider the version of problem  $Zi$  without the dual variables  $w_i$  (see Observation 3.5). Let  $V^* = (p_i^*, d_i^*, y_i^*, \alpha_i^*, \gamma_i^*) \forall i \in \mathcal{I}$  be an optimal solution for the relaxed problem with the corresponding objective  $\Pi^*$ . We divide the proof into two parts. First, we show that given any optimal solution, one can construct a new optimal solution for which all the variables  $\alpha_i^* \forall i \in \mathcal{I}$  are integer. Second, we construct from the latter solution a new solution with all the variables  $\gamma_i^* \forall i \in \mathcal{I}$  integer as well. Assume that the initial optimal solution has at least one fractional component i.e.,  $\exists j \in \mathcal{I}$  s.t.  $0 <$

$\alpha_j^* < 1$ . Now, consider three other feasible solutions  $\tilde{V}$ ,  $\bar{V}$  and  $\underline{V}$  to the relaxed problem as follows:  $Z_i$  given by:

$$\begin{aligned} \tilde{V} = (\tilde{p}_i, \tilde{d}_i, \tilde{y}_i, \tilde{\alpha}_i, \tilde{\gamma}_i) &= \begin{cases} (p_i^*, d_i^*, y_i^*, 1, \gamma_i^*) & \text{if } i = j \\ V_i^* & \forall i \in S_j \\ V_i^* & \text{otherwise} \end{cases} \\ \bar{V} = (\bar{p}_i, \bar{d}_i, \bar{y}_i, \bar{\alpha}_i, \bar{\gamma}_i) &= \begin{cases} (p_i^*, d_i^*, 0, 1, 1) & \text{if } i = j \\ (p_i^* + (1 - \gamma_j^*)g_{ji}, d_i^*, y_i^*, \alpha_i^*, \gamma_i^*) & \forall i \in S_j \\ V_i^* & \text{otherwise} \end{cases} \\ \underline{V} = (\underline{p}_i, \underline{d}_i, \underline{y}_i, \underline{\alpha}_i, \underline{\gamma}_i) &= \begin{cases} (p_i^*, d_i^*, 0, 1, 0) & \text{if } i = j \\ (p_i^* - \gamma_j^*g_{ji}, d_i^*, y_i^*, \alpha_i^*, \gamma_i^*) & \forall i \in S_j \\ V_i^* & \text{otherwise} \end{cases} \end{aligned}$$

where  $S_j$  denotes the set of neighbors of agent  $j$  (excluding  $j$ ). We observe that all three solutions are feasible to the problem for the following reasons. First, since  $0 < \alpha_j^* < 1$  it implies that  $(g_{jj} + \sum_i \gamma_i^* g_{ij} - p_j^*) = 0$  as otherwise it cannot be a best response for agent  $j$  and cannot satisfy the equilibrium constraints. In addition, to ensure feasibility, we should have either  $\gamma_j^* = 0$  or  $d_j^* = t_j$  and therefore  $y_j^* = 0$ . Therefore, changing  $\alpha_j^*$  to 1 or 0 does not affect the best response of agent  $j$  as far as  $\alpha_j$  is concerned. Note that we construct the dual variable for agent  $j$  to satisfy all the feasibility constraints. Second, we have modified the prices of the neighbors of agent  $j$  exactly by the change in the level of influence from

agent  $j$  on them and therefore the purchasing decisions for agents  $i \in S_j$  remain the same. Third, since the agents in  $\mathcal{I} \setminus (\{j\} \cup S_j)$  are unaffected by the change in  $\alpha_j^*$  or  $p_j^* \forall i \in S_j$ , the solution remains feasible for them as well.

Let us denote the objective corresponding to these new solutions by  $\tilde{\Pi}$ ,  $\bar{\Pi}$  and  $\underline{\Pi}$  respectively. We observe that  $\Pi^* - \tilde{\Pi} = -(1 - \alpha_j^*)(p_j^* - c)$ . Since,  $V^*$  is an optimal solution, it has to be the case that  $p_j^* - c \leq 0$  because otherwise  $\tilde{V}$  is a better solution. In addition, we observe that  $\Pi^* - \bar{\Pi} = -(1 - \alpha_j^*)(p_j^* - c) + (1 - \gamma_j^*)d_j^* - \sum_{i \in S_j} \alpha_i^* g_{ji}(1 - \gamma_j^*)$  and  $\Pi^* - \underline{\Pi} = \alpha_j^*(p_j^* - c) - \gamma_j^*d_j^* + \sum_{i \in S_j} \alpha_i^* g_{ji}\gamma_j^*$ . Since  $\Pi^*$  is the optimal value of the objective and  $0 < \alpha_j^* < 1$ , it should be the case that both  $\bar{\Pi}$  and  $\underline{\Pi}$  are lower or equal than  $\Pi^*$ . By requiring  $\Pi^* - \bar{\Pi} \geq 0$  together with  $\Pi^* - \underline{\Pi} \geq 0$  and using the fact that  $p_j^* - c \leq 0$ , we obtain the condition:  $\alpha_j \geq \gamma_j$ . However, from the feasibility constraint, we know that  $\alpha_j \leq \gamma_j$  and therefore  $\alpha_j = \gamma_j$ . By using this fact, we obtain:  $\Pi^* - \bar{\Pi} = -(1 - \alpha_j^*)(p_j^* - c - d_j^* + \sum_{i \in S_j} \alpha_i^* g_{ji})$  and  $\Pi^* - \underline{\Pi} = \alpha_j^*(p_j^* - c - d_j^* + \sum_{i \in S_j} \alpha_i^* g_{ji})$ . Since  $0 < \alpha_j^* < 1$ , it has to be the case that both  $\bar{V}$  and  $\underline{V}$  are optimal solutions as well since they are feasible and yield the same objective than  $V^*$ . We therefore have reduced the number of fractional components by one. One can now repeat the same procedure for each fractional value  $\alpha_j^*$  to derive a constructive way of identifying a feasible integral solution to the original problem with an objective function that is as good as the initial fractional solution. Note that since the number of agents is finite, this step is repeated at most  $N$  times. One can conclude that the continuous relaxation of problem Zi is tight, meaning that for any feasible fractional solution, one can find an integral solution with at least the same objective if not better.

We now know that there exists an optimal solution with  $\alpha_i^*$  integer  $\forall i \in \mathcal{I}$ . We next show the following result that allows to guarantee the integrality of  $\gamma_i^* \forall i \in \mathcal{I}$  at optimality. In other words, it is optimal for each buyer to either fully influence (i.e.,  $\alpha_i^* = \gamma_i^* = 1$ ) and receive the full discount or not to influence at all (i.e.,  $\gamma_i^* = 0$ ) and pay the full price. Consider the optimal integer purchasing decisions  $\alpha_i^* \forall i \in \mathcal{I}$ . For all the agents  $k$  with  $\alpha_k^* = 0$ , it is clear from feasibility that  $\gamma_k^* = 0$ . Consider a given optimal solution denoted by  $V^*$  with  $\alpha_j^* = 1$  and assume by contradiction that  $0 < \gamma_j^* < 1$ . Consider two alternative

feasible solutions  $\bar{V}$  and  $\underline{V}$  to the relaxed problem  $Z_i$  given by:

$$\bar{V} = (\bar{p}_i, \bar{d}_i, \bar{y}_i, \bar{\alpha}_i, \bar{\gamma}_i) = \begin{cases} (p_i^*, d_i^*, y_i^*, 1, 1) & \text{if } i = j \\ (p_i^* + (1 - \gamma_j^*)g_{ji}, d_i^*, y_i^*, \alpha_i^*, \gamma_i^*) & \forall i \in S_j \\ V_i^* & \text{otherwise} \end{cases}$$

$$\underline{V} = (\underline{p}_i, \underline{d}_i, \underline{y}_i, \underline{\alpha}_i, \underline{\gamma}_i) = \begin{cases} (p_i^*, d_i^*, y_i^*, 1, 0) & \text{if } i = j \\ (p_i^* - \gamma_j^*g_{ji}, d_i^*, y_i^*, \alpha_i^*, \gamma_i^*) & \forall i \in S_j \\ V_i^* & \text{otherwise} \end{cases}$$

where  $S_j$  denotes the set of neighbors of agent  $j$  (excluding  $j$ ). We observe that both solutions are feasible to the problem for the following reasons. We first note that we construct the dual variables for agent  $j$  to satisfy all the feasibility constraints. Indeed, since  $0 < \gamma_j^* < 1$ , it has to be the case that  $d_j^* = t_j$  as otherwise it cannot be a best response for agent  $j$  and cannot satisfy the equilibrium constraints. Second, we have modified the prices of the neighbors of agent  $j$  exactly by the change in the level of influence from agent  $j$  on them and therefore the purchasing decisions for agents  $i \in S_j$  remain the same. Third, since the agents in  $\mathcal{I} \setminus (\{j\} \cup S_j)$  are unaffected by the change in  $\alpha_j^*$  or  $p_i^* \forall i \in S_j$ , the solution remains feasible for them as well.

Let us denote the objective corresponding to these new solutions by  $\bar{\Pi}$  and  $\underline{\Pi}$  respectively. We observe that  $\Pi^* - \bar{\Pi} = (1 - \gamma_j^*) \left[ d_j^* - \sum_{i \in S_j} \alpha_i^* g_{ji} \right]$  and  $\Pi^* - \underline{\Pi} = -\gamma_j^* \left[ d_j^* - \sum_{i \in S_j} \alpha_i^* g_{ji} \right]$ . Since  $\Pi^*$  is the optimal value of the objective and  $0 < \gamma_j^* < 1$ , it should be the case that  $d_j^* - \sum_{i \in S_j} \alpha_i^* g_{ji} = 0$ . Because if not, one of the solutions we constructed is strictly better than the optimal solution and this is a contradiction. Consequently, one can see that both  $\bar{V}$  and  $\underline{V}$  are optimal solutions as well since they are feasible and yield the same objective than  $V^*$ . In the process, we have therefore reduced the number of fractional components

by one. One can now repeat the same procedure for each fractional value  $\gamma_j^*$  to derive a constructive way of identifying a feasible integral solution to the original problem with an objective function that is as good as the fractional solution. Note that since the number of agents is finite, this step is repeated at most  $N$  times. In conclusion, the continuous relaxation of problem  $Zi$  always has an optimal solution such that not only the purchasing decisions are integer but the variables  $\gamma$  are integer too.

# Appendix C

## Proofs for Chapter 4

### Proof of Result 4.1

The actual inverse demand curve is  $P_A(Q)$ , and  $P_A(0) = P_m$ . One can write:  $P_A(Q) = P_m - bQ + f(Q)$ , with  $f(0) = 0$  and  $P'_A(Q) = f'(Q)$ . Equating marginal revenue with marginal cost:

$$Q^{**} = \frac{P_m - c + f(Q^{**}) + f'(Q^{**})Q^{**}}{2b}.$$

This yields an expression for the optimal price as a function of  $Q^{**}$ :

$$P^{**} = P_A(Q^{**}) = P_m - \frac{1}{2}(P_m - c + f(Q^{**}) + f'(Q^{**})Q^{**}) + f(Q^{**}).$$

Recall that  $P^* = (P_m + c)/2$  and therefore:  $P^{**} = P^* + 0.5[f(Q^{**}) - f'(Q^{**})Q^{**}]$ . From the first order Taylor expansion, we have for any differentiable function  $f(\cdot)$ :  $f(x) = f(a) + f'(a)(x - a) + R_1$ , where  $R_1 = 0.5f''(\zeta)(x - a)^2$ , for some  $\zeta \in [x, a]$ . Then:

$$f(Q^{**}) - f'(Q^{**})Q^{**} = -R_1 = \frac{f''(\zeta)}{2}(Q^{**})^2 = \frac{P''_A(\zeta)}{2}(Q^{**})^2.$$

Therefore  $P^{**} - P^* = -P''_A(\zeta)(Q^{**})^2/4$ , for some  $\zeta \in [0, Q^{**}]$ . As a result, if  $P_A(Q)$  is convex,  $P''_A(\cdot) \geq 0$  so that  $P^{**} \leq P^*$ , and if  $P_A(Q)$  is concave,  $P''_A(\cdot) \leq 0$  so that  $P^{**} \geq P^*$ .

## Proof of Result 4.2

**Convex case:** The profit  $\Pi^*$ , the optimal price  $P^{**}$  and quantity  $Q^{**}$  are:

$$\begin{aligned}\Pi^* &= \frac{P_m - c}{2} \left[ \frac{1}{2b_2} \left( b_1 - \sqrt{b_1^2 - 2b_2(P_m - c)} \right) \right] \\ P^{**} &= P_m - \frac{b_1}{3b_2} \left[ b_1 - \sqrt{b_1^2 - 3b_2(P_m - c)} \right] + \frac{1}{9b_2} \left[ b_1 - \sqrt{b_1^2 - 3b_2(P_m - c)} \right]^2 \\ Q^{**} &= \frac{b_1 - \sqrt{b_1^2 - 3b_2(P_m - c)}}{3b_2}\end{aligned}$$

The optimal profit is  $\Pi^{**} = (P^{**} - c)Q^{**}$ . One can express the profit and price ratios as functions of  $c$  and  $b_2$  and check the monotonicity to conclude that the profit and price ratios are largest when  $c = 0$  and  $b_2 = b_1^2/4P_m$ , in which case  $\Pi^*$  and  $\Pi^{**}$  are:  $P^{**} = (4/9)P_m$  and  $Q^{**} = (2P_m)/(3b_1)$ . Finally, we can now compute both profits:

$$\Pi^* = \frac{b_1 P_m}{4b_2} \left( 1 - \frac{1}{\sqrt{2}} \right) = \frac{P_m^2}{b_1} \left( 1 - \frac{1}{\sqrt{2}} \right); \quad \Pi^{**} = \frac{2b_1 P_m}{27b_2} = \frac{8P_m^2}{27b_1}$$

Then the profit and price ratios are:  $\Pi^{**}/\Pi^* = 8\sqrt{2}/[27(\sqrt{2}-1)] = 1.0116$ ;  $P^{**}/P^* = 8/9$ . These are the largest values for the ratios, so in general we have inequalities.

**Concave case:** The optimal quantity  $Q^{**}$  equates marginal revenue with marginal cost:

$$Q^{**} = \frac{-b_1 \pm \sqrt{b_1^2 - 3b_2(P_m - c)}}{-3b_2}.$$

Since  $Q^{**} > 0$ , the positive root applies. Then the optimal price is :

$$P^{**} = P_A(Q^{**}) = \frac{2P_m + c}{3} - \frac{b_1^2}{9b_2} + \frac{1}{9b_2} b_1 \sqrt{b_1^2 - 3b_2(P_m - c)}.$$

Finally, the optimal profit as a function of  $P_m$ ,  $c$ ,  $b_1$  and  $b_2$  follows from  $\Pi^{**} = (P^{**} - c)Q^{**}$ .

Our pricing rule is  $P^* = (P_m + c)/2$ , so

$$Q_A(P^*) = \frac{-b_1 \pm \sqrt{b_1^2 - 2b_2(P_m - c)}}{-2b_2}.$$

We select the positive root in order to satisfy  $Q^* > 0$ . The profit is:

$$\Pi^* = (P^* - c)Q_A(P^*) = \frac{P_m - c}{2} \left[ \frac{1}{-2b_2} \left( \sqrt{b_1^2 - 2b_2(P_m - c)} - b_1 \right) \right].$$

Expressing the profit and price ratios in terms of  $b_1$  and checking the monotonicity, one can see that the worst case for both ratios occurs when  $b_1 = 0$ . Intuitively, the larger is  $b_1$ , the more linear is the function, making the ratios closer to 1. If  $b_1 = 0$ ,  $P^{**} = (2P_m + c)/3$  and  $P^* = (P_m + c)/2$ , so

$$\Pi^{**} = \frac{2(P_m - c)}{3} \frac{\sqrt{-3b_2(P_m - c)}}{-3b_2}; \quad \Pi^* = \frac{P_m - c}{2} \frac{\sqrt{-2b_2(P_m - c)}}{-2b_2}$$

Then, the profit and price ratios are:

$$\frac{\Pi^{**}}{\Pi^*} = \frac{4\sqrt{2}}{3\sqrt{3}} = 1.0887; \quad \frac{P^{**}}{P^*} = \frac{2}{3} \frac{2P_m + c}{P_m + c} \leq \frac{4}{3} = 1.33$$

For  $b_1 > 0$ , we have inequalities for both ratios.

### Proof of Result 4.3

Equating marginal revenue and marginal cost,  $MR_A(Q^{**}) = P_m - (n+1)\gamma(Q^{**})^n = c$ . Thus:  $Q^{**} = [(P_m - c)/(n+1)\gamma]^{1/n}$ , and the optimal price is:  $P^{**} = P_A(Q^{**}) = (nP_m - c)/(n+1)$ .

Note that  $P^{**}$  is independent of  $\gamma$ . Next, the optimal profit is:

$$\Pi^{**} = (P^{**} - c)Q^{**} = \frac{n}{(n+1)^{\frac{1}{n}+1}\gamma^{1/n}} (P_m - c)^{\frac{1}{n}+1}.$$

Recall that  $P^* = (P_m + c)/2$ , so the corresponding quantity is  $Q_A(P^*) = [(P_m - c)/(2\gamma)]^{1/n}$ . Therefore,  $\Pi^* = (P^* - c)Q_A(P^*) = [(P_m - c)^{\frac{1}{n}+1}]/[2^{\frac{1}{n}+1}\gamma^{1/n}]$ . We can now compute both ratios:

$$\frac{\Pi^{**}}{\Pi^*} = \frac{2^{\frac{1}{n}+1}n}{(n+1)^{\frac{1}{n}+1}} \leq 2; \quad 1 \leq \frac{P^{**}}{P^*} = \frac{2(nP_m + c)}{(n+1)(P_m + c)} \leq 2$$

## Proof of Result 4.4

First, consider that  $c = 0$ . Equating marginal revenue and marginal cost, we obtain:  $MR_A(Q^{**}) = P_m e^{-\alpha Q^{**}} - \alpha P_m Q^{**} e^{-\alpha Q^{**}} = 0$ , so  $Q^{**} = 1/\alpha$ . Then  $P^{**} = P_m e^{-1}$ ;  $\Pi^{**} = P_m e^{-1} \alpha^{-1}$ . If the firm prices at  $P^*$ , profit is  $\Pi^* = (P^* - c)Q_A(P^*) = 0.5P_m Q_A(P^*)$ . Since  $c = 0$  and  $P^* = 0.5P_m$ , we obtain:  $Q_A(P^*) = -(1/\alpha) \log(0.5)$ . Therefore:  $\Pi^* = 0.5P_m \log(2)/\alpha$ . Therefore:

$$\frac{\Pi^{**}}{\Pi^*} = \frac{P_m e^{-1} 2\alpha}{\alpha P_m \log(2)} = \frac{2e^{-1}}{\log(2)} = 1.0615; \quad \frac{P^{**}}{P^*} = \frac{P_m e^{-1}}{P_m/2} = 2e^{-1} = 0.7357$$

We now show that when  $c > 0$ , both ratios are closer to 1. Start with the price ratio; we show that  $\frac{\partial}{\partial c} \left[ \frac{P^{**}}{P^*} \right] \geq 0$ ;  $\forall 0 \leq c \leq P_m$ . We have:

$$\frac{\partial}{\partial c} \left[ \frac{P^{**}}{P^*} \right] = \frac{\frac{\partial P^{**}}{\partial c} P^* - \frac{\partial P^*}{\partial c} P^{**}}{(P^*)^2}. \quad (\text{C.0.1})$$

For eqn. (C.0.1) to be nonnegative, we need:  $\frac{\partial P^{**}}{\partial c} \frac{1}{P^{**}} \geq \frac{\partial P^*}{\partial c} \frac{1}{P^*}$ . Recall that  $P^* = (P_m + c)/2$  and therefore:  $\partial P^*/\partial c = 0.5$ . As a result, we need to show:

$$\frac{\partial P^{**}}{\partial c} \geq \frac{P^{**}}{P_m + c}. \quad (\text{C.0.2})$$

From the first order condition:  $MR_A(Q^{**}) = P_m e^{-\alpha Q^{**}} - \alpha P_m Q^{**} e^{-\alpha Q^{**}} = P^{**}(1 - \alpha Q^{**}) = c$ . By differentiating both sides with respect to  $c$  and isolating  $\partial P^{**}/\partial c$ :

$$\frac{\partial P^{**}}{\partial c} = \frac{1 + \alpha P^{**} \frac{\partial Q^{**}}{\partial c}}{1 - \alpha Q^{**}}. \quad (\text{C.0.3})$$

Recall that  $P^{**} = P_m e^{-\alpha Q^{**}}$  and hence by differentiating with respect to  $c$ :

$$\frac{\partial P^{**}}{\partial c} = -\alpha P^{**} \frac{\partial Q^{**}}{\partial c}. \quad (\text{C.0.4})$$

By combining (C.0.3) and (C.0.4), we obtain:  $\partial P^{**}/\partial c = 1/(2 - \alpha Q^{**})$ . Since the demand curve is convex, from Result 4.1:  $P^{**} \leq P^* = (P_m + c)/2$  and therefore:  $P^{**}/(P_m + c) \leq 0.5$ . From the FOC,  $0 \leq 1 - \alpha Q^{**} \leq 1$  (so that  $P^{**} \geq c$ ). Thus  $1 \leq 2 - \alpha Q^{**} \leq 2$ , so  $1/(2 - \alpha Q^{**}) \geq 0.5$ , implying that (C.0.2) is satisfied. This concludes the proof for the price

ratio.

The same logic applies for the profit ratio, i.e.,  $\partial[\frac{\Pi^{**}}{\Pi^*}]/\partial c \leq 0$ ;  $\forall 0 \leq c \leq P_m$ .

## Proof of Result 4.5

Equating marginal revenue to marginal cost,  $MR_A(Q^{**}) = P_m \left(1 - \frac{1}{\beta}\right) (Q^{**}/Q_0)^{-1/\beta} = c$ .

Thus:  $Q^{**} = Q_0 \left[\frac{\beta c}{(\beta-1)P_m}\right]^{-\beta}$ . Note that the  $Q^{**}$  is larger than the truncation value  $Q_0$ .

The optimal price and profit are:  $P^{**} = \beta c / (\beta - 1)$ ;  $\Pi^{**} = Q_0 c / (\beta - 1) \left[\frac{\beta c}{(\beta-1)P_m}\right]^{-\beta}$ . By

requiring  $\beta \geq P_m / (P_m - c)$  we ensure that  $P^{**} \leq P_m$ . We next compute the profit by using

$P^*$ :  $\Pi^* = (P^* - c)Q_A(P^*) = 0.5(P_m - c)Q_A(P^*)$ . We have:  $Q_A(P^*) = Q_0 \left(\frac{P_m + c}{2P_m}\right)^{-\beta} \geq Q_0$ .

Then:  $\Pi^* = 0.5Q_0(P_m - c) \left(\frac{P_m + c}{2P_m}\right)^{-\beta}$ . We can now compute both ratios:

$$\frac{\Pi^{**}}{\Pi^*} = \frac{2}{(P_m/c - 1)(\beta - 1)} \left( \frac{2\beta}{(P_m/c + 1)(\beta - 1)} \right)^{-\beta}; \quad \frac{P^{**}}{P^*} = \frac{2\beta}{(P_m/c + 1)(\beta - 1)}$$

## Proof of Result 4.6

Consider any non-increasing concave inverse demand curve. We know from Result 4.1:

$P^* \leq P^{**}$ . Recall that  $P^* = 0.5(P_m + c)$  and therefore,  $P^* \leq P^{**} \leq P_m = 2P^* - c \leq 2P^*$ .

We next show the inequality for the profits:  $\Pi^{**} = (P^{**} - c)Q_A(P^{**}) \leq 2(P^* - c)Q_A(P^{**}) \leq 2(P^* - c)Q_A(P^*) = 2\Pi^*$ , where the last inequality follows from the fact that  $Q_A(\cdot)$  is non-

increasing. In conclusion, we have  $1 \leq \Pi^{**}/\Pi^* \leq 2$  and  $1 \leq P^{**}/P^* \leq 2$ .

## Expressions for Section 4.3

We next present the closed form expressions of the profit ratio  $\Pi^{**}/\Pi^*$  as a function of  $\epsilon$  for the demand models we considered, for  $c = 0$ . (Setting  $\epsilon = 0$  yields the expressions in Section 4.2.)

- **Linear:**  $P_A(Q) = P_m - bQ$      $\Pi^{**}/\Pi^*(\epsilon) = 1/(1 - \epsilon^2)$
- **Quadratic convex:**  $P_A(Q) = P_m - b_1Q + b_2Q^2$ ;  $b_1, b_2 \geq 0$  and  $b_2 \leq b_1^2/4P_m$

$$\frac{\Pi^{**}}{\Pi^*}(\epsilon) \leq \frac{8\sqrt{2}}{27(1 + \epsilon)} \frac{1}{\sqrt{2} - \sqrt{1 + \epsilon}}$$

- **Quadratic concave:**  $P_A(Q) = P_m - b_1Q + b_2Q^2$ ;  $b_1 \geq 0$  and  $b_2 \leq 0$

$$\frac{\Pi^{**}}{\Pi^*}(\epsilon) \leq \frac{4\sqrt{2}}{3\sqrt{3}} \frac{1}{(1+\epsilon)\sqrt{(1-\epsilon)}}$$

- **Monomial:**  $P_A(Q) = P_m - \gamma Q^n$   $\Pi^{**}/\Pi^*(\epsilon) = \frac{2^{\frac{1}{n}+1}n}{(n+1)^{\frac{1}{n}+1}} \frac{1}{(1+\epsilon)(1-\epsilon)^{1/n}}$

- **Semi-log:**  $P_A(Q) = P_m e^{-\alpha Q}$   $\Pi^{**}/\Pi^*(\epsilon) = \frac{2e^{-1}}{(1+\epsilon)\log(\frac{2}{1+\epsilon})}$

- **Log-log (truncated):**  $P_A(Q) = \begin{cases} P_m; & \text{if } Q < Q_0 \\ P_m(Q/Q_0)^{-1/\beta}; & \text{if } Q \geq Q_0 \end{cases}$

$$\frac{\Pi^{**}}{\Pi^*}(\epsilon) = \frac{2}{\left[\frac{P_m}{c}(1+\epsilon) - 1\right](\beta-1)} \left[ \frac{2\beta}{(\beta-1)\left[\frac{P_m}{c}(1+\epsilon) + 1\right]} \right]^{-\beta}$$

# Appendix D

## Proofs for Chapter 5

### Proof of Theorem 5.1

1. Equations (5.4.3) and (5.4.4) are obtained by applying the first order conditions on the objective function of problem (5.3.1) with respect to  $q$  and then to  $p$ .
2. We next prove the second claim about the fact that the optimal solution of the government problem is obtained when the target adoption level is exactly met. Using condition (5.4.2), one can compute the optimal value of  $p^*(r)$  by using a binary search algorithm (note that equation (5.4.3) is monotonic in  $p$  for any given value of  $r$ ). In particular, for any given  $r$ , there exists a single value  $p^*(r)$  that satisfies the optimal equation (5.4.3) and since all the involved functions are continuous, we may also conclude that  $p^*(r)$  is a continuous well defined function. As a result, the objective function of the government when using the optimal policy of the supplier is also a continuous function of  $r$ . In addition, the target level cannot be attained when  $r = 0$  by Assumption 8. We then conclude that the optimal solution of the government problem is obtained when the inequality target constraint is tight. In addition, one can see that the expected adoption target equation is monotonic in  $r$  so that one can solve it by applying a binary section method.
3. Finally, let us show the third part. For the deterministic demand model, we have:  $q_{det} = y(z_{det}) = \Gamma$ . On the other hand, when demand is stochastic, we have  $\mathbb{E}[\min(q_{sto}, D(z_{sto}, \epsilon))] = \Gamma$  so that we obtain:  $q_{sto} \geq q_{det}$ . In addition, the above

expression yields:  $y(z_{sto}) = \Gamma - K(p_{sto}) \geq \Gamma$ . Therefore, we obtain:  $y(z_{sto}) \geq y(z_{det})$ . Since  $y(z)$  is non-increasing with respect to  $z = p - r$  (from Assumption 8), we may infer the following relation for the effective price:  $z_{det} \geq z_{sto}$ . We next compute the optimal price for the deterministic model  $p_{det}$  by differentiating the supplier's objective function with respect to  $p$  and equate it to zero (first order condition):

$$\frac{\partial}{\partial p_{det}} [q_{det} \cdot (p_{det} - c)] = y'(z_{det}) \cdot (p_{det} - c) + y(z_{det}) = 0.$$

One can see that in both models, we have obtained the same optimal equation for the price:  $y(z) = -y'(z) \cdot (p - c)$ . Namely, the optimal price satisfies:  $p = c + \frac{\Gamma}{|y'(z)|}$ . We note that the previous expression is not a closed form expression as both sides depend on the optimal price  $p$ . This is not an issue as our goal here is to compare the optimal quantities in the two models rather than deriving the closed form expressions. Assuming that the deterministic part of demand  $y(z)$  is a convex function, we know that  $y'(z)$  is a non-decreasing function and then:  $0 > y'(z_{det}) \geq y'(z_{sto})$ . We then have the following inequality for the optimal prices:  $p_{sto} \leq p_{det}$ . We note that the optimal subsidy in both models does not follow such a clear relation and it will actually depend on the specific demand function. We next proceed to compare the optimal supplier's profit in both models. For the deterministic demand model, the optimal profit is given by:  $\Pi_{det} = q_{det} \cdot (p_{det} - c) = \frac{\Gamma^2}{|y'(z_{det})|}$ . In the stochastic model, the expression of the optimal profit is given by:

$$\Pi_{sto} = p_{sto} \cdot \mathbb{E}[\min(q_{sto}, D(z_{sto}, \epsilon))] - c \cdot q_{sto} = \frac{\Gamma^2}{|y'(z_{sto})|} - c \cdot (q_{sto} - \Gamma) \leq \Pi_{det}.$$

## Proof of Theorem 5.2

For the linear demand model in (5.4.7), the optimal solution of the supplier's optimization problem has to satisfy the following first order condition:

$$\bar{d} + \alpha \cdot (r + c - 2p) + \frac{c}{p} \cdot F_{\epsilon}^{-1}\left(\frac{p-c}{p}\right) + \frac{p-c}{p} \cdot \mathbb{E}[\epsilon | \epsilon \leq F_{\epsilon}^{-1}\left(\frac{p-c}{p}\right)] = 0.$$

Note that it does not seem easy to obtain a closed form solution for  $p^*(r)$ . In addition, since the previous equation is not monotone one cannot use a binary search method. Instead, one can express  $r$  as a function of  $p$ :  $r = 2p - c - \frac{\bar{d}}{\alpha} + \frac{a \cdot c^2}{\alpha \cdot p^2}$ . We next proceed to solve the government optimization problem by using the tightness of the inequality target adoption constraint:  $\mathbb{E}[\min(q^*(p^*(r), r), D(p^*(r) - r, \epsilon))] = \alpha \cdot (p^*(r) - c) = \Gamma$ . One very interesting conclusion from this analysis is that we have a very simple closed form expression for the optimal price, that is the same than for the deterministic case:

$$p_{sto} = p_{det} = c + \frac{\Gamma}{\alpha}. \quad (\text{D.0.1})$$

We can at this point derive the optimal supplier's profit for both models. In the deterministic case, the profit of the supplier is given by:  $\Pi_{det} = q_{det} \cdot (p_{det} - c) = \frac{\Gamma^2}{\alpha}$ . In the stochastic model, the optimal profit is given by:  $\Pi_{sto} = p_{sto} \cdot \mathbb{E}[\min(q_{sto}, D(z_{sto}, \epsilon))] - c \cdot q_{sto} = \frac{\Gamma^2}{\alpha} - c \cdot (q_{sto} - \Gamma) \leq \Pi_{det}$ . We next derive the optimal production level for both models. In the deterministic case, we obtained that:  $q_{det} = \Gamma$ . For the stochastic case, after substituting all the corresponding expressions we obtain:

$$q_{sto} = \bar{d} + \alpha \cdot (r_{sto} - p_{sto}) + F_{\epsilon}^{-1}\left(\frac{p_{sto} - c}{p_{sto}}\right) = \Gamma + F_{\epsilon}^{-1}\left(\frac{p_{sto} - c}{p_{sto}}\right) - K(p_{sto}) \geq q_{det}.$$

We finally compare the effect of demand uncertainty on the optimal subsidy. One can show after some appropriate manipulations that the optimal subsidy for the deterministic linear demand model is given by:  $r_{det} = \frac{2\Gamma}{\alpha} - \frac{\bar{d}}{\alpha} + c$ . For the stochastic demand model, we have the following optimal equation:  $\alpha \cdot (2p_{sto} - r_{sto} - c) - \bar{d} = K(p_{sto})$ . Hence, one can find the expression for the optimal subsidies as a function of the optimal price  $p_{sto}$ :  $r_{sto} = 2p_{sto} - c - \frac{\bar{d}}{\alpha} - \frac{1}{\alpha} \cdot K(p_{sto})$ . By replacing:  $p_{sto} = c + \frac{\Gamma}{\alpha}$  from (D.0.1), we obtain:  $r_{sto} = \frac{2\Gamma}{\alpha} + c - \frac{\bar{d}}{\alpha} - \frac{1}{\alpha} \cdot K(p_{sto}) = r_{det} - \frac{1}{\alpha} \cdot K(p_{sto}) \geq r_{det}$ .

## Proof of Corollary 5.1

1. We provide the proof of Corollary 5.1.1 by using the facts  $\frac{d\rho}{dc} \leq 0$  and  $\frac{d\rho}{d\Gamma} \geq 0$ . We then show that the production gap widens as the service level  $\rho$  increases. Note that  $q_{sto} - q_{det} = F_{\epsilon}^{-1}(\rho) - K(p_{sto}) = E[\max(F_{\epsilon}^{-1}(\rho) - \epsilon, 0)]$ . Taking the derivative with

respect to the cost  $c$ , we obtain:

$$\frac{d(q_{sto} - q_{det})}{dc} = \frac{d(F_\epsilon^{-1}(\rho))}{d\rho} \cdot \frac{d\rho}{dc} \cdot \rho = \frac{1}{f(F_\epsilon^{-1}(\rho))} \cdot \frac{d\rho}{dc} \cdot \rho \leq 0.$$

Similarly, for the target level  $\Gamma$ :

$$\frac{d(q_{sto} - q_{det})}{d\Gamma} = \frac{d(F_\epsilon^{-1}(\rho))}{d\rho} \cdot \frac{d\rho}{d\Gamma} \cdot \rho = \frac{1}{f(F_\epsilon^{-1}(\rho))} \cdot \frac{d\rho}{d\Gamma} \cdot \rho \geq 0.$$

2. To prove Corollary 5.1.2, we show that the subsidy gap decreases with respect to  $\rho$ . Note that  $r_{sto} - r_{det} = -K(p_{sto})/\alpha = -E[\min(F_\epsilon^{-1}(\rho), \epsilon)]/\alpha$ . Taking the derivative with respect to the cost  $c$ , we obtain:

$$\frac{d(r_{sto} - r_{det})}{dc} = -\frac{1}{\alpha} \left[ \frac{d(F_\epsilon^{-1}(\rho))}{d\rho} \cdot \frac{d\rho}{dc} \cdot (1 - \rho) \right] = -\frac{1}{\alpha} \cdot \frac{1}{f(F_\epsilon^{-1}(\rho))} \cdot \frac{d\rho}{dc} \cdot (1 - \rho) \geq 0.$$

Similarly, for the target level  $\Gamma$ :

$$\frac{d(r_{sto} - r_{det})}{d\Gamma} = -\frac{1}{\alpha} \left[ \frac{d(F_\epsilon^{-1}(\rho))}{d\rho} \cdot \frac{d\rho}{d\Gamma} \cdot (1 - \rho) \right] = -\frac{1}{\alpha} \cdot \frac{1}{f(F_\epsilon^{-1}(\rho))} \cdot \frac{d\rho}{d\Gamma} \cdot (1 - \rho) \leq 0.$$

3. We next present the proof of Corollary 5.1.3. We assume that  $\epsilon$  is an additive random variable with support  $[a_1, a_2]$ , not necessarily with a symmetric PDF. For the linear demand model from (5.4.7), we have:

$$r_{sto} = r_{det} - \frac{1}{\alpha} \cdot K(p_{sto}). \tag{D.0.2}$$

First, let us prove the first inequality by showing that the term on the right in equation (D.0.2) is non-positive for a general parameter  $y$ .

We have:  $\mathbb{E}[\min(y, \epsilon)] = y \cdot \mathbb{P}(y \leq \epsilon) + \mathbb{E}[\epsilon | \epsilon < y] \cdot \mathbb{P}(\epsilon < y)$ . Now, let us divide the analysis into two different cases according to the sign of  $y$ . If  $y \leq 0$ , we obtain:  $\mathbb{E}[\min(y, \epsilon)] = y \cdot \mathbb{P}(y \leq \epsilon) + \mathbb{P}(\epsilon < y) \cdot \mathbb{E}[\epsilon | \epsilon < y] \leq 0$ . In the previous equation, both terms are non-positive. For the case where  $y > 0$ , we have:  $\mathbb{E}[\min(y, \epsilon)] < \mathbb{E}[\epsilon] = 0$ . Therefore,  $\mathbb{E}[\min(F_\epsilon^{-1}(\frac{p-c}{p}), \epsilon)] \leq 0$ , showing the first inequality:  $r_{det} \leq r_{sto}$ . We now show the second inequality. We know from the optimality that:  $p \geq c$ . Let us

evaluate the expression of  $r_{sto}$  in (D.0.2) for different values of  $p$ . If  $p = c$ , we obtain:  $F_\epsilon^{-1}\left(\frac{p-c}{p}\right) = F_\epsilon^{-1}(0) = a_1 < 0$ . Then, we have:  $r_{sto} = r_{det} - \frac{1}{\alpha} \cdot \mathbb{E}[\min(a_1, \epsilon)] = r_{det} - \frac{a_1}{\alpha} > r_{det}$ . If  $p \gg c$ , we obtain:  $F_\epsilon^{-1}\left(\frac{p-c}{p}\right) \rightarrow F_\epsilon^{-1}(1) = a_2 > 0$ . Therefore, we obtain:  $r_{sto} \rightarrow r_{det} - \frac{1}{\alpha} \cdot \mathbb{E}[\min(a_2, \epsilon)] = r_{det} - \frac{1}{\alpha} \cdot \mathbb{E}[\epsilon] = r_{det}$ . Since  $r_{sto}$  is continuous and non-increasing in  $p$  for any  $p \geq c$ , the second inequality holds.

## Proof of Proposition 5.2

By applying a similar methodology as in the proof of Theorem 5.1, one can derive the following expressions (the steps are not reported for conciseness):

$$\begin{aligned} q_{det} &= \Gamma; & q_{sto} &> \Gamma \\ p_{det} &= c + \frac{1}{\alpha} \cdot \left(\frac{\bar{d}}{\Gamma}\right)^{\frac{1}{\alpha}}; & p_{sto} &= \frac{F_\epsilon^{-1}\left(\frac{p_{sto}-c}{p_{sto}}\right)}{K(p_{sto})}c + \frac{1}{\alpha} \cdot \left(\frac{\bar{d}}{\Gamma} \cdot K(p_{sto})\right)^{\frac{1}{\alpha}} \\ r_{det} &= c + \left(\frac{\bar{d}}{\Gamma}\right)^{\frac{1}{\alpha}} \cdot \left(\frac{1}{\alpha} - 1\right); & r_{sto} &= \frac{F_\epsilon^{-1}\left(\frac{p_{sto}-c}{p_{sto}}\right)}{K(p_{sto})}c + \left(\frac{\bar{d}}{\Gamma} \cdot K(p_{sto})\right)^{\frac{1}{\alpha}} \cdot \left(\frac{1}{\alpha} - 1\right) \end{aligned}$$

Here,  $K(p_{sto}) = \mathbb{E}[\min(F_\epsilon^{-1}\left(\frac{p_{sto}-c}{p_{sto}}\right), \epsilon)]$ , so that the above expressions for  $p_{sto}$  and  $r_{sto}$  are not in closed form. Indeed, in this case, one cannot analytically derive a closed form expression. However, we are still able to compare the optimal subsidy between the deterministic and stochastic settings. Since  $\mathbb{E}[\epsilon] = 1$  and  $\epsilon \geq 0$ , we have  $0 \leq K(p_{sto}) \leq 1$ . Consequently, one can see that: since  $\alpha > 1$ :  $r_{sto} \geq r_{det}$ .

## Proof of Proposition 5.3

For the linear additive demand model presented in (5.4.7), one can compute the consumer surplus for given values of  $p$ ,  $r$  and  $q$ :

$$CS_{sto}(\epsilon) = \begin{cases} \frac{D(z, \epsilon)^2}{2\alpha}; & \text{if } D(z, \epsilon) \leq q \\ \frac{D(z, \epsilon) \cdot q}{2\alpha}; & \text{if } D(z, \epsilon) > q \end{cases} = \frac{D(z, \epsilon)}{2\alpha} \cdot [\min(D(z, \epsilon), q)].$$

Therefore, we have:  $CS_{sto}(\epsilon) \geq \frac{[\min(D(z, \epsilon), q)]^2}{2\alpha}$ . By applying the expectation operator, we obtain:

$$\mathbb{E}[CS_{sto}(\epsilon)] \geq \frac{\mathbb{E}\left\{[\min(D(z, \epsilon), q)]^2\right\}}{2\alpha} \geq \frac{\left\{\mathbb{E}[\min(D(z, \epsilon), q)]\right\}^2}{2\alpha} = \frac{\Gamma^2}{2\alpha} = CS_{det}.$$

Where, the second inequality follows by Jensen's inequality (or the fact that the variance of any random variable is always non-negative). The last equality follows from the previous result that the target inequality constraint is tight at optimality.

We next compute the consumer surplus defined in (5.4.12) for the iso-elastic demand from (5.4.10). In particular, we observe that  $z_{max}(\epsilon) = \infty$  for any value of  $\epsilon$  (if we assume that  $\epsilon$  is strictly positive and finite). In addition, we have:  $z_{sto} = p_{sto} - r_{sto} = \left[\frac{\bar{d}}{\Gamma} \cdot K(p_{sto})\right]^{\frac{1}{\alpha}}$ . Note that  $K(p_{sto})$  is a deterministic constant and does not depend on the realization of the noise  $\epsilon$ . Therefore, when computing  $CS_{sto}(\epsilon)$  for a given  $\epsilon$ , since demand is multiplicative with respect to the noise, one can see that  $\epsilon$  cancels out and that simplifies the calculation. We obtain:  $CS_{sto}(\epsilon) = \frac{1}{\alpha-1} \cdot \left[\frac{\bar{d}}{\Gamma} \cdot K(p_{sto})\right]^{\frac{1-\alpha}{\alpha}} \cdot \left[\frac{\bar{d}}{\Gamma} \cdot K(p_{sto})\right] \cdot \min(D(z_{sto}, \epsilon), q_{sto})$ . Then, by taking the expectation operator, we obtain:  $\mathbb{E}[CS_{sto}(\epsilon)] = \frac{\bar{d}}{\alpha-1} \cdot \left(\frac{\bar{d}}{\Gamma}\right)^{\frac{1-\alpha}{\alpha}} \cdot (K(p_{sto}))^{\frac{1}{\alpha}} = CS_{det} \cdot [K(p_{sto})]^{\frac{1}{\alpha}}$ . Here, we have used the fact that the inequality adoption constraint is tight at optimality, that is:  $\mathbb{E}[\min(D(z_{sto}, \epsilon), q_{sto})] = \Gamma$ . In addition, this is the only term that depends on the noise  $\epsilon$ . Since we have  $0 \leq K(p_{sto}) \leq 1$ , one conclude that for any  $\alpha > 1$ :  $\mathbb{E}[CS_{sto}(\epsilon)] \leq CS_{det}$ .

### Proof of Theorem 5.3

We present first the proof for the deterministic demand model and then the one for stochastic demand. Let us first consider the unconstrained optimization problem faced by the central planner. If demand is deterministic, the objective function is given by:  $J(p, r) = q(p, r) \cdot (p - r - c)$ . We assume that demand is a function of the effective price (denoted by  $z$ ), that is:  $q(p, r) = y(p - r) = y(z)$ . Next, we compute the unconstrained optimal solution denote by  $z^*$  by imposing the first order condition:  $\frac{dJ(z)}{dz} = 0 \Rightarrow z^* = c - \frac{y(z^*)}{y'(z^*)}$ . Although, we did not derive a closed form expression for  $z^*$ , we know that it should satisfy the above fixed point equation. We now show that any unconstrained optimal solution is infeasible for the

constrained original problem since it violates the target inequality constraint:

$$q(p^*, r^*) = y(z^*) = y\left(c - \frac{y(z^*)}{y'(z^*)}\right) \leq y(c) < \Gamma.$$

We used the facts that demand is positive, differentiable and a decreasing function of the effective price (see Assumption 8). In addition, since we assumed that the target level cannot be achieved without subsidies, we have shown that the unconstrained optimal solution is not feasible. Therefore, we conclude that the target inequality constraint has to be tight at optimality, namely:  $q(p_{det}, r_{det}) = \Gamma$ . In other words, the optimal effective price and production level are the same than in the decentralized model.

We now proceed to present the proof for the case where demand is stochastic. Let us consider the constrained optimization problem faced by the central planner. We denote by  $J$  the objective function (multiplied by minus 1) and by:  $\lambda_i$ ;  $i = 1, 2, 3$  the corresponding KKT multipliers of the three constraints. The KKT optimality conditions are then given by:

$$\frac{\partial J}{\partial q} - \lambda_2 \cdot \mathbb{P}(q \leq D) = 0; \quad \frac{\partial J}{\partial p} - \lambda_1 - A \cdot \lambda_2 = 0; \quad \frac{\partial J}{\partial r} - \lambda_3 - B \cdot \lambda_2 = 0 \quad (\text{D.0.3})$$

Here,  $A$  and  $B$  are given by:

$$\begin{aligned} A &= \frac{\partial}{\partial p} \mathbb{E}[\min(q, D(z, \epsilon))] = y'(z) \cdot F_{D(z, \epsilon)}(q) \\ B &= \frac{\partial}{\partial r} \mathbb{E}[\min(q, D(z, \epsilon))] = -y'(z) \cdot F_{D(z, \epsilon)}(q) = -A \end{aligned}$$

If in addition the noise is additive, we have:  $F_{D(z, \epsilon)}(q) = F_\epsilon(q - y(z))$ . We also have:

$$\frac{\partial J}{\partial q} = c - z \cdot [1 - F_{D(z, \epsilon)}(q)]; \quad \frac{\partial J}{\partial p} = -\mathbb{E}[\min(q, D(z, \epsilon))] - z \cdot A = -\frac{\partial J}{\partial r}$$

We note that the last two equations are symmetric and hence equivalent. Equivalently, the central planner decides only upon the effective price  $z = p - r$  and not  $p$  and  $r$  separately. We next assume that  $\lambda_1 = \lambda_3 = 0$ . This corresponds (from the complementary slackness conditions) to assume that both corresponding constraints are not tight. Indeed, clearly the optimal subsidies may be assumed to be strictly positive since we assumed that when

$r = 0$ , the adoption constraint cannot be satisfied. We further assume that the supplier wants to achieve positive profits, so that the optimal price is strictly larger than the cost. Therefore, the KKT conditions (both stationarity and complementary slackness) can be written as follows:

$$c - (z + \lambda_2) \cdot [1 - F_\epsilon(q - y(z))] = 0 \quad (\text{D.0.4})$$

$$- (z - \lambda_2) \cdot y'(z) \cdot F_\epsilon(q - y(z)) = \mathbb{E}[\min(q, D(z, \epsilon))] \quad (\text{D.0.5})$$

$$\lambda_2 \cdot (\Gamma - \mathbb{E}[\min(q, D(z, \epsilon))]) = 0 \quad (\text{D.0.6})$$

We now have two possible cases depending on the value of  $\lambda_2$ . Let us investigate first the case where  $\lambda_2 = 0$ . From equation (D.0.4), we have:  $F_\epsilon(q - y(z)) = \frac{z-c}{z}$ . By using equation (D.0.5), we obtain:  $z = c - \frac{\mathbb{E}[\min(q, D(z, \epsilon))]}{y'(z)}$ . Now, we have:  $\mathbb{E}[\min(q, D(z, \epsilon))] = y\left(c - \frac{\mathbb{E}[\min(q, D(z, \epsilon))]}{y'(z)}\right) + \mathbb{E}[\min(F_\epsilon^{-1}\left(\frac{p-c}{p}\right), \epsilon)]$ . Since we assume that the function  $y(z)$  is a decreasing function of the effective price  $z$  and that both  $q$  and  $D(z, \epsilon)$  are non-negative, we obtain:  $\mathbb{E}[\min(q, D(z, \epsilon))] < y(c) + \mathbb{E}[\min(F_\epsilon^{-1}\left(\frac{p-c}{p}\right), \epsilon)] \leq y(c)$ . In the last step, we used the fact that the expression  $\mathbb{E}[\min(F_\epsilon^{-1}\left(\frac{p-c}{p}\right), \epsilon)]$  is non-positive. Therefore, we conclude that  $\mathbb{E}[\min(q, D(z, \epsilon))]$  is strictly less than  $\Gamma$ . In other words, the solution is not feasible since it violates the target inequality constraint. Hence, we must have  $\lambda_2 > 0$  and the inequality constraint is tight at optimality:  $\mathbb{E}[\min(q, D(z, \epsilon))] = \Gamma$ . Now, by using equation (D.0.4), we obtain:  $F_\epsilon(q - y(z)) = \frac{z+\lambda_2-c}{z+\lambda_2}$ . We then substitute the above expression in equation (D.0.5):  $z = c - \lambda_2 - \frac{\Gamma}{y'(z)}$ . Now, we have:  $\Gamma = y(z) + \mathbb{E}[\min(F_\epsilon^{-1}\left(\frac{z+\lambda_2-c}{z+\lambda_2}\right), \epsilon)]$ . By expressing the previous equation in terms of the effective price, we obtain:  $\Gamma = y(z) + \mathbb{E}[\min(F_\epsilon^{-1}\left(\frac{-\frac{\Gamma}{y'(z)}}{c-\frac{\Gamma}{y'(z)}}\right), \epsilon)]$ . Therefore, one can solve the previous equation and find the optimal effective price  $z$ . We note that this is exactly the same equation as in the decentralized case, so that the effective prices are the same. The optimal production levels are given by:  $q = y(z) + F_\epsilon^{-1}\left(\frac{-\frac{\Gamma}{y'(z)}}{c-\frac{\Gamma}{y'(z)}}\right)$ . Similarly, the equations are the same in both the decentralized and centralized models so that the optimal production levels are identical. Finally, we just need to show that  $\lambda_2 > 0$  in order to complete the

proof. We have:

$$\Gamma = y\left(c - \lambda_2 - \frac{\Gamma}{y'(z)}\right) + \mathbb{E}\left[\min\left(F_\epsilon^{-1}\left(\frac{-\frac{\Gamma}{y'(z)}}{c - \frac{\Gamma}{y'(z)}}\right), \epsilon\right)\right] \leq y\left(c - \lambda_2 - \frac{\Gamma}{y'(z)}\right) < y(c - \lambda_2)$$

If we assume by contradiction that  $\lambda_2 < 0$ , we obtain:  $\Gamma < y(c - \lambda_2) < y(c)$ . This is a contradiction so that:  $\lambda_2 > 0$  and the proof is complete.

## Proof of Proposition 5.4

We first consider the scenario where the supplier is non-sophisticated. In this case, the optimal decision variables are still  $r_{det}$ ,  $q_{det}$  and  $p_{det}$ . However, in reality demand is uncertain and therefore the expected sales are given by:

$$\mathbb{E}\left[\min(q_{det}, D(z_{det}, \epsilon))\right] = \mathbb{E}\left[\min(q_{det}, y(z_{det}) + \epsilon)\right] = \Gamma + \mathbb{E}[\min(0, \epsilon)] \leq \Gamma. \quad (\text{D.0.7})$$

Here, we have used the fact that:  $q_{det} = y(z_{det}) = \Gamma$ .

Next, we assume that the supplier is sophisticated. Note that in this case, the optimal subsidies set by the government are still equal to  $r_{det}$ . In other words, the government does not have any distributional information on demand uncertainty and believes neither does the supplier. In particular, the subsidies are set such that:  $y(p_{det} - r_{det}) = \Gamma$ . However, the supplier is sophisticated in the sense that he uses distributional information on demand uncertainty in order to decide the optimal price and production. In a similar way as in equations (5.4.3) and (5.4.4) from Theorem 5.1, the optimal price when  $r = r_{det}$  can be obtain as the solution of the following non-linear equation:

$$y(p - r_{det}) + \mathbb{E}\left[\min\left(F_\epsilon^{-1}\left(\frac{p - c}{p}\right), \epsilon\right)\right] + y'(p - r_{det}) \cdot (p - c) = 0. \quad (\text{D.0.8})$$

In addition, one can compute the optimal production level as follows:

$$q^*(p, r_{det}) = y(p - r_{det}) + F_\epsilon^{-1}\left(\frac{p - c}{p}\right).$$

For the linear model, (D.0.8) becomes:  $\bar{d} - \alpha \cdot (2p - r_{det} - c) + \mathbb{E}\left[\min\left(F_\epsilon^{-1}\left(\frac{p - c}{p}\right), \epsilon\right)\right] = 0$ .

Equivalently, the optimal price denoted by  $p^*$  follows the following relation:

$$p^* = c + \frac{\Gamma}{\alpha} + \frac{1}{2\alpha} \cdot \mathbb{E}\left[\min\left(F_\epsilon^{-1}\left(\frac{p^* - c}{p^*}\right), \epsilon\right)\right].$$

Note that the optimal price when demand is deterministic is equal to  $p_{det} = c + \frac{\Gamma}{\alpha}$  and therefore:  $p^* \leq p_{det}$ . As a result, the expected demand is given by:

$$y(p^* - r_{det}) = \bar{d} - \alpha \cdot (p^* - r_{det}) = \Gamma - \frac{1}{2} \cdot \mathbb{E}\left[\min\left(F_\epsilon^{-1}\left(\frac{p^* - c}{p^*}\right), \epsilon\right)\right] \geq \Gamma.$$

We next proceed to compute the expected sales:

$$\begin{aligned} \mathbb{E}\left[\min(q^*, D(p^* - r_{det}, \epsilon))\right] &= \mathbb{E}\left[\min\left(y(p^* - r_{det}) + F_\epsilon^{-1}\left(\frac{p^* - c}{p^*}\right), y(p^* - r_{det}) + \epsilon\right)\right] \\ &= \Gamma + \frac{1}{2} \cdot \mathbb{E}\left[\min\left(F_\epsilon^{-1}\left(\frac{p^* - c}{p^*}\right), \epsilon\right)\right] \leq \Gamma. \end{aligned}$$