

Optimization Driven Approaches for Subsidy Allocation and Supply Chain Procurement

by

Gonzalo Romero

B.Sc., University of Chile (2008)

Civil Industrial Engineer, University of Chile (2008)

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

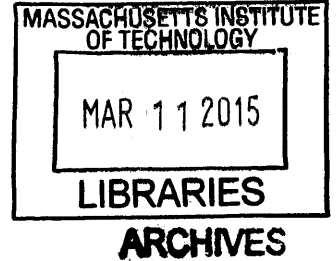
Doctor of Philosophy in Management

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2014

© Massachusetts Institute of Technology 2014. All rights reserved.



Signature redacted

Author
Alfred P. Sloan School of Management

Signature redacted August 8, 2014

Certified by
Retsef Levi

J. Spencer Standish (1945) Professor of Management
Thesis Supervisor

Certified by
Georgia Perakis

William F. Pounds Professor of Management
Thesis Supervisor

Accepted by . **Signature redacted**

Ezra W. Zuckerman
Director, Sloan School of Management PhD Program

Optimization Driven Approaches for Subsidy Allocation and Supply Chain Procurement

by

Gonzalo Romero

Submitted to the Alfred P. Sloan School of Management
on August 8, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Management

Abstract

This thesis introduces several new models in operations management, that are motivated by practical settings. It studies these models in an optimization-driven approach, employing mathematical programming techniques to derive important structural and algorithmic insights on the corresponding problems.

In the first part of the thesis, we study subsidy allocation problems under budget constraints and endogenous market response, where the central planner's objective is to maximize the market consumption of a good. We first consider co-payment subsidies, that are paid to manufacturing firms per unit sold. We focus on "uniform co-payments", in which each firm receives the same co-payment, regardless of its cost structure, or efficiency. Uniform co-payments are frequently implemented in practice. Therefore, a natural question is whether uniform co-payments are in fact the best that the central planner can do; or, more generally, how do they perform compared to the optimal co-payment allocation? Notably, we first identify relatively general sufficient conditions such that uniform co-payments are optimal, even if the firms are heterogeneous, and if the central planner is uncertain about the market response. We then complement the effectiveness of uniform co-payments, by studying a very relevant setting where they are not optimal. We show that, for any instance of this model, uniform co-payments are guaranteed to induce at least 85% of the optimal market consumption. In summary, uniform co-payments turn out to be *surprisingly* powerful in maximizing the market consumption of a good. We then consider lump sum subsidies, which are an alternative subsidy mechanism also implemented in practice. We show that the problem of optimally allocating lump sum subsidies is *NP*-hard, and discuss two simple allocation policies that have good performance guarantees.

In the second part of the thesis, we introduce a model to incorporate the cost of handling orders at a central distribution center, into the procurement decisions of a company. We show how structural results for this model lead to a practical method to select the best case pack size per SKU in procurement contracts, as well as to serve orders at the distribution center. Furthermore, we test this method on real data from a large utility company, finding significant total cost reductions.

Thesis Supervisor: Retsef Levi
Title: J. Spencer Standish (1945) Professor of Management

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

Acknowledgments

First, I would like to thank my advisors, Georgia Perakis and Retsef Levi, for all their help and guidance during these 5 years. Their passion for research and teaching, as well as their broad interests and dedication, have set a very high bar for me as an academic. I appreciate how they always made me feel their support, while at the same time they pushed me to be the best researcher I can be. For all this, I will always be thankful.

I also want to thank the other members of my Thesis Committee, Mahesh Nagarajan and Steve Graves. I admire their research and their sharpness. It was a privilege to have them in my committee. I thank all the faculty at MIT that I had the pleasure to interact with, and learn from; specially Vivek Farias, for his generosity and encouragement, and Gabriel Bitran, for welcoming me to the program, and for comforting me during the ups and downs of my first two years at MIT.

I want to thank Florin (or Dragos?), for his generous friendship, and particularly for sharing my taste for dumb jokes. I also thank all my friends in the Ph.D. program, it has been a long journey together, and their friendship was fundamental for me to have such a great experience. The list is too long to include it here, but I want to specially thank Andre, Fernanda and Leon. Additionally, a fundamental part of living in Westgate was having good friends visiting us with their families: Hordur and Gyda, Ángel and Carla, Ami and Esti, Edwin and Julie, Rodrigo and Tati, and Johann and Anna, we are grateful to you, and we treasure your friendship.

I am grateful to my academic mentor at the Industrial Engineering Department at the University of Chile, Andrés Weintraub, for encouraging me to pursue a Ph.D. at MIT. Together with Rafael Epstein, and Guillermo Durán, they initiated me in operations research and operations management. It is thank to them, and their example, that I decided I wanted to be an academic in the first place.

I am also grateful to my parents, Gonzalo and Marité, my siblings Javiera, Camila and Matías, my family in law, specially Arturo, and my friends in Chile, for their continuous support and encouragement during all my studies. I am happy that we

had the pleasure to have most of them visiting us, at some point or another, during these years. Discovering, and sharing with them, many of the places New England has to offer was an unforgettable experience.

Finally, and most importantly, I thank my family: there are no words to describe how important it is for me to have Maca and Martina by my side. Nothing makes me happier than knowing that this has been a great experience for them too. We will miss Boston, but I know that we will continue to be very happy in the next chapter of our lives. This thesis is dedicated to them.

Contents

Introduction	15
I Subsidy Allocation with Endogenous Market Response	18
1 On the Effectiveness of Uniform Co-payments in Maximizing Market Consumption	19
1.1 Introduction	19
1.1.1 Application: Global Subsidy for Antimalarial Drugs	22
1.2 Literature Review	25
1.3 Model	30
1.3.1 Co-payment Allocation Problem	33
1.3.2 Special Cases	35
1.3.3 An Upper Bound Problem	39
1.4 Optimality of Uniform Co-payments	41
1.4.1 Sufficient Condition for Optimality	42
1.5 Incorporating Market State Uncertainty	44
1.5.1 An Upper Bound Problem	47
1.5.2 Optimality of Uniform Co-payments	48
1.6 Maximizing Social Welfare	52
1.6.1 Optimality of Uniform Co-payments	54
1.7 Numerical Results	56

1.7.1	Results for some Special Cases from Section 1.3.2	56
1.7.2	Results for Cournot Competition with Nonlinear Demand . . .	61
1.8	Conclusions	64
2	85% Worst-Case Performance Guarantee for Uniform Co-Payments	67
2.1	Introduction	67
2.1.1	Main Contributions	68
2.2	Literature Review	69
2.3	Model	71
2.3.1	Optimal Co-payments Allocation	74
2.3.2	Uniform Co-payments Allocation	78
2.3.3	Consistency Constraints on Parameters a, b, c_i and B	79
2.4	Preliminary Results and Problem Statement	81
2.4.1	Problem Statement	83
2.5	Worst-Case Performance of Uniform Co-payments	85
2.5.1	Candidate to Worst Case Instance	85
2.5.2	Tight Worst-Case Performance Guarantees	86
2.6	Auxiliary Results	93
2.7	Conclusions	97
3	A Continuous Knapsack Problem with Separable Convex Utilities: Approximation Algorithms and Applications	99
3.1	Introduction	99
3.2	Problem formulation	100
3.2.1	A simple 1/2-approximation algorithm	102
3.2.2	An $(1 - e^{-1})$ -approximation algorithm	107
3.3	Application: Allocating technology subsidies to minimize a good's market price	111
3.4	Conclusions	117

II	Supply Chain Procurement	119
4	Optimizing Purchasing and Handling Costs for Supply Chain Procurement	121
4.1	Introduction	121
4.1.1	Main Contributions	123
4.2	Related Literature	125
4.3	Model	126
4.3.1	DP Formulation for Selecting the Best Case Pack Size	128
4.4	Structural Insights	132
4.4.1	Selecting the Best Case Pack Size per SKU	133
4.4.2	Selecting Multiple Case Pack Sizes per SKU	135
4.5	Numerical Experiments on Real Data	142
4.6	Conclusions and Future Work	146
	Conclusions	147
A	Proofs	149
A.1	Chapter 1	149
A.2	Chapter 2	157

List of Figures

1-1	Boxplot of the Relative Performance of Uniform Co-payments - Cournot Competition with Constant Marginal Costs	57
1-2	Boxplot of the Relative Performance of Uniform Co-payments - Price Taking Firms	59
1-3	Boxplot of the Relative Performance of Uniform Co-payments - Cournot Competition with Nonlinear Demand	61
2-1	Worst-Case Guarantee for Uniform Co-Payments	87
4-1	Empirical Distribution of the Order Size for the Square Washer in 2013	143
4-2	Purchasing and Handling Costs Induced per Case Pack Size	144
4-3	Summary of the Experimental Results	145

List of Tables

1.1	Summary Statistics of the Relative Performance of Uniform Co-payments	
	- Cournot Competition with Constant Marginal Costs	57
1.2	Summary Statistics of the Relative Performance of Uniform Co-payments	
	- Price Taking Firms	59
1.3	Summary Statistics of the Relative Performance of Uniform Co-payments	
	- Cournot Competition with Nonlinear Demand	60
4.1	An Instance of the Problem of Selecting Multiple Case Pack Sizes . .	137
4.2	Examples of Optimal Solutions to Problem (<i>IP</i>)	137

Introduction

This thesis introduces several new models in operations management, that are motivated by practical settings. It studies these models in an optimization-driven approach, employing mathematical programming techniques to derive important structural insights on the corresponding problems, as well as developing new, and theoretically efficient, algorithms to solve them. The main research question addressed in Part I of the thesis is, how should a central planner increase the market consumption of a good, by allocating subsidies to its competing and selfish producers, when there is a budget constraint? The motivation to allocate such subsidies stems from the positive societal externalities generated by the aggregated market consumption, and from the fact that, left alone, the resulting market equilibrium induced by the selfish competing producers might not be socially optimal. To address this problem, we study different subsidy allocation problems under budget constraints and endogenous market response. We focus on settings where, due to practical reasons, it is more convenient to allocate subsidies to the firms producing the good rather than to the consumers.

We first consider co-payment subsidies, that are paid to the manufacturing firms for each unit sold in the market. We focus on “uniform co-payments”, in which each firm receives the same co-payment, regardless of its cost structure and efficiency. Uniform co-payments are frequently implemented in practice. Therefore, a natural question is whether uniform co-payments are in fact the best that the central planner can do; or, more generally, how do they perform in maximizing the market consumption of a good compared to the optimal co-payment allocation? Notably, in Chapter 1 we identify relatively general sufficient conditions on the firms’ cost structure, under

which uniform co-payments are optimal, even if the firms' efficiency levels are arbitrarily different. Moreover, we show that this insight is preserved, under slightly less general conditions, even when the central planner faces uncertainty about the endogenous market response, or when the objective of the central planner is to maximize the social welfare. We then complement the effectiveness of uniform co-payments, by presenting extensive simulation results in relevant settings where they are not optimal. The computational experiments suggest that uniform co-payments induce a market consumption that is, on average, very close to the optimal.

In Chapter 2 we focus on the important case of Cournot competition with linear demand and constant marginal costs. This is a fundamental model, that generally provides interesting insights. For this model, we characterize the optimal co-payment allocation, which consists of giving larger co-payments to less efficient firms. We argue that this policy is hard to implement in practice, and thus we study the performance of the more practical, and conceptually simpler, uniform co-payments. We show a tight worst-case parametric performance guarantee, which depends on the total number of firms in the market. This leads to an *asymptotically tight* 85% uniform worst-case performance guarantee. Namely, we show that uniform co-payments will induce at least 85% of the market consumption induced by optimal co-payments, in *any instance* of this model. Taken together, the results of the first two chapters of the thesis suggest that uniform co-payments are *surprisingly* powerful in maximizing the market consumption of a good. Therefore, the decision makers facing this type of problems should not spend time and resources designing a more sophisticated co-payments allocation policy, as the simple uniform co-payments policy is likely to provide most of the potential benefits.

In Chapter 3 we consider lump sum subsidies, which are an alternative subsidy mechanism also implemented in practice. We model this problem as a novel application of a continuous knapsack problem with separable convex utilities. We show that the problem is *NP*-hard, and we provide two simple algorithms that have worst-case performance guarantees, as well as a practical interpretation. Moreover, we identify special settings where these simple algorithms are actually optimal. These results sug-

gest that simple subsidy allocation policies have a good performance in minimizing the market price of a good.

In part II of the thesis we focus on supply chain procurement. Procurement decisions are often made in a silo, without taking into consideration the effect that they might have on the internal supply chain costs of the company. In Chapter 4, we introduce a novel optimization framework to incorporate the cost of handling orders at a central distribution center, into the procurement decisions. Specifically, our model explicitly considers the effects of the case pack selection in procurement contracts, on the purchasing and handling costs of a company.

We show how structural results for this model lead to a practical method to select the best case pack size per SKU, as well as to serve orders at the distribution center. Furthermore, we test our method on real data from a large utility company. The simulation results suggest that our method has the potential to significantly reduce the purchasing and handling costs for the company. Importantly, the optimal policy suggested by our method is simple to implement, and to communicate. It only requires to compare the easily computable long run average purchasing and handling costs induced by each available case pack size, therefore facilitating the incorporation of the distribution center's handling costs into the procurement department decisions.

We additionally consider the problem of choosing multiple case pack sizes per SKU. For this problem we show that, under some assumptions, selecting at most three sizes can provide a guaranteed performance when compared to the optimal policy. This is important because the optimal policy can potentially imply selecting every case pack size available from the supplier, making it unlikely to be applied in practice.

Part I

Subsidy Allocation with Endogenous Market Response

Chapter 1

On the Effectiveness of Uniform Co-payments in Maximizing Market Consumption

1.1 Introduction

In this chapter we provide a new modeling framework to analyze a subsidy allocation problem with endogenous market response, under a budget constraint on the total amount of subsidies that the central planner can pay. The central planner's objective is to maximize the aggregated market consumption of a good. Using our framework, we identify sufficient conditions on the firms' marginal cost functions, such that *uniform subsidies are optimal*. That is, the simple policy that allocates the same subsidy to every firm is optimal, even if the firms are heterogeneous, and their efficiency levels are arbitrarily different. This is an important insight because uniform subsidies is a policy commonly used in practice, primarily because of its simplicity and perceived fairness. Moreover, we prove that, in many cases, uniform subsidies do not only obtain the optimal aggregated market consumption, but at the same time obtain the best social welfare solution. Furthermore, we show that the optimality of uniform subsidies is usually preserved, even if the central planner is uncertain about the spe-

cific market conditions. Finally, we present simulation results in relevant settings where uniform subsidies are not optimal. They suggest that the aggregated market consumption induced by uniform subsidies is relatively close to the one induced by optimal subsidies.

We study the important setting in which a central planner aims to impact a given market. Specifically, her goal is to increase the aggregated market consumption of a good, by providing *co-payment subsidies*, which are paid for each unit that is produced to competing (profit maximizers) heterogeneous firms. The motivation to provide such subsidies stems from the positive societal externalities that can be obtained by increasing the aggregated market consumption, and from the fact that left alone the resulting market equilibrium induced by the selfish competing producers might not be socially optimal. A current example are the recent efforts around the production of infectious disease treatments to the developing world, such as antimalarial drugs (e.g., Arrow et al. (2004)), and vaccines (e.g., Snyder et al. (2011)).

Furthermore, typically the central planner makes her subsidy allocation in the presence of a budget constraint, which is often determined prior to the actual subsidy allocation decision. For example, in some cases the central planner could be a foundation that raised a certain amount of money to address a related issue, and it is then facing the challenge of how to allocate the budget towards co-payment subsidies. Another challenge typically faced by the central planner is that the intervention in the market through the allocation of subsidies will likely change the market equilibrium induced by the competing producers. Hence, to optimally allocate the subsidies, the central planner has to take into account these complex dynamics.

We propose a novel modeling framework to study strategic and operational issues related to co-payment subsidies allocation. The models that we develop explicitly capture the setting of a central planner aiming to maximize the aggregated market consumption of a good, in the presence of a budget constraint, and market competition between heterogeneous profit maximizing firms. The firms are heterogeneous in terms of their respective efficiency and cost structure. This is modeled through firm-specific marginal cost functions. The models that we develop fall into the class

of *Mathematical Program with Equilibrium Constraints* (MPEC). They are relatively general and capture different cost structures, inverse demand functions, as well as a range of market dynamics of quantity competition that are typical to the settings being studied. For example, the models capture as special cases Cournot Competition with linear demand, as well as Cournot Competition under yield uncertainty with linear demand and linear marginal cost functions. MPEC models are typically computationally challenging, both to solve optimally and to analyze, see for example Luo et al. (1996). However, by reformulating these problems, we are able to develop tractable mathematical programs that provide upper bounds on the optimal objective value, and allow the development of efficient algorithms. Even more importantly, they allow analyzing the effectiveness of practical policies. In particular, the chapter focuses attention on the effectiveness of the commonly used *uniform co-payments*, in which the per-unit co-payment is the same for all competing firms in the market. The common use of uniform co-payments, in spite of the existence of heterogeneous firms, each with potentially different efficiency level, is primarily driven by the simplicity of implementation, as well as some notion of fairness. The chapter addresses the important question of to what extent uniform co-payments are effective in increasing the aggregated market consumption, compared to potentially more sophisticated policies that could allow the co-payment to be firm-specific. Through the mathematical programming upper bound relaxation that we develop, the chapter provides some surprising insights. First, we can show that for a large class of firm-specific cost structures, uniform co-payments are in fact *optimal*. That is, there is no loss of efficiency in using uniform co-payments in these settings compared to any other possible co-payment allocation. Second, this insight is maintained even if one considers the case in which there exists uncertainty about the future market state, and the central planner has to set up the subsidies prior to the realization of the market condition. Third, in many cases uniform subsidies do not only obtain the optimal (maximal) aggregated market consumption, but at the same time obtain the best social welfare solution. Finally, in other settings, where uniform subsidies are not optimal, extensive computational experiments suggest that they still perform, on average, very close to

optimal.

To demonstrate the applicability of the model and the relevance of the issues studied in the chapter, we next discuss in detail the case of antimalarial drugs.

1.1.1 Application: Global Subsidy for Antimalarial Drugs

A motivating example, where the setting modeled in this chapter is observed in practice, is the global fight against malaria. This has been a long standing challenge for the healthcare industry. It is estimated that in 2012 about 200 million cases occurred worldwide, and more than 600,000 people died of malaria, see the world malaria report by the World Health Organization (2013). To make matters worse, recently chloroquine, the traditional drug for treating malaria, has become less effective due to growing resistance to this medication. Artemisinin combination therapies (ACT) have been identified as the successor drugs to chloroquine in order to treat malaria; however, they are at least ten times more costly, see White (2008).

In 2004 the Institute of Medicine (IoM) reviewed the economics of antimalarial drugs. It identified that several manufacturers compete in an unregulated market, and concluded that the most effective way of ensuring access to ACTs for the greatest number of patients would be to provide a centralized subsidy to the producers. The goal would be achieving high overall coverage of ACTs, therefore, the subsidized price to the end user should be at least as low as chloroquine's. Moreover, the IoM recognized that firms had not invested in producing ACTs on the scale needed to supply Africa, because there had been no assured market, therefore, the global capacity to produce ACTs was quite limited, see Arrow et al. (2004).

In this context, the Roll Back Malaria Partnership and the World Bank, developed in 2007 the Affordable Medicines Facility for malaria (AMFm), a concrete initiative to improve access to safe, effective, and affordable antimalarial medicines. In 2008, the Global Fund started hosting and managing the AMFm, which began operations in July 2010. By July 2012, the AMFm had managed a budget of US\$336 millions -pledged by UNITAID, the governments of the United Kingdom and Canada, and the Bill & Melinda Gates Foundation- to pursue its main objective: increasing the con-

sumption of ACTs, as detailed in their evaluation report online AMFm Independent Evaluation Team (2012).

As usually implemented in practice, the policy proposed by AMFm consisted of giving a *uniform co-payment*, see AMFm Independent Evaluation Team (2012). Namely, each firm receives the same co-payment, for each unit sold, regardless of any differences among them. Moreover, there are 11 firms participating in the AMFm program, and they range from large pharmaceuticals like Novartis, with manufacturing plants in USA and China, and Sanofi, with manufacturing plants in Germany and Morocco, to smaller firms with manufacturing plants in Uganda, India and Korea, see the market intelligence aggregator, funded by UNITAID, A2S2 (2014). Note that the firms receiving the uniform co-payments are highly heterogeneous, both in their market size and location-wise. One additional relevant characteristic of the AMFm program is that all the ACT manufacturers that receive co-payments commit to supply antimalarials on a no profit/no loss basis, see the report by Boulton (2011). Giving the right incentives to the firms producing these drugs can increase access to them, hence, it has the potential to have a significant impact on this global problem, see Arrow et al. (2004).

Results and Contributions. The main contributions of this chapter are the following:

New modeling framework for a subsidy allocation problem. We introduce a general optimization framework to analyze subsidy allocation problems with endogenous market response, under a budget constraint on the total amount of subsidies the central planner can pay. The central planner's objective is to maximize the aggregated market consumption of a good. Our models allow general inverse demand and marginal cost functions, assuming only that the inverse demand function is decreasing in the aggregated market consumption, and that the firms' marginal costs are increasing. These are standard assumptions in the literature. In fact, they are more general than assumptions usually considered.

Sufficient conditions for the optimality of uniform co-payments. We compare

uniform co-payments to the optimal, and potentially differentiated, co-payment allocation, which provides more flexibility, but it is potentially significantly harder to implement. The main result in this chapter shows that uniform co-payments are in fact *optimal* for a large family of marginal cost functions. This family of marginal cost functions includes homogeneous functions of the same degree as a special case. This result is surprising, considering that firms are heterogeneous, and particularly since the assumptions on the inverse demand function are very general (essentially only monotonicity and continuity). More importantly, it establishes sufficient conditions such that the policy that is frequently being used in practice is actually optimal. Additionally, we provide sufficient conditions for uniform co-payments to simultaneously maximize the social welfare. In particular, we show that homogeneous functions of the same degree satisfy these conditions as well.

Incorporate market state uncertainty. We extend the models by assuming that the central planner does not know the exact market state with certainty (i.e., the specific inverse demand function is uncertain), but she has a set of possible scenarios, and beliefs on the likelihood that each scenario will materialize. We model this setting as a stochastic MPEC, where the central planner decides her co-payment allocation policy with the objective of maximizing the expected aggregated market consumption. This model is considerably harder to analyze, see Patriksson and Wynter (1999). However, we show that uniform co-payments are still optimal in this setting, for a large family of firms' marginal cost functions. In particular, this family includes convex homogeneous functions of the same degree. Moreover, the analysis suggests that the central planner only needs to consider the scenario with the highest aggregated market consumption at equilibrium, regardless of the exact distribution over the different market states.

Tractable upper bound problems. Based on an innovative mathematical programming reformulation of our model, we develop tractable upper bound prob-

lems. These are used extensively in the analysis mentioned above. In addition, we use them to conduct a numerical study of the performance of uniform co-payments in relevant settings where they are not optimal. Specifically, we consider Cournot Competition with linear demand and constant marginal costs, and a more general setting with non-linear demand, and non-linear marginal cost. The results obtained on data generated at random suggest that the aggregated market consumption induced by uniform subsidies is on average 96% optimal. We believe that the innovative reformulation of the model, and the resulting upper bounds, would be useful to study additional interesting and important research questions.

The rest of the chapter is structured as follows. Section 1.2 reviews related literature from operations management and economics. In Section 1.3 we present our model, the uniform co-payments allocation problem, and a relaxation of this problem. Section 1.4 presents the main result on sufficient conditions for the optimality of uniform co-payments in the deterministic model. In Section 1.5 we extend our model to consider the case when the central planner is uncertain about the market state. Section 1.6 considers the alternative objective of maximizing social welfare, and presents sufficient conditions for the optimality of uniform subsidies. Section 1.7 presents a numerical study of the relative performance of uniform subsidies in settings where they are not optimal. Finally, Section 1.8 provides concluding remarks.

1.2 Literature Review

The subject of taxes and subsidies allocation and incidence has a vast literature in the economics community. Fullerton and Metcalf (2002) present a thorough review of classical and recent result in this area. The main areas of research in this literature are imperfect competition, partial equilibrium models, and general equilibrium models. This chapter is closely related to the study of subsidies in imperfect competition models. However, the traditional approach in this literature assumes homogeneous firms, and focuses on studying the impact of taxes, or subsidies, on the number of firms

participating in the market in a symmetric equilibrium, see Fullerton and Metcalf (2002). Alternatively, models of differentiated products are considered, which give the firms some monopoly power, and the focus is again on the number of firms active in the market in equilibrium. The reason for this is that the number of competitors in the market is directly related to the ability to pass taxes forward to the consumer. More generally, when analyzing comparative static properties in oligopoly models, like the subsidy allocation in our case, it is fairly common to focus on symmetric equilibria with homogeneous firms in order to obtain more precise insights, see, for example, Vives (2001). In contrast, in our model we take an operational view: we assume heterogeneous firms that produce a commodity, and we focus on the specific subsidy allocation among them. Additionally, an important modeling characteristic we consider is the presence of a budget constraint, in terms of the total amount of funding that can be allocated to these subsidies. This feature allows us to investigate the interplay between the optimal subsidies structure, and the budget available.

Within theoretical research in economics, one particular area that studies a problem related to the one considered in this chapter is the strategic trade policy literature, particularly the “third market model”, see Brander (1995). In this model, n home firms and n^* foreign firms export a commodity to a third market, where the market price is set through Cournot Competition, with constant marginal costs, among all the firms. The government can allocate subsidies to the home firms, increasing their profit at the expense of the foreign competitors. The government’s utility is equal to the profit earned by the home firms, minus the cost of the subsidy payments. Let us emphasize that the government does not face a budget constraint, and that the firms’ profit is equally weighted with the cost of the subsidy payments. An exception to the latter is found in Leahy and Montagna (2001), where the cost of the subsidy payments is weighted by a parameter δ , interpreted as the social cost of funds. An alternative interpretation of δ is to let it be the Lagrange multiplier of a budget constraint for the government, relating it to our model. We focus here on the case with heterogeneous firms. In this setting, Collie (1993) and Long and Soubeyran (1997) *assume* a uniform subsidy and study its effect in the market shares of the firms. Later, Leahy

and Montagna (2001) assume linear demand, and derive closed form expressions for the optimal subsidies. They conclude that the optimal subsidy policy is generally not uniform; and if the social cost of funds is sufficiently low then the government should allocate higher export subsidies to more efficient firms. Note that this result is consistent with our numerical study in Section 1.7, where effectively uniform subsidies are not optimal for Cournot competition with linear demand and constant marginal costs in our model. Nonetheless, we find evidence that the relative performance of uniform subsidies is very good. In contrast, in our model we assume more general increasing marginal cost functions, and find conditions under which uniform subsidies are optimal.

On the other hand, the economics literature in this area has shifted towards empirical research. In particular, Cohen et al. (2014) show in a recent randomized controlled trial in Kenya that a very high subsidy for ACT antimalarials dramatically increases access to them, and they suggest that this program should be complemented with the introduction of rapid malaria tests over-the-counter to reduce the risk that the treatment goes to patients without malaria. This is an important insight, as it shows empirically that subsidies for ACTs, as the one considered as a motivation in Section 1.1.1, work in practice. Similarly, Dupas (2014) also showed in a randomized controlled trial in Kenya that short-run subsidies for an antimalarial bed net had a positive impact on the willingness to pay for the bed net a year later. This is in contrast to the belief that consumers may anchor around the subsidized price and become unwilling to pay more for the product later. This result suggests that short run subsidy programs, as the one also considered to increase the consumption of ACTs, are expected to be beneficial in the long run as well.

Mathematical Programs with Equilibrium Constraints (MPECs) are very hard to solve and analyze, both in practice and in theory. In particular, even the simplest case with linear demand and linear constraints is NP-hard, see Luo et al. (1996) Luo et al. (1996). Moreover, Stochastic Mathematical Programs with Equilibrium Constraints can be even harder to solve in practice, and are as hard as their deterministic counterparts in theory, see Patriksson and Winter (1999) Patriksson and

Wynter (1999). In this context, the best that we can hope for is to identify interesting structure in particular cases that may lead to structural or algorithmic results. The co-payment allocation problem (*CAP*) presented and analyzed in Sections 3 and 4 is a particular case of an MPEC, while the co-payment allocation problem with market uncertainty (*SCAP*) presented and analyzed in Section 5 is particular case of an SMPEC. In both cases, our main methodological contribution consists in identifying a fairly general model of a practical problem, whose structure allows us to prove surprising structural results, such as the optimality of uniform subsidies for a family of marginal cost functions and for any inverse demand function. Examples in the operations research and operations management literature that study similar models, and give structural or algorithmic results include DeMiguel and Xu (2009) DeMiguel and Xu (2009), and Adida and DeMiguel (2011) Adida and DeMiguel (2011). Additionally, in the operations research and operations management communities, a growing literature has been devoted to analyzing oligopoly models with congestion, e.g., Acemoglu and Ozdaglar (2007), and Johari et al. (2010). Recently, Correa et al. (2014) study markup equilibria, a particular case of supply function equilibria, with firms that have increasing marginal costs. In supply function equilibria firms are assumed to choose functions which map the quantity produced to prices, see Klemperer and Meyer (1989). In markup equilibria firms are restricted to choose a supply function of the form of a scalar times their marginal cost. Correa et al. (2014) find sufficient conditions for the existence of markup equilibria for marginal cost functions very similar to the ones were uniform co-payments are optimal in our model. On the other hand, the problem of controlling and reducing the contagion of infectious diseases has been studied in the operations management literature mainly focusing on the analysis of vaccine's markets, particularly the influenza vaccine, its supply chain coordination -e.g. Chick et al. (2008) and Mamani et al. (2012)- and the market competition under yield uncertainty -e.g. Deo and Corbett (2009) and Arifoglu et al. (2012)- as opposed to our interest in subsidy allocation. In particular, we consider the case of allocating subsidies to Cournot competitors under yield uncertainty, and we show that if the demand and the marginal costs functions are linear, then uniform co-payments are

optimal in this setting.

The motivation problem of allocating subsidies to increase the aggregated market consumption of new antimalarial drugs is also studied by Taylor and Xiao (2014), however they study a different question. Specifically, they consider the case of one manufacturer selling to one retailer facing stochastic demand at a fixed price c . Their analysis focuses on the placement of the subsidy by the central planner in the supply chain, comparing the possibility of subsidizing either sales or purchases (from the retailer point of view). They conclude that the central planner should only subsidize purchases, which is equivalent to subsidizing the manufacturer and thus consistent with our modeling framework. They show that this insight is maintained for the case of multiple heterogeneous retailers and one manufacturer. Furthermore, their Lemma 3 characterizes the order up to level of the retailers, which is decreasing in the wholesale price c . Therefore, this model can be characterized by an arbitrary decreasing inverse demand function from the manufacturer's point of view. In this chapter we focus directly on subsidies allocated to heterogeneous manufacturers which face an arbitrary decreasing inverse demand function, and we incorporate market competition among them. In this sense, our model is consistent with the insights provided by Taylor and Xiao (2014), and we extend the analysis to focus on the effectiveness of allocating the same co-payment to competing heterogeneous manufacturers. The combined message of these two papers to the policy makers is that, on the one hand, allocating co-payments to the manufacturers make sense when the objective is to maximize the aggregated market consumption, and on the other hand, even if the manufacturers are heterogeneous, the very simple and practical policy of allocating the same co-payments to each firm will most likely obtain most of the potential benefits. On the other hand, there is a growing trend in the operations management literature that studies the problem of a central planner deciding rebates that are directed to the *consumers*, with the goal of incentivizing the adoption of some technology, such as green technology, see for example Aydin and Porteus (2009), Lobel and Perakis (2012), Cohen et al. (2012), Chemama et al. (2013), Krass et al. (2013) Raz and Ovchinnikov (2013), and Cohen et al. (2013). In contrast, motivated by a

different set of practical applications, we focus on co-payments that are allocated to the producers, for each unit sold in the market. More generally, our work is related to the operations management literature that analyzes the impact of contract design on the behavior of firms in a supply chain. A comprehensive overview of this literature is provided in Cachon (2003). However, the focus of this framework is set on firms designing contracts to maximize their profits, while we are interested in a central planner designing incentives to maximize the aggregated market consumption of a good.

1.3 Model

In this section, we introduce a mathematical programming formulation of the subsidies allocation problem. We then use this formulation to obtain a relaxation of the problem, which provides an upper bound on the largest aggregated market consumption that can be induced with the available budget.

We consider a market for a commodity composed by $n \geq 2$ heterogeneous competing firms. Each firm $i \in \{1, \dots, n\}$ decides its output q_i independently, with the goal of maximizing its own profit. We assume that the introduction of subsidies in the market will induce an increase in the aggregated market consumption, and that the firms do not have the installed capacity to provide all of it. This implies that capacity is scarce in the market. We model this effect by assuming that the marginal cost of each firm is increasing. Specifically, *we assume that the firms have a firm-specific non-negative, increasing and differentiable marginal cost function on its production quantity, denoted by $h_i(q_i)$.*

Consumers are described by an inverse demand function $P(Q)$, where $Q \equiv \sum_{i=1}^n q_i$ is the aggregated market consumption. *We assume that $P(Q)$ is non-negative, decreasing and differentiable in $[0, \bar{Q}]$, where \bar{Q} is the smallest value such that $P(\bar{Q}) = 0$.* This is equivalent to assuming that the aggregated market demand for the good is bounded. This assumption captures the antimalarial drugs example from Section 1.1.1, where even if the new malaria treatment was given away for free there would

be a finite demand for it. Additionally, we assume that $P(0) = \bar{P} > 0$. This is equivalent to assuming that there exists a finite price such that the demand for the good becomes zero. This could be motivated by the consumers of the good switching to a substitute product, or simply not being able to afford it. In the antimalarial drugs example, there exist alternative treatments, which are less effective, that consumers may choose instead. Moreover, this is precisely the motivation for introducing a subsidy for the new malaria treatment in the first place.

The assumption on the market equilibrium dynamics is that each firm participating in the market equilibrium produces up to the point where its marginal cost equals the market price; and firms that do not participate in the market equilibrium must have a marginal cost of producing zero units which is larger than the market price. This can be expressed in the following condition:

$$\text{For each } i, j, \text{ if } q_i > 0, \text{ then } h_i(q_i) = P(Q) \leq h_j(q_j). \quad (1.1)$$

At this level of generality, in both the firms' marginal costs and the inverse demand function, an interpretation for this equilibrium condition is that firms act as price takers and compete on quantity. Note that this simple model captures the behavior of ACT manufacturers given in Section 1.1.1, where all the firms receiving co-payments operate in a no-profit/no loss basis. More generally, this is a reasonable model whenever the firms in the market have little market power, for example when there are many firms competing in the market, or when firms face threat of entry to their market, see Tirole (1988). However, we will show that for more specific families of marginal cost functions, or inverse demand functions, well known imperfect market competition models will be special cases of our model. These include Cournot Competition with linear demand, and Cournot Competition under yield uncertainty, with linear demand and linear marginal cost functions. Assuming a decreasing inverse demand function, and increasing firms' marginal cost functions, ensure that there exists a unique market equilibrium, see, for example, Marcotte and Patriksson (2006).

Note that the generality of an arbitrary decreasing inverse demand function allows

to model complex demand mechanisms that have been considered in the operations management literature. One such example is the case of multiple competing retailers under demand uncertainty. Specifically, Bernstein and Federgruen (2005) have shown that in a model where each retailer chooses its retail price p_i and its order quantity y_i , and faces multiplicative random demand, the distribution of which may depend on its own retail price as well as those of the other retailers, there exists a unique Nash equilibrium in which all the retailer prices decrease when the wholesale price is reduced. Moreover, under additional mild assumptions this leads to each equilibrium order quantity being decreasing in the wholesale price, resulting in a decreasing inverse demand function. This model may potentially capture in more details the mechanism by which the price to the final consumer is reduced when the central planner allocates co-payments to the manufacturers, however, as long as it induces a demand to the suppliers which is decreasing in the wholesale price, this model is a particular case of a general decreasing demand function, therefore our results apply.

In more details, in the model considered in Theorem 4 of Bernstein and Federgruen (2005) each retailer faces multiplicative random demand $D_i(\mathbf{p}) = d_i(\mathbf{p})\epsilon_i$, where ϵ_i is a random variable with cdf $G_i(\cdot)$, independent of the vector of prices \mathbf{p} . Their Theorem 4 shows that if $\log d_i(\mathbf{p})$ has increasing differences in (p_i, p_j) , for each i , and if the demand functions satisfy two technical assumptions (satisfied for example by $d_i(p)$ being Linear, Logit, Cobb-Douglas or CES, and ϵ_i being exponential, normal with mean one and standard deviation $\sigma \leq 1$, or having a power distribution, for each i) then there exists a unique Nash equilibrium, where all prices $p_i(w)$ are increasing in the wholesale price w . While the equilibrium order quantity is $y_i(\mathbf{p}(w), w) = d_i(\mathbf{p}(w))G_i^{-1}\left(\frac{p_i(w)-w}{p_i(w)}\right)$. Note that the additional assumption that the Jacobian of $d_i(\mathbf{p})$ is diagonally dominant (on p_i) is sufficient for $y_i(\mathbf{p}(w), w)$ to be decreasing in the wholesale price w , as long as the decrease in the retailer prices is smaller than the reduction in the wholesale price that induced them. The latter condition is the behavior we expect to see in practice, as the retailers would tend to keep a fraction of the benefits generated by the wholesale price reduction for themselves, as opposed to transferring all of it to the consumers. Finally, in Bernstein and

Federgruen (2005) all the retailers buy from one supplier, while in our model, they would buy at the wholesale price induced by multiple price taking suppliers engaging in quantity competition.

1.3.1 Co-payment Allocation Problem

We will refer to the problem faced by the central planner as the *co-payment allocation problem (CAP)*. The co-payment allocation problem is a particular case of a Stackelberg game, or a bilevel optimization problem. In the first stage, the central planner allocates a given budget $B > 0$, in the form of co-payments $y_i \geq 0$, to each firm $i \in \{1, \dots, n\}$, per each unit provided in the market. Moreover, she anticipates that, in the second stage, the equilibrium output of each firm will satisfy a modified version of the equilibrium condition. The difference in the market equilibrium condition is given by the fact that, from firm i 's perspective, the effective price, for each unit sold, is now $P(Q) + y_i$, or equivalently its marginal cost is reduced by y_i .

The central planner's objective is to maximize the aggregated market consumption. Note that this is equivalent to maximizing the consumer surplus, which in this model is equal to $\int_0^Q P(x)dx - P(Q)Q$. Specifically, the derivative on the consumer surplus with respect to the equilibrium aggregated market consumption is $-P'(Q)Q > 0$, which is positive for any decreasing inverse demand function. This is the appropriate objective in many applications, where the central planner is a supra-national authority, like the World Bank, whose main interest is effectively to maximize the aggregated market consumption, say of an infectious disease treatment, without explicitly taking into account the additional surplus obtained by local producers (see Arrow et al. (2004) for further discussion on this topic for the case of antimalarial drugs). Additionally, in Section 1.6, we will also analyze the case where the central planner's objective is to maximize the social welfare, including both the consumer and the producer surplus.

Finally, let us emphasize that the central planner can only allocate co-payments, and never charge a tax for the units produced in the market. In other words, the co-payments being allocated have to be non-negative. A formulation of the co-payment

allocation problem is given in the following:

$$\begin{aligned} \max_{\mathbf{y}, \mathbf{q}, Q} \quad & Q \\ \text{s.t.} \quad & \sum_{i=1}^n q_i y_i \leq B \end{aligned} \tag{1.2}$$

$$y_i \geq 0, \text{ for each } i \in \{1, \dots, n\} \tag{1.3}$$

$$\sum_{i=1}^n q_i = Q \tag{1.4}$$

$$q_i \geq 0, \text{ for each } i \in \{1, \dots, n\} \tag{1.5}$$

$$P(Q) + y_i = h_i(q_i), \text{ for each } i \in \{1, \dots, n\}. \tag{1.6}$$

This is a valid formulation even if there are firms that have a positive marginal cost of producing zero units, which prevents them from participating in the market equilibrium. Namely, if for some firm i we have $h_i(0) \geq P(Q)$, then we can just set $q_i = 0$ and $y_i = h_i(0) - P(Q) \geq 0$. This is without loss of generality, because setting $q_i = 0$ ensures that firm i does not have any impact in the budget constraint (1.2), and the non-negativity constraint on the co-payment y_i ensures that the market equilibrium condition is satisfied. In other words, constraint (1.6) does *not* imply that every firm has to participate in the market equilibrium.

From the equilibrium condition given in constraint (1.6), it follows that we can replace all the co-payment variables y_i by $h_i(q_i) - P(Q)$. Namely, we can reformulate the co-payment allocation problem as if the central planner was deciding the output of each firm, as long as there exist feasible co-payments that can sustain the outputs chosen as the market equilibrium. The feasibility of the co-payments will be given by both the budget constraint (1.2), and the non-negativity of the co-payments (1.3). We summarize this observation in the following proposition.

Proposition 1. *The co-payments allocation problem faced by the central planner can*

be formulated as follows:

$$\begin{aligned} \max_{\mathbf{q}, Q} \quad & Q \\ \text{s.t.} \quad & \sum_{j=1}^n q_j h_j(q_j) - P(Q)Q \leq B \end{aligned} \quad (1.7)$$

$$(CAP) \quad h_i(q_i) \geq P(Q), \text{ for each } i \in \{1, \dots, n\} \quad (1.8)$$

$$\sum_{j=1}^n q_j = Q \quad (1.9)$$

$$q_i \geq 0, \text{ for each } i \in \{1, \dots, n\}. \quad (1.10)$$

The co-payments that the central planner must allocate to induce outputs \mathbf{q} are,

$$y_i(\mathbf{q}) = h_i(q_i) - P(Q), \text{ for each } i.$$

Constraint (1.7) is equivalent to the budget constraint (1.2). Note that it has a budget balance interpretation, namely, the total cost in the market, minus the total revenue in the market, has to be less or equal than the budget introduced by the central planner. Constraint (1.8) is equivalent to the non-negativity of the co-payments (1.3).

1.3.2 Special Cases

Our model is fairly general. In particular, in this section, we discuss some well known imperfect competition models that are captured as special cases.

Cournot Competition with Linear Demand. The classical oligopoly model proposed by Cournot is defined in a very similar setting. The only difference is that, given all the other firms production levels, each firm sets its output q_i at a level such that it maximizes their profit Π_i , where

$$\Pi_i = P(Q)q_i - \int_0^{q_i} h_i(x_i)dx_i.$$

If we assume $P(Q)$ is decreasing and $h_i(q_i)$ are increasing, for each i , as well as $P'(Q) + q_i P''(Q) \leq 0$, then there exists a unique market equilibrium defined by the solution to the first order conditions of the firms' profit maximization problem, see Vives (2001). Namely, at equilibrium, each firm sets its output at a level such that,

$$\text{For each } i, \text{ if } q_i > 0 \text{ then } \frac{\partial \Pi_i}{\partial q_i} = 0, \text{ or equivalently, } P(Q) = h_i(q_i) - P'(Q)q_i. \quad (1.11)$$

In the equilibrium condition (1.11), the marginal cost must be equal to the marginal revenue, while in the equilibrium condition (1.1), the marginal cost must be equal to the market price. Moreover, the term $P'(Q)q_i$ is not independent for each firm.

Now, for the commonly assumed special case where the inverse demand function is linear, namely $P(Q) = a - bQ$, it follows that $P'(Q) = -b$. Define $\tilde{h}_i(q_i) \equiv h_i(q_i) + bq_i$, for each i , then we can rewrite the equilibrium condition as follows:

$$\text{For each } i, \text{ if } q_i > 0 \text{ then } P(Q) = \tilde{h}_i(q_i).$$

This equilibrium condition is a special case of condition (1.1), but written for a modified cost function $\tilde{h}_i(q_i)$.

Cournot Competition under Yield Uncertainty with Linear Demand and Linear Marginal Costs. We consider the Cournot Competition under yield uncertainty model used in Deo and Corbett (2009). We assume that each firm $i \in \{1, \dots, n\}$ decides its production target \bar{q}_i , while the actual output is uncertain and given by $q_i = \alpha_i \bar{q}_i$, where α_i is a random variable reflecting the random yield for firm i . We assume that the random variables α_i are identically and independently distributed for all firms, with $\mathbb{E}[\alpha_i] = \mu$, and $\text{Var}[\alpha_i] = \sigma^2$. Additionally, we assume a linear inverse demand function $P(Q) = a - bQ$, where $Q = \sum_{i=1}^n q_i$ is again the aggregated market consumption.

We consider two marginal costs: (i) $\bar{h}(\bar{q}_i)$ per unit of production target, and (ii) $h(q_i)$ per unit actually produced. The first cost is driven by the amount of raw

materials needed for production, while the second cost corresponds to the cost of packaging the actual output. Finally, we assume Cournot Competition among the firms. Namely, given the production target of all the other firms, each firm sets its production target \bar{q}_i to the level that maximizes its expected profit. We generalize the model used in Deo and Corbett (2009) in two ways. First, we consider heterogeneous firms while Deo and Corbett consider homogeneous firms. Second, Deo and Corbett assume a constant marginal cost function and a fixed cost to enter the market, while we assume more general marginal cost functions. Moreover, we extend the model to include a central planner allocating subsidies to the competing firms, anticipating the market reaction to the subsidy allocation, and facing a budget constraint. In order to do so, we assume that both marginal cost functions are linear. Namely, we assume that $\bar{h}(\bar{q}_i) = \bar{g}_i \bar{q}_i$, and $h(q_i) = g_i q_i$. Note that we consider heterogeneous firms, where some of them may be more efficient than the others, depending on the values of the firm specific parameters \bar{g}_i and g_i .

Let us start by considering the second stage problem. Assume that the central planner allocates a co-payment $y_i \geq 0$ to each firm $i \in \{1, \dots, n\}$. Each firm sets its production target \bar{q}_i to the level that maximizes its expected profit, given by

$$\begin{aligned} & \mathbb{E} \left[P(Q)q_i + y_i q_i - \int_0^{\bar{q}_i} \bar{h}(x_i) dx_i - \int_0^{q_i} h(x_i) dx_i \right] \\ &= \mathbb{E} \left[\left(a - b \sum_{i=1}^n \alpha_i \bar{q}_i \right) \alpha_i \bar{q}_i + y_i \alpha_i \bar{q}_i - \bar{g}_i \frac{\bar{q}_i^2}{2} - g_i \frac{\alpha_i^2 \bar{q}_i^2}{2} \right]. \end{aligned}$$

The expectation is taken with respect to the random variables α_i . This is a concave maximization problem in \bar{q}_i , therefore, the first order condition is sufficient for optimality. In order to write the first order condition in a compact form, define

$$\tilde{g}_i \equiv \frac{\bar{g}_i}{\mu} + \frac{\sigma^2 + \mu^2}{\mu} g_i + b\mu + 2b \frac{\sigma^2}{\mu}, \quad \tilde{h}_i(\bar{q}_i) \equiv \bar{g}_i \bar{q}_i, \quad \bar{Q} = \sum_{j=1}^n \bar{q}_j, \quad \bar{P}(\bar{Q}) = a - \mu b \bar{Q}.$$

Additionally, note that the expected market price has the following closed form ex-

pression,

$$\mathbb{E}[P(Q)] = a - \mu b \sum_{i=1}^n \bar{q}_i = a - \mu b \bar{Q} = \bar{P}(\bar{Q}).$$

Hence, we can write the first order condition of the firms' profit maximization problem as follows:

$$\bar{P}(\bar{Q}) = \tilde{h}_i(\bar{q}_i) - y_i. \quad (1.12)$$

In order to define the co-payment allocation problem in this setting, it remains to address how will the yield uncertainty be considered in the budget constraint. We consider two possible approaches that will lead to optimization problems with similar structure.

First, assume that the central planner would like to find a co-payment allocation, such that it satisfies the budget constraint in expectation, then we can write the budget constraint as follows:

$$\mathbb{E} \left[\sum_{i=1}^n q_i y_i \right] = \mu \sum_{i=1}^n \bar{q}_i y_i \leq B.$$

Alternatively, assume that the central planner takes a robust approach. Namely, she would like to satisfy the budget constraint in each possible yield uncertainty realization. We will assume, for simplicity, that the i.i.d. random yields for each firm have a bounded support, that is $\alpha_i \in [\underline{\alpha}, \bar{\alpha}]$, for each i . Then, we can write the budget constraint as follows:

$$\bar{\alpha} \sum_{i=1}^n \bar{q}_i y_i \leq B.$$

Finally, assuming that the budget constraint must be satisfied in expectation (the robust approach is analogous), we can use Equation (1.12) to write the central

planner's problem, like in Proposition 1, as follows:

$$\begin{aligned} \max_{\mathbf{y}, \bar{\mathbf{q}}} \quad & \mathbb{E}[Q] = \mu \sum_{i=1}^n \bar{q}_i = \mu \bar{Q} \\ \text{s.t.} \quad & \sum_{i=1}^n \bar{q}_i \tilde{h}_i(\bar{q}_i) - \bar{P}(\bar{Q}) \bar{Q} \leq \frac{B}{\mu} \end{aligned} \quad (1.13)$$

$$\tilde{h}_i(\bar{q}_i) \geq \bar{P}(\bar{Q}), \text{ for each } i \in \{1, \dots, n\} \quad (1.14)$$

$$\sum_{j=1}^n \bar{q}_j = \bar{Q} \quad (1.15)$$

$$\bar{q}_i \geq 0, \text{ for each } i \in \{1, \dots, n\}. \quad (1.16)$$

The co-payments that the central planner must allocate to induce the production targets $\bar{\mathbf{q}}$, are $y_i = \tilde{h}_i(\bar{q}_i) - \bar{P}(\bar{Q})$, for each i . The resulting problem formulation is a special case of the co-payment allocation problem (*CAP*).

1.3.3 An Upper Bound Problem

Note that under our assumptions, the co-payment allocation problem (*CAP*) is not necessarily a convex optimization problem. In fact, we have only assumed that the marginal cost functions $h_i(q_i)$ are increasing, for each i , and that the inverse demand function $P(Q)$ is decreasing. In order to gain some insights into the structure of the optimal solution, we ignore the non-negativity of the co-payments and analyze the following relaxation, which provides an upper bound on the aggregated market consumption that can be induced with the available budget B .

$$\begin{aligned} \max_{\mathbf{q}, Q} \quad & Q \\ \text{s.t.} \quad & \sum_{j=1}^n q_j h_j(q_j) - P(Q) Q \leq B \end{aligned} \quad (1.17)$$

$$(UBP) \quad \sum_{j=1}^n q_j = Q \quad (1.18)$$

$$q_i \geq 0, \text{ for each } i \in \{1, \dots, n\}. \quad (1.19)$$

This upper bound problem may still be non-convex, because of the budget constraint (1.17). However, Lemma 1 below asserts that at optimality the budget constraint is tight (i.e. holds with equality), and each active firm i must have a value of $(h_i(q_i)q_i)'$ equal to each other, and no larger than any inactive firm. This property will be have a central role in proving the optimality of uniform subsidies.

Lemma 1. *Assume that the marginal cost functions $h_i(q_i)$ are non-negative, increasing, and differentiable in $[0, \bar{Q}]$; and that the inverse demand function $P(Q)$ is non-negative, decreasing, and differentiable in $[0, \bar{Q}]$. Then, any optimal solution to the upper bound problem (UBP) must satisfy the budget constraint (1.17) with equality, and also satisfy the following condition:*

$$\text{If } q_i > 0, \text{ then } (h_i(q_i)q_i)' \leq (h_j(q_j)q_j)', \text{ for each } i, j \in \{1, \dots, n\}.$$

Proof. The feasible set of problem (UBP) is closed and bounded. It is bounded because $q_i \in [0, \check{q}_i]$, for each i , where \check{q}_i is such that $h_i(\check{q}_i)\check{q}_i = B$. Similarly, $Q \in [0, \check{Q}]$, where $\check{Q} = \max_{i \in \{1, \dots, n\}} \{\check{q}_i\}$. On the other hand, it is closed because it is defined by inequalities on continuous functions. Additionally, the objective function of problem (UBP) is continuous. It follows that there exists an optimal solution to problem (UBP).

Let (\mathbf{q}^*, Q^*) be an optimal solution to problem (UBP). Assume by contradiction that the budget constraint is not tight for (\mathbf{q}^*, Q^*) . Namely,

$$\sum_{j=1}^n q_j^* h_j(q_j^*) - P(Q^*)Q^* = \sum_{j=1}^n q_j^* (h_j(q_j^*) - P(Q^*)) < B.$$

Then, we can increase the value of q_i^* , for any index i , by $\epsilon > 0$ sufficiently small, maintain feasibility, and obtain a strictly larger objective value. This contradicts the optimality of (\mathbf{q}^*, Q^*) . Therefore, the budget constraint must be tight *for every optimal solution* to problem (UBP).

Assume by contradiction that there exist indexes i, j such that $q_i^* > 0$ and $(h_i(q_i^*)q_i^*)' > (h_j(q_j^*)q_j^*)'$. Then, we can decrease the value of q_i^* , and increase the

value of q_j^* , both by the same $\epsilon > 0$ sufficiently small, and maintain feasibility. Specifically, the marginal change in the left hand side of the budget constraint (1.17) is $-(h_i(q_i^*)q_i^*)' + (h_j(q_j^*)q_j^*)' < 0$. Therefore, the budget constraint for this modified solution is satisfied, and not tight. However, this modified solution attains the same objective value Q^* , and it is therefore optimal. This is a contradiction to the fact that the budget constraint must be tight for every optimal solution to problem (UBP). ■

1.4 Optimality of Uniform Co-payments

The result obtained in this section asserts that uniform co-payments are optimal for the co-payment allocation problem (CAP), for a large class of marginal costs functions $h_i(q_i)$. Specifically, we show that if the marginal cost functions satisfy Property 1 below, then uniform subsidies are optimal.

Property 1. *For each i, j and each $q_i, q_j \geq 0$, if $h_i(q_i) > h_j(q_j)$ then $(h_i(q_i)q_i)' \neq (h_j(q_j)q_j)'$.*

Next, we show that there exists a large class of marginal cost functions that satisfy Property 1 above. Consider the case in which $h_i(q_i) = h(g_i q_i)$, where $h(x)$ is non-negative, increasing, and differentiable over $x \geq 0$, and $g_i > 0$ is a firm specific parameter. This captures the setting where all firms use a similar technology, but can differ in their efficiency. Specifically, $h(x)$ models the industry specific marginal cost function, while $g_i > 0$ models the efficiency of firm i .

In this setting there is no loss of generality in assuming $h(0) = 0$. Specifically, any positive value for $h(0)$ will affect each firm in the same way, therefore, it will only shift the market price by a constant that can be re-scaled to zero. This assumption implies that all firms have a positive output in the market equilibrium, for any positive market price. Therefore, the underlying assumption is that all firms have already entered the market *before* the subsidy is decided, and there is no subsequent entry or exit of firms into the market. This assumption is reasonable in our setting, where the subsidy is not permanent (it only applies until the budget is exhausted), and it is paid ex-post to the firms, for each unit already sold.

In this setting, any function $h(x)$ such that $h(x) + h'(x)x$ is monotone will satisfy Property 1. Specifically, for each such function we would have that $h_i(q_i) > h_j(q_j)$ is equivalent, by definition, to $h(g_i q_i) > h(g_j q_j)$. However, $h(x)$ increasing implies $g_i q_i > g_j q_j$. Moreover, $h(x) + h'(x)x$ monotone implies $h(g_i q_i) + h'(g_i q_i)g_i q_i \neq h(g_j q_j) + h'(g_j q_j)g_j q_j$. Which is, again by definition, equivalent to $(h_i(q_i)q_i)' \neq (h_j(q_j)q_j)'$. Some functions that satisfy this condition, and the respective marginal cost functions associated to them, are:

- $h(x) = e^x - 1$, $h_i(q_i) = e^{g_i q_i} - 1$.
- $h(x) = x^u$, for $u > 0$, $h_i(q_i) = g_i q_i^u$.
- $h(x) = \ln(x + 1)$, $h_i(q_i) = \ln(g_i q_i + 1)$.
- Any polynomial with positive coefficients.

Specifically, all these functions have the property that $h(x)x$ is convex over $x \geq 0$, therefore, $h(x) + h'(x)x$ is increasing. Note that the marginal cost functions $h(x)$ are allowed to be concave, e.g., $h(x) = x^u$ for $0 < u < 1$, and $h(x) = \ln(x + 1)$. Moreover, note that $h_i(q_i) = g_i q_i^u$ corresponds exactly to the only homogeneous function of degree $u > 0$ in one variable.

1.4.1 Sufficient Condition for Optimality

The next one is the main result in this section.

Theorem 1. *Assume that the marginal cost functions $h_i(q_i)$ are non-negative, increasing, and differentiable in $[0, \bar{Q})$; the inverse demand function $P(Q)$ is non-negative, decreasing, and differentiable in $[0, \bar{Q}]$. If the marginal cost functions satisfy Property 1, then uniform co-payments are optimal for the co-payment allocation problem (CAP).*

Proof. The existence of an optimal solution to problem (UBP) was shown in Lemma 1. Let (\mathbf{q}, Q) be an optimal solution to problem (UBP). We will show that if the marginal cost functions satisfy Property 1, then (\mathbf{q}, Q) induces uniform co-payments

for every firm with a positive output $q_i > 0$. Moreover, (\mathbf{q}, Q) is feasible for the co-payment allocation problem (*CAP*), therefore optimal.

From Lemma 1 it follows that (\mathbf{q}, Q) is such that the budget constraint is binding, and for each i, j with $q_i > 0$ and $q_j > 0$, we must have $(h_i(q_i)q_i)' = (h_j(q_j)q_j)'$. The assumption that the marginal cost functions satisfy Property 1 implies $h_i(q_i) = h_j(q_j)$, which implies that uniform subsidies are optimal. Specifically, because the budget constraint is tight, it follows that $h_i(q_i) - P(Q) = \frac{B}{Q} > 0$ for every i such that $q_i > 0$.

In order to show that the solution (\mathbf{q}, Q) is feasible for the co-payment allocation problem (*CAP*), it remains to show that the firms that do not participate in the market equilibrium effectively have a marginal cost of producing zero units which is larger than the induced market price. Specifically, (\mathbf{q}, Q) is such that, for each i, j with $q_i > 0$ and $q_j = 0$, we have,

$$h_j(0) - P(Q) \geq h_i(q_i) + h_i'(q_i)q_i - P(Q) \geq h_i(q_i) - P(Q) = \frac{B}{Q} > 0.$$

The first inequality follows from Lemma 1, and the second inequality follows from $h_i(q_i)$ increasing. The equality follows from the fact that the budget constraint is tight, and $q_i > 0$.

Hence, (\mathbf{q}, Q) is also feasible for the co-payment allocation problem (*CAP*), therefore optimal. Moreover, (\mathbf{q}, Q) induces uniform co-payments. Therefore, uniform co-payments are optimal for the co-payment allocation problem (*CAP*). ■

This result is surprising, considering that the assumptions on the inverse demand function are very general, and particularly since firms can be heterogeneous and the central planner has the freedom to allocate differentiated co-payments to each firm.

The intuition behind this result comes from the market equilibrium condition and the budget constraint. Essentially, if the central planner allocates a larger co-payment to a firm, then its resulting market share will increase, which is exactly the rate at which it will consume budget. This will in turn make less budget available to the rest of the firms, therefore, their co-payments would have to decrease. Theorem 1 shows that if the marginal cost functions of the firms satisfy Property 1, then the net effect

of this change will never be positive.

In particular, Theorem 1 applies for the special cases we considered in Section 1.3.2. For Cournot Competition with linear demand, uniform co-payments are optimal for any marginal cost functions $h_i(q_i)$, such that the functions $\tilde{h}_i(q_i) = h_i(q_i) + bq_i$ satisfy Property 1. Specifically, if the marginal cost functions are linear, that is $h_i(q_i) = g_i q_i$, for each i , then uniform co-payments are optimal. Similarly, for Cournot Competition under yield uncertainty with linear demand, if both marginal costs are linear, then uniform subsidies are optimal. Note that in both cases we allow for heterogeneous firms, where some of them can be significantly more efficient than others.

1.5 Incorporating Market State Uncertainty

A natural extension of the model discussed in Section 1.3 is to consider the setting where the central planner does not know the market state (defined by the inverse demand function) with certainty, but generally she will have a set of possible market state scenarios, and beliefs on the likelihood that each scenario will materialize.

In more details, we assume that she has a discrete description of the market state uncertainty, where each scenario $s \in \{1, \dots, m\}$ is realized with probability p_s . Each scenario s is characterized by a scenario dependent inverse demand function $P^s(Q^s)$. For each scenario $s \in \{1, \dots, m\}$, we make assumptions like in Section 1.3. Namely, we assume that each inverse demand function $P^s(Q)$ is non-negative, decreasing and differentiable in $[0, \bar{Q}^s]$, where \bar{Q}^s is the smallest value such that $P^s(\bar{Q}^s) = 0$. Similarly, for the market equilibrium condition we assume that if scenario s realizes, then firms set their output q_i^s at a level such that, for each i, j , if $q_i^s > 0$, then $h_i(q_i^s) = P^s(Q^s) \leq h_j(q_j^s)$.

Similar to Section 1.3, a formulation of the co-payments allocation problem under

market state uncertainty can be written as follows:

$$\begin{aligned} & \max_{(\mathbf{q}^s, Q^s)_{s \in \{1, \dots, m\}}, \mathbf{y}} && \sum_{s=1}^m Q^s p_s \\ & \text{s.t.} && \sum_{j=1}^n q_j^s y_j \leq B, \text{ for each } s \in \{1, \dots, m\} \end{aligned} \quad (1.20)$$

$$y_i \geq 0, \text{ for each } i \in \{1, \dots, n\} \quad (1.21)$$

$$\sum_{j=1}^n q_j^s = Q^s, \text{ for each } s \in \{1, \dots, m\} \quad (1.22)$$

$$q_i^s \geq 0, \text{ for each } i \in \{1, \dots, n\}, s \in \{1, \dots, m\} \quad (1.23)$$

$$h_i(q_i^s) - P^s(Q^s) - y_i \geq 0, \text{ for each } \begin{matrix} i \in \{1, \dots, n\} \\ s \in \{1, \dots, m\} \end{matrix} \quad (1.24)$$

$$q_i^s (h_i(q_i^s) - P^s(Q^s) - y_i) = 0, \text{ for each } \begin{matrix} i \in \{1, \dots, n\} \\ s \in \{1, \dots, m\} \end{matrix}. \quad (1.25)$$

The objective is to maximize the expected aggregated market consumption. Constraint (1.20) is the budget constraint, for each market state scenario. Constraint (1.21) corresponds to the non-negativity of the co-payments. Constraint (1.22) defines the aggregated market consumption for each scenario. Finally, constraints (1.23)-(1.25) are the complementarity constraints, which tie together the different scenarios. They state that, in each scenario, each firm either participates in the market equilibrium, in which case it produces the quantity that equates its marginal cost with the market price plus the co-payment; or its marginal cost of producing zero units is strictly larger than the market equilibrium price plus the co-payment, in which case the firm is inactive. Naturally, each firm must get the same co-payment in each possible scenario.

In other words, constraints (1.23)-(1.25) correspond to the non-anticipativity constraints, and they state that co-payments are a first stage decision made by the central planner *before* the uncertainty is realized. This is precisely what prevents us from using the co-payments to eliminate the complementarity constraints from the model formulation, similarly to Proposition 1. This makes the problem significantly harder to analyze. In order to somewhat simplify this formulation, we make the additional assumption that producing zero units has a marginal cost of zero, as stated in the

following Proposition.

Proposition 2. *If we additionally assume*

- $h_i(0) = 0$, for each $i \in \{1, \dots, n\}$.

Then, the co-payments allocation problem under market state uncertainty faced by the central planner can be re-written as follows:

$$\begin{aligned} \max_{(\mathbf{q}^s, Q^s)_{s=1, \dots, m}} \quad & \sum_{s=1}^m Q^s p_s \\ \text{s.t.} \quad & \sum_{j=1}^n q_j^s h_j(q_j^s) - P^s(Q^s) Q^s \leq B, \quad s \in \{1, \dots, m\} \end{aligned} \quad (1.26)$$

$$h_i(q_i^s) \geq P^s(Q^s), \text{ for each } \begin{matrix} i \in \{1, \dots, n\} \\ s \in \{1, \dots, m\} \end{matrix} \quad (1.27)$$

$$\text{(SCAP)} \quad \sum_{j=1}^n q_j^s = Q^s, \text{ for each } s \in \{1, \dots, m\} \quad (1.28)$$

$$q_i^s \geq 0, \text{ for each } i \in \{1, \dots, n\}, \quad s \in \{1, \dots, m\} \quad (1.29)$$

$$h_i(q_i^s) - P^s(Q^s) = h_i(q_i^{s'}) - P^{s'}(Q^{s'}), \quad \begin{matrix} i \in \{1, \dots, n\}, \\ s, s' \in \{1, \dots, m\}. \end{matrix} \quad (1.30)$$

The co-payments that the central planner must allocate in order to induce outputs $\{\mathbf{q}^s\}_{s \in \{1, \dots, m\}}$ are,

$$y_i = h_i(q_i^s) - P^s(Q^s), \text{ for each } i \in \{1, \dots, n\}, \quad s \in \{1, \dots, m\}. \quad (1.31)$$

Proposition 2 states that, if the marginal cost of producing zero units is zero, then every firm will participate in the market equilibrium, for any non-negative market price. Therefore, Equation (1.31) holds, and we can eliminate the variables y_i from the problem formulation.

Like in Proposition 1, constraint (1.26) corresponds to the budget constraint. Namely, for each scenario s , the total cost minus the total revenue in the market has to be less or equal than the budget introduced by the central planner. Constraint (1.27) is the non-negativity of the co-payments, it ensures that the solution proposed by the central planner can be sustained as a market equilibrium by allocating only subsidies, and not taxes. Like before, the only constraint that ties all the scenarios

together is the non-anticipativity constraint (1.30), which states that each firm must get the same co-payment in each possible scenario.

This problem is still hard to analyze directly, which motivates us to develop a relaxation that provides an upper bound on the expected aggregated market consumption that can be induced with the available budget, as shown below. All the proofs in this section are presented in Appendix A.1.

1.5.1 An Upper Bound Problem

We start with a simple observation that is derived from the structure of problem (*SCAP*).

Lemma 2. *For any feasible solution to the co-payments allocation problem under market state uncertainty (*SCAP*), without loss of generality, the scenarios can be renumbered, such that the following inequalities are true:*

$$P^1(Q^1) \geq P^2(Q^2) \geq \dots \geq P^m(Q^m), \quad (1.32)$$

$$h_i(q_i^1) \geq h_i(q_i^2) \geq \dots \geq h_i(q_i^m), \text{ for each } i, \quad (1.33)$$

$$q_i^1 \geq q_i^2 \geq \dots \geq q_i^m, \text{ for each } i, \quad (1.34)$$

$$Q^1 \geq Q^2 \geq \dots \geq Q^m, \quad (1.35)$$

$$\sum_{j=1}^n q_j^1 y_j \geq \sum_{j=1}^n q_j^2 y_j \geq \dots \geq \sum_{j=1}^n q_j^m y_j, \quad (1.36)$$

where $\sum_{j=1}^n q_j^s y_j$ is the total amount spent in co-payments in scenario s .

Let $(\mathbf{q}^s, Q^s)_{s=1, \dots, m}^*$ be an optimal solution to the co-payment allocation problem under scenario uncertainty (*SCAP*), and assume that the scenarios are numbered such that Equations (1.32)-(1.36) above hold. Then, we claim that the solution to problem (*SUBP*) below provides an upper bound on the expected aggregated market consumption that can be induced with the available budget. Specifically, problem (*SUBP*) is derived from problem (*SCAP*) by adding constraints (1.41) and (1.42) below, and replacing the non-anticipativity constraint (1.30), with the relaxed version (1.43).

$$\begin{aligned} \max_{\mathbf{q}^s, Q^s} \quad & \sum_{s=1}^m Q^s p_s \\ \text{s.t.} \quad & \sum_{j=1}^n q_j^s h_j(q_j^s) - P^s(Q^s)Q^s \leq B, \quad s \in \{1, \dots, m\} \end{aligned} \quad (1.37)$$

$$h_i(q_i^s) \geq P^s(Q^s), \quad \text{for each } \begin{matrix} i \in \{1, \dots, n\} \\ s \in \{1, \dots, m\} \end{matrix} \quad (1.38)$$

$$(SUBP) \quad \sum_{j=1}^n q_j^s = Q^s, \quad \text{for each } s \in \{1, \dots, m\} \quad (1.39)$$

$$q_i^s \geq 0, \quad \text{for each } i \in \{1, \dots, n\}, \quad s \in \{1, \dots, m\} \quad (1.40)$$

$$P^1(Q^1) \geq P^s(Q^s), \quad \text{for each } s \in \{1, \dots, m\} \quad (1.41)$$

$$Q^1 \geq Q^s, \quad \text{for each } s \in \{1, \dots, m\} \quad (1.42)$$

$$h_i(q_i^s) - P^s(Q^s) \leq h_i(q_i^1) - P^1(Q^1), \quad \text{for each } \begin{matrix} i \in \{1, \dots, n\}, \\ s \in \{1, \dots, m\}. \end{matrix} \quad (1.43)$$

Problem (SUBP) is a valid relaxation of problem (SCAP). Specifically, the optimal solution of problem (SCAP), $(\mathbf{q}^s, Q^s)_{s=1, \dots, m}^*$, is feasible for problem (SUBP), and attains the same objective value. To argue the feasibility of solution $(\mathbf{q}^s, Q^s)_{s=1, \dots, m}^*$ for problem (SUBP), recall from Lemma 2 that $s = 1$ is the scenario that attains the largest value for both the induced market price (see (1.32)), and the induced aggregated market consumption (see (1.35)), in solution $(\mathbf{q}^s, Q^s)_{s=1, \dots, m}^*$. It follows that adding constraints (1.41) and (1.42), does not cut-off solution $(\mathbf{q}^s, Q^s)_{s=1, \dots, m}^*$. Finally, solution $(\mathbf{q}^s, Q^s)_{s=1, \dots, m}^*$ satisfies constraint (1.43) with equality.

1.5.2 Optimality of Uniform Co-payments

In this section, we consider again the setting where all firms use a similar technology, but they can differ in their respective efficiency, similar to the assumptions in Section 1.4. Specifically, we consider the case in which $h_i(q_i) = h(g_i q_i)$, where $h(x)$ is non-negative, increasing, and differentiable over $x > 0$, and $g_i > 0$ is a firm specific parameter. The function $h(x)$ models the industry specific marginal cost function, while g_i models the efficiency of firm i . Recall from Section 1.4 that we can assume, without loss of generality, that $h(0) = 0$, therefore, we will refer to the co-payment

allocation problem under market state uncertainty (*SCAP*), and its upper bound problem (*SUBP*).

We now present sufficient conditions, which ensure that uniform subsidies maximize the expected aggregated market consumption in this setting. Specifically, we show that if the firms' marginal cost functions satisfy Property 2 below, then uniform subsidies are optimal for the co-payments allocation problem under market state uncertainty (*SCAP*).

Property 2. *The function $h(x)$ is convex, and such that for any $x_1 > x_2 \geq 0$, and $x_1 > x_3 > x_4 \geq 0$, if $\frac{h(x_2)}{h(x_1)} > \frac{h(x_4)}{h(x_3)}$ then $\frac{h'(x_2)}{h'(x_1)} > \frac{h'(x_4)}{h'(x_3)}$.*

Note that Property 2 implies Property 1, discussed in Section 1.4. Specifically, $h(x)$ increasing and convex implies that $h(x) + h'(x)x$ is increasing. This is a sufficient condition for Property 1 to hold.

Remark 1. *The functions $h_i(q_i) = g_i q_i^m$, for $m \geq 1$, and $h_i(q_i) = e^{g_i q_i} - 1$, satisfy Property 2.*

Note, from Remark 1, that from the examples of marginal cost functions that satisfy Property 1 given in Section 1.4, all the ones that are also *convex* satisfy Property 2 as well. In this sense, the extra requirements in Property 2, with respect to Property 1, are mainly driven by the convexity assumption. Finally, note that functions $h_i(q_i) = g_i q_i^m$, for $m \geq 1$, are the unique convex homogeneous functions in one variable.

Theorem 2 below shows that there exists an optimal solution to the upper bound problem (*SUBP*), such that the co-payments induced in scenario $s = 1$, the one that attains the largest aggregated market consumption (see (1.35)), and the largest amount spent in co-payments (see (1.36)), are uniform. This result will play a central role in proving the main result in this section.

Theorem 2. *Assume that the inverse demand function $P(Q)$ is non-negative, decreasing, and differentiable in $[0, \bar{Q}]$. Assume that the marginal costs functions are given by $h_i(q_i) = h(g_i q_i)$ for each i , for any increasing and continuously differentiable*

function $h(x)$, such that $h(0) = 0$. If $h(x)$ satisfies Property 2, then there exists an optimal solution to the upper bound problem (SUBP), $(\hat{q}^s, \hat{Q}^s)_{s=1,\dots,m}$, such that, $h_i(\hat{q}_i^1) - P^1(\hat{Q}^1) = y^1$ for each $i \in \{1, \dots, n\}$, for some value $y^1 > 0$.

To prove Theorem 2, we show the following lemmas that will be useful in the analysis. Specifically, we will consider the optimal solution to problem (SUBP) with the smallest difference between $(\max_{i \in \{1, \dots, n\}} \{h_i(q_i^1)\})$ and $(\min_{i \in \{1, \dots, n\}} \{h_i(q_i^1)\})$ (from Lemma 3 below we know that such a solution does exist). Note that proving Theorem 2 is equivalent to showing that this difference is zero. We will assume by contradiction that this difference is strictly positive, and show that then we can construct another optimal solution with an even smaller difference, a contradiction.

When constructing the modified optimal solution, Lemma 4 allows us to focus only on constraint (1.43). On the other hand, using the convexity assumption on $h(x)$, Lemma 5 provides bounds on the impact that the modifications to the optimal solution have on constraint (1.43). These bounds will allow us to complete the proof by arguing that the modified solution is feasible and optimal, while attaining a smaller difference between the maximum marginal cost in scenario $s = 1$, and the minimum marginal cost in scenario $s = 1$.

Lemma 3. *Under the assumptions of Theorem 2, there exists an optimal solution to problem (SUBP) that attains the minimum of the gaps between the maximum marginal cost in scenario $s = 1$, and the minimum marginal cost in scenario $s = 1$, induced by any optimal solution.*

Lemma 4. *Under the assumptions of Theorem 2, for any feasible solution to problem (SUBP), $(q^s, Q^s)_{s=1,\dots,m}$, if $h_i(q_i^s) > h_j(q_j^s)$ for some i, j, s , then we can transfer a sufficiently small $\epsilon^s > 0$, from q_i^s to q_j^s , without violating any of the constraints (1.37)-(1.42) related to scenario s .*

Lemma 5. *Under the assumptions of Theorem 2, for any feasible solution to problem (SUBP), $(q^s, Q^s)_{s=1,\dots,m}$, for any $\epsilon^1 > 0$, and for any scenario $s \neq 1$, the following*

conditions must hold:

$$\text{If } \epsilon^s \geq 0 \text{ satisfies } \epsilon^1 \leq \frac{h'(g_i(q_i^s - \epsilon^s))}{h'(g_i q_i^1)} \epsilon^s, \text{ then } h_i(q_i^s - \epsilon^s) - P^s(Q^s) \leq h_i(q_i^1 - \epsilon^1) - P^1(Q^1). \quad (1.44)$$

$$\text{If } \epsilon^s \geq 0 \text{ satisfies } \frac{h'(g_i(q_i^s + \epsilon^s))}{h'(g_i q_i^1)} \epsilon^s = \epsilon^1, \text{ then } h_i(q_i^s + \epsilon^s) - P^s(Q^s) \leq h_i(q_i^1 + \epsilon^1) - P^1(Q^1). \quad (1.45)$$

Theorem 3 below concludes this section characterizing a family of firms' marginal cost functions such that uniform co-payments are optimal, even if the central planner is uncertain about the market state. This family includes convex homogeneous functions of the same degree.

Theorem 3. *Assume that the inverse demand function $P(Q)$ is non-negative, decreasing, and differentiable in $[0, \bar{Q}]$. Assume that the marginal costs functions are given by $h_i(q_i) = h(g_i q_i)$ for each i , for any increasing and continuously differentiable function $h(x)$, such that $h(0) = 0$. If $h(x)$ satisfies Property 2, then allocating the largest feasible uniform co-payment is an optimal solution for the co-payment allocation problem under market state uncertainty (SCAP).*

This result is surprising, as it shows that, with some additional conditions, the optimality of uniform subsidies is preserved, even if the central planner is uncertain about the market state. Different market states induce different inverse demand functions, which can be arbitrarily different. Moreover, the assumption on the inverse demand functions of each scenario are very mild. Specifically, we only assume that they are decreasing. This is a very relevant setup, as it corresponds to a more realistic representation of the problem faced in practice, where there are large uncertainties about different characteristics of the market state, which ultimately define the effective response of the demand side to different market prices.

Moreover, the analysis suggests that the central planner only needs to consider the scenario with the highest aggregated market consumption at equilibrium (see (1.35)), i.e., scenario $s = 1$, regardless of the exact distribution over the different market states. Specifically, Theorem 2 shows that uniform subsidies are optimal for

scenario $s = 1$ in the relaxed upper bound problem (*SUBP*), while Theorem 3 shows that the uniform subsidies induced by scenario $s = 1$, are in fact optimal for the co-payment allocation problem under market state uncertainty (*SCAP*). This insight suggests that the central planner only needs to identify the scenario with the highest aggregated market consumption at equilibrium, and implement the uniform subsidies induced by it, as opposed to taking into consideration her beliefs on the likelihood that each market state will be realized, and the effect that the subsidy allocation will have on each possible market state scenario.

1.6 Maximizing Social Welfare

In this section, we assume that the central planner's objective is in fact to maximize social welfare. Given some $\delta \in (0, 1]$, which represents the social cost of funds, the central planner problem of allocating subsidies to maximize social welfare can be written as follows:

$$\begin{aligned} \max_{\mathbf{y}, \mathbf{q}, Q} \quad & \int_0^Q P(x)dx - \sum_{i=1}^n \int_0^{q_i} (h(x_i) - y_i)dx_i - \delta \left(\sum_{i=1}^n q_i y_i \right) \\ \text{s.t.} \quad & y_i \geq 0, \text{ for each } i \in \{1, \dots, n\} \end{aligned} \tag{1.46}$$

$$\sum_{i=1}^n q_i = Q \tag{1.47}$$

$$q_i \geq 0, \text{ for each } i \in \{1, \dots, n\} \tag{1.48}$$

$$P(Q) + y_i = h_i(q_i), \text{ for each } i \in \{1, \dots, n\}. \tag{1.49}$$

The first two terms in the objective function correspond to the sum of the consumer and producer surplus, including the co-payments y_i . The third term in the objective function corresponds to the social cost of financing the subsidies. Note that in this problem there is no budget constraint. Specifically, the social cost of funds $\delta \in (0, 1]$ will induce a total amount invested in subsidies at optimality, which can be interpreted as the implicit budget available. Constraint (1.46) states that the central planner is only allowed to allocate subsidies, and not taxes, to the firms. Like in Section 1.3,

constraint (1.49) does *not* imply that every firm has to participate in the market equilibrium.

From the equilibrium condition given in constraint (1.49), it follows again that we can replace all the co-payment variables y_i by $h_i(q_i) - P(Q)$, as stated in the proposition below.

Proposition 3. *The social welfare maximization problem can be written as follows:*

$$\begin{aligned} \max_{\mathbf{q}, Q} \quad & \int_0^Q P(x)dx - \sum_{i=1}^n \int_0^{q_i} h(x_i)dx_i + (1 - \delta) \left(\sum_{i=1}^n h(q_i)q_i - P(Q)Q \right) \\ \text{s.t.} \quad & h(q_i) \geq P(Q), \text{ for each } i \in \{1, \dots, n\} \end{aligned} \quad (1.50)$$

$$(CAP - SW) \quad \sum_{i=1}^n q_i = Q \quad (1.51)$$

$$q_i \geq 0, \text{ for each } i \in \{1, \dots, n\}. \quad (1.52)$$

The co-payments that the central planner must allocate to induce outputs \mathbf{q} are,

$$y_i(\mathbf{q}) = h_i(q_i) - P(Q), \text{ for each } i.$$

The first two terms in the objective function correspond to the sum of the consumer and producer surplus, with no subsidies. The third term corresponds to the increase in social welfare induced by subsidies, minus the social cost of financing them. Constraint (1.50) states that the central planner is only allowed to allocate subsidies, and not taxes, to the firms.

We will make the natural assumption that the social cost of funds, $\delta \in (0, 1]$, is such that objective function of problem (CAP - SW) is coercive¹, therefore, there exists an optimal solution, see, for example, Bertsekas (1999). Then, the budget B that the central planner spends in subsidies, in order to maximize social welfare, can be written as follows. Let (\mathbf{q}^*, Q^*) be an optimal solution of problem (CAP - SW), then

$$B \equiv \sum_{i=1}^n h(q_i^*)q_i^* - P(Q^*)Q^*.$$

¹Let us denote the objective function of problem (CAP - SW) by $SW(\mathbf{q}, Q)$. Recall that $SW(\mathbf{q}, Q)$ is coercive if $SW(\mathbf{q}^k, Q^k) \rightarrow -\infty$ for any feasible sequence such that $\|(\mathbf{q}^k, Q^k)\| \rightarrow \infty$.

1.6.1 Optimality of Uniform Co-payments

We conclude this section by characterizing settings where, in addition to maximizing the aggregated market consumption, uniform co-payments also maximize social welfare. Specifically, we show that if the marginal cost functions satisfy Property 3 below, then uniform subsidies are optimal for the problem (*CAP – SW*).

Property 3. For each i, j and each $q_i, q_j \geq 0$, if $h_i(q_i) > h_j(q_j)$ then $\frac{h_i(q_i)}{h'_i(q_i)q_i} \geq \frac{h_j(q_j)}{h'_j(q_j)q_j}$.

Two examples of marginal cost functions that satisfy Property 3, are

- $h_i(q_i) = g_i q_i^u$ for $u > 0$.
- $h_i(q_i) = \ln(g_i q_i + 1)$.

Note that these marginal cost functions also satisfy Property 1. Therefore, for these two marginal cost functions uniform subsidies maximize both the consumer surplus and social welfare.

This leads to the main result of this section.

Theorem 4. Assume that the marginal cost functions $h_i(q_i)$ are non-negative, increasing, and differentiable in $[0, \bar{Q}]$; the inverse demand function $P(Q)$ is non-negative, decreasing, and differentiable in $[0, \bar{Q}]$; and the social cost of funds $\delta \in (0, 1]$ is such that it induces a finite central planner's budget B . If the marginal cost functions satisfy Property 3, then uniform co-payments are optimal for the social welfare maximization problem (*CAP – SW*).

Proof. Let (\mathbf{q}^*, Q^*) be the optimal solution of problem (*CAP – SW*). First, we show that (\mathbf{q}^*, Q^*) must satisfy,

$$(1 - \delta) < \frac{h_i(q_i^*)}{h_i(q_i^*) + h'_i(q_i^*)q_i^*}, \text{ for each } i. \quad (1.53)$$

Specifically, the expression in the objective function of problem (*CAP – SW*) related

to the aggregated market consumption Q , is strictly increasing in Q . Namely,

$$\frac{\partial}{\partial q_i} \left(\int_0^Q P(x) dx - (1 - \delta)P(Q)Q \right) = \delta P(Q) - (1 - \delta)P'(Q)Q > 0,$$

where the inequality follows from the inverse demand function $P(Q)$ being non-negative and decreasing. On the other hand, the remaining expression in the objective function of problem $(CAP - SW)$, related to firm i 's output q_i , is such that,

$$\frac{\partial}{\partial q_i} \left((1 - \delta) \sum_{i=1}^n h(q_i)q_i - \sum_{i=1}^n \int_0^{q_i} h(x_i) dx_i \right) = (1 - \delta)(h_i(q_i) + h'_i(q_i)q_i) - h_i(q_i).$$

Assume for a contradiction that Equation (1.53) does not hold. Namely, there exists an index i such that, $(1 - \delta)(h_i(q_i^*) + h'_i(q_i^*)q_i^*) - h_i(q_i^*) \geq 0$. It follows that we can increase q_i^* by $\epsilon > 0$ sufficiently small, and obtain a feasible solution that attains a strictly larger objective value. This is a contradiction to the optimality of (\mathbf{q}^*, Q^*) .

Second, assume by contradiction that there exist indexes i, j , with $q_i^* > 0$ and $q_j^* > 0$, such that $h_i(q_i^*) > h_j(q_j^*)$. The fact that the marginal cost functions satisfy Property 3 implies that $\frac{h_i(q_i^*)}{h'_i(q_i^*)q_i^*} \geq \frac{h_j(q_j^*)}{h'_j(q_j^*)q_j^*}$. From direct algebraic manipulations, it follows that,

$$\frac{h_i(q_i^*)}{h_i(q_i^*) + h'_i(q_i^*)q_i^*} \leq \frac{h_i(q_i^*) - h_j(q_j^*)}{h_i(q_i^*) + h'_i(q_i^*)q_i^* - h_j(q_j^*) - h'_j(q_j^*)q_j^*}. \quad (1.54)$$

Now, Equations (1.53) and (1.54) imply that

$$(1 - \delta) < \frac{h_i(q_i^*) - h_j(q_j^*)}{h_i(q_i^*) + h'_i(q_i^*)q_i^* - h_j(q_j^*) - h'_j(q_j^*)q_j^*}.$$

Therefore, we can transfer $\epsilon > 0$ sufficiently small from q_i^* to q_j^* , and obtain the following positive marginal change in the objective function, $h_i(q_i^*) - h_j(q_j^*) - (1 - \delta)(h_i(q_i) + h'_i(q_i)q_i - h_j(q_j^*) - h'_j(q_j^*)q_j^*) > 0$. Namely, there exists a feasible solution with a strictly larger objective value. This contradicts the optimality of (\mathbf{q}^*, Q^*) .

Hence, we conclude that for each i, j , with $q_i^* > 0$ and $q_j^* > 0$, it must be the case that $h_i(q_i^*) = h_j(q_j^*)$. Therefore, uniform subsidies maximize social welfare. \blacksquare

1.7 Numerical Results

In Section 1.4, we have identified conditions on the firms' marginal cost functions that guarantee the optimality of uniform co-payments to maximize the aggregated market consumption of a good. In this section we study the performance of uniform co-payments, in relevant settings where they are sub-optimal. More precisely, in Section 1.7.1 we consider Cournot Competition with linear demand and constant marginal cost, as well as a more general setting with price taking firms having non-linear marginal costs and facing non-linear demand. On the other hand, in Section 1.7.2 we extend this study to consider Cournot competition with non-linear demand and non-linear marginal costs, which require additional modeling and machinery in order to be simulated. Our goal here is to study numerically the performance of uniform co-payments on problems with data generated at random.

1.7.1 Results for some Special Cases from Section 1.3.2

In order to evaluate the relative performance of uniform subsidies, we need to be able to compute the aggregated market consumption induced by them. Proposition 4 below addresses this issue.

Proposition 4. *Assume that the marginal cost functions $h_i(q_i)$ are non-negative, increasing, and differentiable in $[0, \bar{Q}]$; and the inverse demand function $P(Q)$ is non-negative, decreasing, and differentiable in $[0, \bar{Q}]$. Then, the market equilibrium induced by the largest feasible uniform co-payment can be computed as the solution to the following convex optimization problem,*

$$\begin{aligned}
 \min_{\mathbf{q}} \quad & \sum_{j=1}^n \int_0^{q_j} h_j(x_j) dx_j - \int_0^{q_{n+1}} P(\bar{Q} - x_{n+1}) dx_{n+1} - B \ln(\bar{Q} - q_{n+1}) \\
 \text{s.t.} \quad & \sum_{j=1}^n q_j + q_{n+1} = \bar{Q} \\
 \text{(UCAP)} \quad & q_i \geq 0, \text{ for each } i.
 \end{aligned}$$

Assuming that the inverse demand function $P(Q)$ is decreasing, and that the

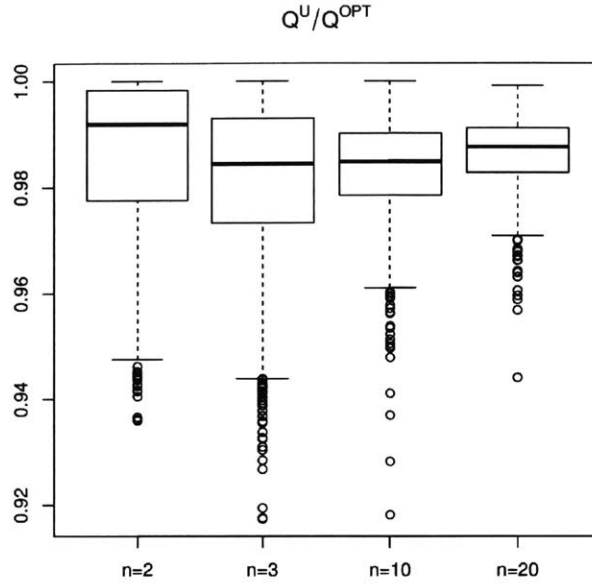


Figure 1-1: Boxplot of the Relative Performance of Uniform Co-payments - Cournot Competition with Constant Marginal Costs

Q^U/Q^{OPT}	n=2	n=3	n=10	n=20
Min.	0.9360	0.9175	0.9182	0.9442
1st Qu.	0.9776	0.9734	0.9785	0.9828
Median	0.9919	0.9845	0.9849	0.9876
Mean	0.9860	0.9806	0.9834	0.9866
3rd Qu.	0.9983	0.9930	0.9902	0.9912
Max.	1.0000	1.0000	1.0000	0.9991

Table 1.1: Summary Statistics of the Relative Performance of Uniform Co-payments - Cournot Competition with Constant Marginal Costs

firms' marginal cost functions $h_i(q_i)$ are increasing, implies that problem (*UCAP*) is a convex optimization problem. On the other hand, in the experimental settings we consider, it will always be the case that at least the upper bound problem (*UBP*) is a convex optimization problem. To solve these problems we used CVX, a package for specifying and solving convex programs, see Grant and Boyd (2012). We will denote by Q^U , Q^{OPT} and Q^{UB} the aggregated market consumption component of the optimal solutions to problems (*UCAP*), (*CAP*) and (*UBP*), respectively.

Cournot Competition with Linear Demand and Constant Marginal Costs

The model presented in Section 1.3 captures Cournot Competition with linear demand and non-decreasing marginal cost functions $h_i(q_i)$. Specifically, this implies that the modified marginal cost function defined in Section 1.3.2, $\tilde{h}_i(q_i) \equiv h_i(q_i) + bq_i$, is increasing. In particular, in this section we consider *constant* marginal costs. Although the constant marginal costs case moves away from the our scarce installed capacity assumption, it is a well understood model where uniform co-payments are not optimal. Therefore, it is interesting to study the performance of uniform co-payments in this setting.

Specifically, in this section we assume $P(Q) = a - bQ$, and $h_i(q_i) = c_i$, for each i . Therefore, the modified marginal cost is $\tilde{h}_i(q_i) = c_i + bq_i$, for each i . Under these assumptions, the co-payment allocation problem (*CAP*) is a convex optimization problem. Therefore, we solve both the uniform co-payments allocation problem (*UCAP*) and the co-payment allocation problem (*CAP*), and we compare their objective functions directly. We consider four cases in the number of firms participating in the market, $n \in \{2, 3, 10, 20\}$. For each one of this four cases, we solve 1,000 instances of the problem. These instances are randomly generated, with parameters sampled from the following distributions: a, b are uniformly distributed in $[0, 50]$, c_i are independent and uniformly distributed in $[0, a]$, for each i .

Figure 1-1 presents a boxplot of the results for the ratio Q^U/Q^{OPT} , while Table 1.1 presents some summary statistics. It is interesting that the minimum value of the ratio Q^U/Q^{OPT} never went below 91% in the simulation results. Moreover, the mean and median values are above 98%, for each value of the number of firms participating in the market n . This suggests that, in most cases, the aggregated market consumption induced by uniform co-payments is fairly close to the aggregated market consumption induced by the optimal co-payment allocation.

Price Taking Firms with Non-linear Demand and Marginal Costs

Now we consider a more general experimental setup, with non-linear demand and non-linear marginal costs, where the firms act as price takers. In this setting we assume

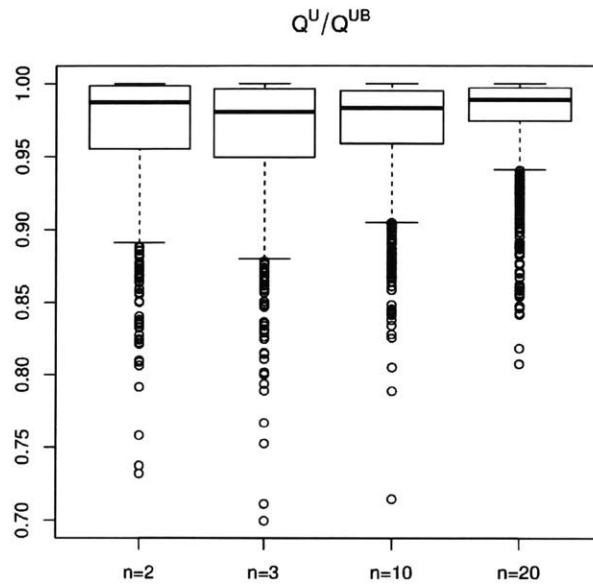


Figure 1-2: Boxplot of the Relative Performance of Uniform Co-payments - Price Taking Firms

Q^U/Q^{UB}	n=2	n=3	n=10	n=20
Min.	0.7321	0.7000	0.7149	0.8076
1st Qu.	0.9554	0.9497	0.9592	0.9747
Median	0.9874	0.9808	0.9836	0.9892
Mean	0.9698	0.9661	0.9710	0.9784
3rd Qu.	0.9985	0.9966	0.9952	0.9972
Max.	1.0000	1.0000	1.0000	1.0000

Table 1.2: Summary Statistics of the Relative Performance of Uniform Co-payments - Price Taking Firms

Q^U/Q^{OPT}	n=2	n=3	n=10	n=20
Min.	0.8360	0.8370	0.8577	0.8884
1st Qu.	0.9846	0.9798	0.9662	0.9681
Median	0.9952	0.9909	0.9791	0.9803
Mean	0.9878	0.9840	0.9735	0.9756
3rd Qu.	0.9991	0.9963	0.9976	0.9870
Max.	1.0000	1.0000	0.9996	0.9997

Table 1.3: Summary Statistics of the Relative Performance of Uniform Co-payments - Cournot Competition with Nonlinear Demand

$P(Q) = a - bQ^{m_0}$, and $h_i(q_i) = c_i + g_i q_i^{m_i}$ for each i . Under these assumptions, the co-payment allocation problem (*CAP*) is a non-convex optimization problem. However, the upper bound problem (*UBP*) is a convex optimization problem. Therefore, we solve both the uniform co-payments allocation problem (*UCAP*) and the upper bound problem (*UBP*), and we compare their objective functions.

We consider again four cases in the number of firms participating in the market, $n \in \{2, 3, 10, 20\}$. For each one of this four cases, we solve 1,000 instances of the problem. These instances are randomly generated, with parameters sampled from the following distributions: a, b are uniformly distributed in $[0, 50]$. For each i , c_i are independent and uniformly distributed in $[0, a]$, g_i are independent and uniformly distributed in $[0, 50]$, and m_i are independent and uniformly distributed in $(0, 20]$. Finally, m_0 is uniformly distributed in $(0, 3]$.

Note that $P(Q) = a - bQ^{m_0}$, $m_0 \in (0, 3]$, captures both convex and concave decreasing inverse demand functions. Similarly, $h_i(q_i) = c_i + g_i q_i^{m_i}$, $m_i \in (0, 20]$ for each i captures both convex and concave marginal cost firms. The results for the ratio Q^U/Q^{UB} are displayed in Figure 1-2 and in Table 1.2. The minimum value of the ratio Q^U/Q^{UB} never went below 70% in the simulation results. Moreover, the mean and median values are above 96%, for each value of the number of firms participating in the market n , where in this case we are not comparing directly to the optimal solution, but to an upper bound. This suggests that again, in most cases, the aggregated market consumption induced by uniform co-payments is fairly close to the aggregated market consumption induced by the optimal co-payment allocation.

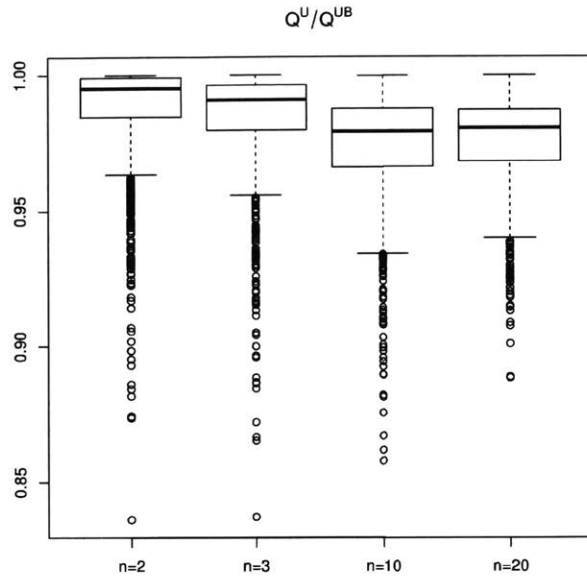


Figure 1-3: Boxplot of the Relative Performance of Uniform Co-payments - Cournot Competition with Nonlinear Demand

1.7.2 Results for Cournot Competition with Nonlinear Demand

In this section we extend the numerical study of the performance of uniform subsidies for Cournot competition with nonlinear demand. Recall from equation (1.11) that this model is not one of the special cases of formulation (*CAP*) given in Section 1.3.2, as it would correspond to each firm having a non-separable marginal cost function $\tilde{h}(q_i, Q) = h_i(q_i) - P'(Q)q_i$, which depends on the market output of all the other firms. Nonetheless, we will use additional modeling techniques that will allow us to compute numerically a bound on the performance of uniform subsidies in experiments for this case as well.

We again consider a fairly general experimental setup, with non-linear demand and non-linear marginal costs; the difference is that now firms are assumed to engage in Cournot competition. Specifically, we assume $P(Q) = a - bQ^{m_0}$, and $h_i(q_i) = c_i + g_i q_i^{m_i}$ for each i , where all the parameters are positive. Under these assumptions, the non-separable marginal cost function of each firm becomes $\tilde{h}(q_i, Q) = c_i + g_i q_i^{m_i} + m_0 b Q^{m_0-1} q_i$. Moreover, we consider the natural generalizations for the

co-payment allocation problem (*CAP*) and the upper bound problem (*UBP*) where we directly replace the marginal cost function $h_i(q_i)$ by the non-separable function $\tilde{h}(q_i, Q)$, therefore we skip the problem statements here.

Note that both problems (*CAP*) and (*UBP*) are non-convex for our experimental setup, however we will be able to solve problem (*UBP*) efficiently as follows. First, from the continuity of the functions $P(Q)$ and $\tilde{h}(q_i, Q)$, and the monotonicity of the objective function it follows that any optimal solution to problem (*UBP*) must be such that the budget constraint is tight. Second, note that the non-separable marginal cost function of each firm is increasing in its own market output, namely $\frac{d\tilde{h}(q_i, Q)}{dq_i} = m_i g_i q_i^{m_i-1} + m_0 b Q^{m_0-1} + m_0(m_0-1)bQ^{m_0-2}q_i > 0$ because $Q + (m_0-1)q_i > 0$ for each $m_0 > 0$, as $q_i \leq Q$ for each firm. Hence, the function $TC(Q)$ defined below is increasing in Q .

$$\begin{aligned}
TC(Q) \equiv \min_{\mathbf{q}} \quad & \sum_{i=1}^n \tilde{h}(q_i, Q) q_i = \sum_{i=1}^n (c_i q_i + g_i q_i^{m_i+1} + m_0 b Q^{m_0-1} q_i^2) \\
\text{s.t.} \quad & \sum_{i=1}^n q_i = Q \\
& q_i \geq 0, \text{ for each } i.
\end{aligned}$$

Finally, note that the total revenue in the market $P(Q)Q = aQ - bQ^{m_0+1}$ is concave for any $m_0 > 0$, and let us denote the total revenue maximizing market output by Q^M . From the first observation it follows that the optimal objective value of problem (*UBP*), denoted by Q^{UB} , must satisfy the budget constraint with equality, namely $P(Q^{UB})Q^{UB} = TC(Q^{UB}) - B$. Moreover, if $B \geq TC(Q^M) - P(Q^M)Q^M$ then the functions $P(Q)Q$ and $TC(Q) - B$ have a unique intersection. Therefore, Q^{UB} can be computed efficiently using binary search, where a convex optimization problem must be solved to evaluate $TC(Q)$ at each iteration.

On the other hand, we need to be able to compute the market consumption induced by uniform subsidies. However, because the marginal costs functions $\tilde{h}(q_i, Q)$ are non-separable and asymmetric in the influence of any two firms, it follows that the market equilibrium induced by uniform subsidies cannot be formulated as a convex

optimization problem, see for example Correa and Stier-Moses (2011). Nonetheless, the market equilibrium can be formulated as an asymmetric variational inequality. Hence, from Corollary 1 in Aghassi et al. (2006) it follows that (\mathbf{q}^U, Q^U) is the market equilibrium induced by uniform subsidies if and only if there exists λ^U such that the solution $(\mathbf{q}^U, Q^U, \lambda^U)$ is feasible for problem (GUCAP) below, and it attains an objective value of zero.

$$\begin{aligned}
& \min_{\mathbf{q}, Q, \lambda} \sum_{i=1}^n h_i(q_i, Q)q_i - \sum_{i=1}^n P'(Q)q_i^2 + \left(P(Q) + \frac{B}{Q}\right) (\bar{Q} - Q) - \lambda\bar{Q} \\
& \text{s.t.} \quad \sum_{i=1}^n q_i = Q \\
(GUCAP) \quad & q_i \geq 0, \text{ for each } i \\
& \lambda \leq h_i(q_i, Q), \text{ for each } i \\
& \lambda \leq P(Q) + \frac{B}{Q}.
\end{aligned}$$

In fact, this implies that this solution is optimal for problem (GUCAP) because the objective is non-negative for any feasible solution, see Aghassi et al. (2006). In our setting problem (GUCAP) is non-convex. However, if a nonlinear solver finds a feasible solution $(\mathbf{q}^U, Q^U, \lambda^U)$ with objective value equal to zero, then it follows that Q^U is the market consumption induced by uniform subsidies. We use LOQO to solve the smooth non-convex problem (GUCAP), see Vanderbei (2006), and in our experiment the solver finds the optimal solution in 94% of the instances considered.

As before, we consider the number of firms being $n \in \{2, 3, 10, 20\}$, and for each case we solve 1,000 instances randomly generated with parameters sampled from the same distributions as in Section 1.7.1. The results for the ratio Q^U/Q^{UB} are displayed in Figure 1-3 and in Table 1.3. The minimum value of the ratio Q^U/Q^{UB} never went below 83% in the simulation results. Moreover, the mean and median values are above 97%, for each value of the number of firms participating in the market n , which are relatively better results compared to the ones we obtained for price taking firms in the same setting. As before, in this case we are not comparing directly to the optimal solution, but to an upper bound. This suggests that, in most cases, the

aggregated market consumption induced by uniform co-payments is fairly close to the aggregated market consumption induced by the optimal co-payment allocation, even for the setting of Cournot competition with nonlinear demand considered here.

1.8 Conclusions

We provide a new modeling framework to analyze the problem of a central planner injecting a budget of subsidies into a competitive market, with the objective of maximizing the aggregated market consumption of a good. This is equivalent to maximizing the consumer surplus. The co-payment allocation policy that is usually implemented in practice is uniform, in the sense that every firm gets the same co-payment. A central question in this chapter is how efficient uniform co-payments are compared to the optimal subsidy allocation, assuming that some firms could be significantly more efficient than others.

Using our framework, we show that uniform co-payments are in fact *optimal* for a large a family of marginal cost functions. Moreover, we show that the optimality of uniform co-payments is preserved, under less general conditions, in the case where the central planner is uncertain about the market state. Furthermore, we show that uniform co-payments also maximize the social welfare for a large a family of marginal cost functions. Finally, we study the performance of uniform co-payments in relevant settings where they are not optimal. Our simulation results suggest that the aggregated market consumption induced by uniform co-payments is relatively close to the aggregated market consumption induced by the optimal co-payment allocation. It is an interesting research question to explore whether there exist theoretical bounds on the effectiveness of uniform subsidies in these settings.

In summary, we present interesting evidence that gives theoretical support to the use of uniform co-payments in practice. Therefore, decision makers facing the problem of allocating subsidies to increase the aggregated market consumption of a good, should not spend time and resources developing sophisticated allocation policies, as it is very likely that the very simple uniform subsidy policy will attain most of the

potential benefits. Future research on this topic should study whether these insights are preserved in dynamic models, where the subsidy allocation may change over time, or under different market equilibrium conditions, such as supply function equilibria.

Chapter 2

85% Worst-Case Performance

Guarantee for Uniform Co-Payments

2.1 Introduction

In many relevant settings the aggregated market consumption of a good is less than what is considered socially optimal. This is generally due to the positive societal externalities generated by its consumption, which, by definition, are not internalized by consumers. Classical examples of such goods include vaccines and infectious disease treatments, see Brito et al. (1991) and Arrow et al. (2004), respectively. One frequently implemented method to address this problem, is having a central planner intervening the market by allocating fixed per unit subsidies to the producers of the good, with the objective of increasing its market consumption. These type of subsidies are known as co-payments.

Let us emphasize that the co-payment allocation policy most often implemented in practice is uniform, in the sense that every firm gets the same co-payment, regardless of any differences in their cost structure or efficiency, see for example AMFm Independent Evaluation Team (2012) for the case of new malaria drugs. This is probably due to the simplicity and ease of implementation of this policy. How close the market consumption induced by uniform co-payments is, to the one induced by the optimal co-payment allocation, can be the key to effectively correct market imperfections,

such as the ones found in markets with large positive externalities. This motivates the goal of understanding the optimal co-payment allocation structure, and providing insights on how suboptimal uniform co-payments can be in the worst case.

2.1.1 Main Contributions

In Section 1.7 in Chapter 1 of this thesis we presented simulation results in relevant settings where uniform co-payments are not optimal. They suggest that, on average, the aggregated market consumption induced by uniform co-payments is relatively close to the one induced by optimal co-payments. In this chapter we focus on one of these settings, namely Cournot competition with linear demand and constant marginal costs. For each instance of this model, we show that uniform co-payments will *always* induce *at least* 85% of the market consumption induced by the optimal co-payments allocation. Namely, we show an 85% worst-case performance guarantee for uniform co-payments in maximizing the market consumption for Cournot competition with linear demand and constant marginal costs.

Specifically, we characterize the optimal co-payment allocation in this setting, as well as the market equilibrium it induces. We show that the optimal allocation consists of giving larger co-payments to less efficient firms. We argue that this policy is hard to implement in practice, and thus we study the performance of the more practical, and frequently implemented, uniform co-payments policy. We also characterize the market equilibrium induced by uniform co-payments. Hence, we are able to write a non-convex optimization problem to minimize the ratio of the market consumption induced by uniform co-payments, over the market consumption induced by optimal co-payments. The variables of this problem are the demand's parameters, the firms' marginal costs, and the central planner's budget; and its optimal solution characterizes the worst case performance of uniform co-payments for this model.

Moreover, we derive a linear program whose optimal solution provides a lower bound for the worst case performance of uniform co-payments in maximizing the market consumption. We then use duality theory to find its optimal objective value in closed form. This result allows us to discard multiple local optima and simplify

the non-convex optimization problem previously described. Finally, we consider other relaxations of the simplified problem, which allow us to solve it in closed form. Its solution shows an *asymptotically tight* worst case performance guarantee of $\frac{2+\sqrt{2}}{4} \approx 85.31\%$. Namely, we show that, in each instance of this model, uniform co-payments will induce at least 85% of the largest market consumption that can be attained. The results in this chapter suggest that the efficiency loss induced by uniform co-payments can be expected to be relatively small. Hence, this bounded efficiency loss should be weighted against the important practical advantages of uniform co-payments, such as their ease of implementation and communication.

2.2 Literature Review

Cournot competition with linear demand and constant marginal costs is a simple oligopoly model where firms compete in quantity. It is a well understood model that provides interesting insights. Therefore, it is frequently used by researchers as a building block to study complex phenomena. Examples of this trend in the operations management and operations research literature include using this model, among others, to study the structure of supply chains, see Corbett and Karmarkar (2001), supply chain contracts, see Cachon (2003), production under yield uncertainty, see Deo and Corbett (2009), firms' profits compared to other equilibrium concepts, see Farahat and Perakis (2011), and facility network design under competition, see Dong et al. (2013). Importantly for us, uniform co-payments are not optimal for Cournot competition with linear demand and constant marginal costs. Hence, in the same spirit as in previous literature, we consider this model as a relevant example to study how much market consumption can be lost, in the worst case, when implementing uniform co-payments instead of the optimal co-payments policy. As pointed out in most of the papers in the aforementioned literature, assuming linear demand and constant marginal costs allows us to write close form expressions for the market equilibrium, in our case for any given co-payments allocation. These are important in order to derive the *asymptotically tight* worst-case performance guarantees for uniform

co-payments that we present in this chapter.

Similarly, the model of Cournot competition with linear demand and constant marginal costs has also been used as a building block in other areas, such as marketing and economics. Some examples in the marketing literature include analyzing channel structure, see Choi (1991), and process innovation and product differentiation, see Gupta and Loulou (1998). On the other hand, some examples in the economics literature include comparing price versus quantity contracts, see Singh and Vives (1984) and Häckner (2000), as well as studying experimentation and learning with uncertain product differentiation, see Harrington (1995).

One particular area in the economics literature that studies a problem related to the one considered in this chapter, is the "third market model" in strategic trade policy, see Brander (1995). In this model, n home firms and n^* foreign firms export a commodity to a third market, where the market price is set through Cournot Competition with constant marginal costs. The government can allocate subsidies to the home firms, increasing their profit at the expense of the foreign competitors. The government's utility is equal to the profit earned by the home firms, minus the cost of the subsidy payments. Note that the government does not face a budget constraint. In a model with heterogeneous firms, Collie (1993) and Long and Soubeyran (1997) *assume* a uniform subsidy, and study its effect in the market shares of the firms. Later, Leahy and Montagna (2001) assume linear demand, and derive closed form expressions for the optimal subsidies. They conclude that the optimal subsidy policy is generally not uniform, and the government should allocate higher export subsidies to the *more* efficient firms. In contrast, in our model we find that it is optimal to allocate higher co-payments to the *less* efficient firms. This difference is driven by the fact that the central planner's objective in our model is to maximize the market consumption, as opposed to maximize the firms' profits. Although in our setting uniform co-payments are not optimal either, the main result in this chapter shows that, in the worst case, they will induce at least 85% of the market consumption induced by optimal co-payments.

2.3 Model

In this section, we introduce a mathematical programming formulation for allocating co-payments, to Cournot competitors with linear demand and constant marginal costs. We consider a market for a commodity composed by $n \geq 2$ heterogeneous competing firms. We assume that each firm $i \in \{1, \dots, n\}$ decides its output $q_i \geq 0$ independently, with the goal of maximizing its own profit, and that it has a constant marginal cost $c_i \geq 0$. Consumers are described by a linear inverse demand function $P(Q) = a - bQ$, where $Q \equiv \sum_{i=1}^n q_i$ is the aggregated market consumption, and $a > 0$, $b > 0$ are the demand parameters. We will assume, without loss of generality, that the firms are labeled such that $c_1 \leq c_2 \leq \dots \leq c_n \leq a$.

In terms of the market equilibrium dynamics, we assume that the firms engage in Cournot competition, see Cournot (1897). Namely, that given all the other firms' outputs, each firm $i \in \{1, \dots, n\}$ sets its output q_i at a level such that it maximizes its own utility $\Pi_i(q_i, Q) = (P(Q) - c_i)q_i = (a - bQ - c_i)q_i$. That is, each firm $i \in \{1, \dots, n\}$ simultaneously solves the following problem

$$\max_{q_i \geq 0} \Pi_i(q_i, Q), \quad (2.1)$$

and the market equilibrium consists of the fixed point attained at the intersection of the best responses of all the firms. The necessary and sufficient first order condition of problem (2.1), implies that each firm i participating in the market equilibrium produces up to the point where its marginal cost equals its marginal revenue, namely $c_i = P(Q) + P'(Q)q_i = a - bQ - bq_i$; and each firm j that does not participate in the market equilibrium, must have a marginal cost, which is larger than the market price, namely $c_j \geq P(Q) = a - bQ$. This can be summarized in the following equilibrium condition:

$$\text{For each } i, j, \text{ if } q_i > 0, \text{ then } c_i + bq_i = a - bQ \leq c_j + bq_j. \quad (2.2)$$

The existence and uniqueness of the market equilibrium in this setting is well known,

see for example Tirole (1988).

As already mentioned in Section 2.1, we focus on settings where the market consumption induced at the market equilibrium is less than what is socially optimal. For this reason, a central planner intervenes the market by allocating fixed per unit subsidies, or co-payments, to each firm. We will refer to the problem faced by the central planner as the *co-payment allocation problem (CAP)*. The co-payment allocation problem is a particular case of a Stackelberg game, see Stackelberg (1952), or a bilevel optimization problem. In the first stage, the central planner allocates her budget $B > 0$ in the form of co-payments $y_i \geq 0$, per unit provided in the market, to each firm $i \in \{1, \dots, n\}$. Moreover, she anticipates that in the second stage the equilibrium output of each firm will satisfy a modified version of the equilibrium condition (2.2), stated below in constraint (2.7). The main difference in the market equilibrium condition (2.7), with respect to (2.2), is given by the fact that, from firm i 's perspective, the effective price for each unit sold is now $P(Q) + y_i = a - bQ + y_i$.

The central planner's objective is to maximize the aggregated market consumption. Finally, let us emphasize that we have assumed that the central planner can only allocate co-payments, and never charge taxes for the units produced in the market. In other words, the allocated co-payments have to be non-negative. This is the case in the practical problems that motivate this research, as the central planner is seldom in the position of charging taxes to firms that may operate in different countries, with the goal of increasing the aggregated market consumption of a good, see for example Arrow et al. (2004) for the case of malaria drugs. Hence, the following is a valid formulation of the co-payment allocation problem

$$\begin{aligned} \max_{\mathbf{y}, \mathbf{q}, Q} \quad & Q \\ \text{s.t.} \quad & \sum_{i=1}^n q_i y_i \leq B \end{aligned} \tag{2.3}$$

$$y_i \geq 0, \text{ for each } i \in \{1, \dots, n\} \tag{2.4}$$

$$\sum_{i=1}^n q_i = Q \tag{2.5}$$

$$q_i \geq 0, \text{ for each } i \in \{1, \dots, n\} \tag{2.6}$$

$$c_i + bq_i = a - bQ + y_i, \text{ for each } i \in \{1, \dots, n\}. \tag{2.7}$$

Constraint (2.3) is the budget constraint, where the total amount spent in co-payments can be at most the available budget B . Constraints (2.4) and (2.7) are the non-negativity of the co-payments, and the modified equilibrium condition previously discussed, respectively.

This is a valid formulation even if there are firms that do not participate in the market equilibrium. Namely, if for some firm i we have $q_i = 0$, then, from the modified market equilibrium condition (2.7), and the non-negativity of the co-payments (2.4), we must have $y_i = c_i - (a - bQ) \geq 0$. Hence, the non-negativity of the co-payment y_i , in constraint (2.4), exactly ensures that the original market equilibrium condition (2.2) is satisfied. Moreover, allocating a co-payment $y_i = c_i - (a - bQ) \geq 0$ to firm i is without loss of generality, because setting $q_i = 0$ ensures that firm i does not have an impact in the budget constraint (2.3). In other words, the fact that we impose the modified market equilibrium condition (2.7) on each firm $i \in \{1, \dots, n\}$, does *not* imply that every firm has to participate in the market equilibrium.

From the equilibrium condition (2.7), it follows that we can replace all the co-payment variables y_i by $c_i + bq_i - (a - bQ)$. Namely, we can reformulate the co-payment allocation problem as if the central planner was deciding the output of each firm, as long as there exist feasible co-payments that can sustain the outputs chosen as the market equilibrium. The feasibility of the co-payments is given by both the budget constraint (2.3), and the non-negativity of the co-payments (2.4). It follows that, the

co-payments allocation problem faced by the central planner can be reformulated as described next

$$\begin{aligned} \max_{\mathbf{q}, Q} \quad & Q \\ \text{s.t.} \quad & \sum_{i=1}^n (c_i q_i + b q_i^2) - (a - bQ)Q \leq B \end{aligned} \quad (2.8)$$

$$(CAP) \quad c_i + b q_i \geq a - bQ, \text{ for each } i \in \{1, \dots, n\} \quad (2.9)$$

$$\sum_{j=1}^n q_j = Q \quad (2.10)$$

$$q_i \geq 0, \text{ for each } i \in \{1, \dots, n\}, \quad (2.11)$$

where constraint (2.8) is equivalent to the budget constraint (2.3), and constraint (2.9) is equivalent to the non-negativity of the co-payments (2.4). The co-payments that the central planner must allocate, to induce outputs \mathbf{q} , are $y_i = c_i + b q_i - a - bQ$, for each $i \in \{1, \dots, n\}$. Note that problem (CAP) is a convex optimization problem.

The first question we will address, in the following section, is whether we can characterize the structure of the optimal co-payments, as well as the structure of the induced market equilibrium.

2.3.1 Optimal Co-payments Allocation

In this section, we will show that we can characterize the structure of any optimal solution to problem (CAP). Specifically, we will provide *closed form expressions* for the market consumption, and each firm's market output, induced by optimal co-payments, which are parametrized by indexes $l, m \in \{1, \dots, n\}$ that are defined below. Additionally, we discuss the practical challenges that would have to be faced in order to implement this solution, which suggest that the uniform co-payment policy can be more attractive from a practical perspective. All the proofs are presented in Appendix A.2.

In more details, Proposition 5 below shows that the market outputs induced by the optimal co-payments have the following intuitive structure: more efficient firms produce more than less efficient firms, up to the point where firms are so inefficient that

they do not participate in the market equilibrium induced by optimal co-payments. The latter is characterized by an index $m \in \{1, \dots, n\}$, associated to the last firm that has a positive output in the market equilibrium induced by optimal co-payments.

Similarly, the optimal co-payments have the following structure. The more efficient firms may not get any co-payments, and only after some index $l \in \{1, \dots, m\}$, firms start getting a co-payment that is increasing in their marginal cost (see Proposition 5 below). Namely, in order to maximize the market consumption at equilibrium, for Cournot competitors with linear demand and constant marginal costs, the best that the central planner can do with her co-payments is to give *more co-payments to less efficient firms*. This structural result is driven by the central planner's objective of maximizing the market consumption. For example, it can be shown that if the central planner's objective was to minimize the total cost instead, then we would obtain the opposite result, where more efficient firms would get a larger co-payment at optimality. This is in agreement with the observations made by researchers in different problem settings, see for example Leahy and Montagna (2001).

Proposition 5. *Any optimal solution of the co-payments allocation problem (CAP), (\mathbf{q}^*, Q^*) , is such that the budget constraint (2.8) is tight, and there exist indexes $l, m \in \{1, \dots, n\}$, with $l \leq m$, such that the optimal co-payments are given by*

$$y_i^* = 0, \text{ for each } i \in \{1, \dots, l-1\}, \quad (2.12)$$

$$y_i^* = y_l^* + \frac{c_i - c_l}{2} > 0, \text{ for each } i \in \{l, \dots, m\}, \quad (2.13)$$

$$y_i^* = y_l^* + c_i - c_l - bq_l^* \geq y_l^* + \frac{c_i - c_l}{2} > 0, \text{ for each } i \in \{m+1, \dots, n\}. \quad (2.14)$$

The optimal market output of each firm are given by

$$q_i^* = q_l^* + \frac{c_l - c_i}{b} - \frac{y_l^*}{b} \geq q_l^* + \frac{c_l - c_i}{2b} > 0, \text{ for each } i \in \{1, \dots, l-1\}, \quad (2.15)$$

$$q_i^* = q_l^* - \frac{c_i - c_l}{2b} > 0, \text{ for each } i \in \{l, \dots, m\}, \quad (2.16)$$

$$q_i^* = 0, \text{ for each } i \in \{m+1, \dots, n\}, \quad (2.17)$$

where the market output of the first firm that receives a positive co-payment q_i^* is given by

$$q_i^* = \frac{a + \sum_{i=1}^m c_i}{(m+1)b} - \frac{c_l}{b} - \sum_{i=l+1}^m \frac{c_i - c_l}{2(m+1)b} + \frac{l}{(m+1)b} y_l^*. \quad (2.18)$$

The expressions (2.13)-(2.18) are written as a function of the first positive co-payment y_l^* , which is given by

$$\begin{aligned} y_l^* = & - \left(\frac{a + \sum_{i=1}^{l-1} c_i - lc_l}{2l} \right) + \\ & \left(\left(\frac{a + \sum_{i=1}^{l-1} c_i - lc_l}{2l} \right)^2 + \frac{m+1}{(m-l+1)l} \sum_{i=l}^m \frac{(c_i - c_l)^2}{4} + \frac{m+1}{(m-l+1)l} bB \right. \\ & \left. - \sum_{i=l}^m \frac{c_i - c_l}{2} \left(a + \sum_{i=1}^{l-1} c_i + \sum_{i=l+1}^m \frac{c_i}{2} - (m+l) \frac{c_l}{2} \right) \frac{1}{(m-l+1)l} \right)^{1/2}. \end{aligned} \quad (2.19)$$

Finally, the aggregated market consumption is given by

$$\begin{aligned} Q^* = & \frac{1}{2(m+1)lb} \left((2lm - m + l - 1)a - (m+l+1) \sum_{i=1}^{l-1} c_i - lc_l - l \sum_{i=l+1}^m c_i \right) \\ & + \frac{m-l+1}{2(m+1)lb} \left(\left(a + \sum_{i=1}^{l-1} c_i - lc_l \right)^2 + \frac{l(m+1)}{m-l+1} \sum_{i=l+1}^m (c_i - c_l)^2 + \frac{4l(m+1)}{m-l+1} bB \right. \\ & \left. - \frac{l}{m-l+1} \left(\sum_{i=l+1}^m (c_i - c_l) \right) \left(2a + 2 \sum_{i=1}^{l-1} c_i + \sum_{i=l+1}^m c_i - (m+l)c_l \right) \right)^{1/2}. \end{aligned} \quad (2.20)$$

Practical Challenges. Proposition 5 characterizes the optimal co-payment y_i^* , and the induced market output q_i^* , for each firm $i \in \{1, \dots, n\}$, as well as the induced market consumption Q^* . Moreover, it provides closed form expressions that are parametrized by the indexes l , of the first firm that receives a co-payment, and m , of the last firm that has a positive market output. Nonetheless, if we wanted to transfer these insights into practice, we would have to keep in mind that the optimal co-payments policy imposes the following challenges. First, the optimal co-payments are a complicated function of the problem parameters, which would make them difficult to communicate. Second, they are different for each firm, which would signifi-

cantly increase the complexity of the process of paying to the manufacturers. More importantly, the optimal co-payments policy requires the central planner to know the marginal cost c_i of *each* firm i , as well as being highly sensitive to changes in the value of the marginal costs. Although in our model we assumed a full information setting, where the central planner knows the marginal cost of each firm, in practice this may not be the case. Therefore, any practical implementation would either require a truthful mechanism to elicit the marginal costs, or alternatively it would have to deal with potential misspecifications. In contrast, the uniform co-payments policy is simple to communicate and control. Additionally, we will see in the next section that it only depends on the *average* marginal cost of the firms, hence it is more stable to misspecifications. These characteristics make the uniform co-payments a more attractive policy for practical purposes, as long as the loss in the induced market consumption, with respect to the optimal co-payments policy, is not very large.

The expression for the market consumption induced by optimal co-payments Q^* in equation (2.20) is nonlinear, and quite complex to work with. Therefore, because we are interested in the worst case performance of uniform co-payments in maximizing the market consumption, it would be desirable to have a simpler expression that is an upper bound on Q^* , to compare with. The following lemma addresses this point by providing an upper bound on Q^* , which has a linear expression in the marginal costs c_i , and in the demand parameter a .

Lemma 6. *If the indexes l , m defined in Proposition 5 satisfy $m \geq 2$, and $l \in \{2, \dots, m\}$, then the following is a valid upper bound on the total market consumption Q^* induced by the optimal co-payment allocation*

$$Q^* \leq \frac{2ma - 2 \sum_{i=1}^{l-2} c_i - (m-l+3)c_{l-1} - \sum_{i=l}^m c_i}{2(m+1)b}. \quad (2.21)$$

This bound is attained when $y_i^ = \frac{a-c_{i-1}}{2}$.*

2.3.2 Uniform Co-payments Allocation

In this section, we will characterize the structure of the market equilibrium induced by uniform co-payments. Specifically, we will provide *closed form expressions* for the market consumption induced by uniform co-payments, as well as for the market output of each firm. These closed form expressions are parametrized by an index $u \in \{1, \dots, n\}$, which denotes the last firm that might have a positive output in the market equilibrium induced by uniform co-payments. Additionally, we will provide a lower bound on the induced market consumption, which has a linear expression in the marginal costs c_i . All the proofs are presented in Appendix A.2.

By definition, the uniform co-payments allocation gives the same co-payment to each firm. In this setting, larger co-payments will clearly lead to a larger market output, therefore we focus on the largest possible uniform co-payment that can be afforded with the central planner's budget B . Namely, if we denote by q_i^U the output of firm i induced by the uniform co-payment y^U , then we will focus on the value of y^U that makes the budget constraint tight. That is, $\sum_{i=1}^n q_i^U y^U = B$, or equivalently $y^U = \frac{B}{Q^U}$.

In words, the amount of the uniform co-payment is obtained by simply dividing the available budget B , by the the largest market consumption that be can attained with this budget under a uniform co-payment policy, denoted by Q^U . In practice, the way this policy is usually implemented is by dividing the budget by a target market consumption that the central planner has set as a goal, see for example AMFm Independent Evaluation Team (2012) for the case of new malaria drugs. Let us emphasize again that the uniform co-payments policy is conceptually simple, and easy to communicate and control. In terms of the parameters in our model, the structure of the market equilibrium induced by uniform co-payments is described in the following lemma.

Lemma 7. *Define, without loss of generality, $c_{n+1} = a$. Then, the market output*

induced by the uniform co-payments allocation, (\mathbf{q}^U, Q^U) , is

$$q_i^U = \frac{a}{b} - Q^U + \frac{B}{bQ^U} - \frac{c_i}{b} \geq 0, \text{ for each } i \in \{1, \dots, u\}, \quad (2.22)$$

$$q_i^U = 0, \text{ for each } i \in \{u+1, \dots, n\}, \quad (2.23)$$

$$Q^U = \frac{ua - \sum_{i=1}^u c_i + \sqrt{(ua - \sum_{i=1}^u c_i)^2 + 4u(u+1)bB}}{2(u+1)b}, \quad (2.24)$$

where $u \in \{1, \dots, n\}$, is the smallest index such that $c_i > a - bQ^U + \frac{B}{Q^U}$, for each $i \in \{u+1, \dots, n+1\}$.

The uniform co-payment that induces this market output is

$$y_i^U = \frac{B}{Q^U} \geq 0, \text{ for each } i \in \{1, \dots, n\}. \quad (2.25)$$

The expression for the market consumption induced by uniform co-payments Q^U , in equation (2.24), is much simpler than the one we have for Q^* , in equation (2.20). In particular, it only depends on the *average* marginal cost of the firms that are active in the market equilibrium. However, this expression is still nonlinear. Similarly to the previous section, it would be desirable to have a simpler expression that is a *lower* bound on Q^U , to compare with. Lemma 8 below provides such a lower bound, which has a linear expression in the marginal costs.

Lemma 8. *The following is a valid lower bound for the market consumption induced by the uniform co-payments allocation.*

$$Q^U \geq \frac{uc_u - \sum_{i=1}^u c_i}{b}. \quad (2.26)$$

Where the bound in equation (2.26) is attained when $q_u^U = 0$.

2.3.3 Consistency Constraints on Parameters a , b , c_i and B

Proposition 5 and Lemma 7 provide closed form expressions for the market consumption induced by optimal co-payments Q^* , and by uniform co-payments Q^U , respec-

tively. These closed form expressions are parametrized by the indexes l , m and u . Recall that l is the index of the first firm that receives a positive co-payment, and m is the index of the last firm with a positive output, in the market equilibrium induced by the optimal co-payments allocation. Similarly, recall that u is the last index of a firm that might have a positive market output, in the market equilibrium induced by uniform co-payments. However, the existence of the indexes l, m, u induces consistency constraints on the parameters a, b, c_i and B . They are given by the natural constraints $q_u^U \geq 0$, $0 \leq y_l^* \leq \frac{c_l - c_{l-1}}{2}$, and the definition of the index u in Lemma 7.

Specifically, in equation (2.22) we have that $q_u^U = \frac{a}{b} - Q^U + \frac{B}{bQ^U} - \frac{c_i}{b} \geq 0$, which implies

$$bB \geq \left(uc_u - \sum_{i=1}^u c_i \right) \left((u+1)c_u - \sum_{i=1}^u c_i - a \right). \quad (2.27)$$

On the other hand, the definition of index u in Lemma 7 states that $c_{u+1} > a - bQ^U + \frac{B}{Q^U}$, which implies

$$bB \leq \left(uc_{u+1} - \sum_{i=1}^u c_i \right) \left((u+1)c_{u+1} - \sum_{i=1}^u c_i - a \right). \quad (2.28)$$

Similarly, equation (2.15) implies $y_l^* \leq \frac{c_l - c_{l-1}}{2}$ (or equivalently $y_{l-1}^* = 0$), which in turn implies

$$bB \leq \frac{1}{4(m+1)} \left(\sum_{i=l}^m (c_i - c_{l-1}) \right) \left(2a + 2 \sum_{i=1}^{l-2} c_i + \sum_{i=l}^m c_i - (m+l-1)c_{l-1} \right) - \sum_{i=l}^m \frac{(c_i - c_{l-1})^2}{4}. \quad (2.29)$$

Finally, defining $y_l^* \geq 0$, implies

$$bB \geq \frac{1}{4(m+1)} \left(\sum_{i=l+1}^m (c_i - c_l) \right) \left(2a + 2 \sum_{i=1}^{l-1} c_i + \sum_{i=l+1}^m c_i - (m+l)c_l \right) - \sum_{i=l+1}^m \frac{(c_i - c_l)^2}{4}. \quad (2.30)$$

Note that constraints (2.27) and (2.28) follow from substituting the expression for Q^U , given in equation (2.24), and solving for the budget B . Similarly, constraints (2.29) and (2.30) follow from substituting the expression for y_i^* , given in equation (2.19), and solving for the budget B . Additionally, note that constraint (2.30) ensures that the square root in the expression for Q^* , in equation (2.20), is well defined.

Our goal in this chapter is to characterize the worst case performance of uniform co-payments, in maximizing the market consumption in this setting. Therefore, a natural approach is to write a mathematical program that minimizes the ratio of the closed form expressions for Q^U and Q^* , parametrized by the indexes l , m and u in Lemma 7 and Proposition 5, respectively, as a function of the problem parameters: the number of firms in the market n , the demand parameters a , b , the marginal cost of each firm c_i , and the budget B . This is precisely what we will do in the remainder of the chapter, where we will need to make sure that the consistency constraints introduced in this section are satisfied.

2.4 Preliminary Results and Problem Statement

In this section, we present a mathematical program whose optimal solution quantifies the worst case performance of uniform co-payments, in maximizing the aggregated market consumption, for Cournot competition with linear demand and constant marginal costs. In order to do so, we will start by giving a set of preliminary results that will allow us to simplify its formulation, as well as providing useful tools for its analysis.

We begin with a simple observation that will allow us, without loss of generality, to scale the marginal costs of each firm c_i , and the demand parameter a . The proofs of Lemmas 9 and 10, as well as the proofs of Propositions 6 and 7 below, are provided in Appendix A.2.

Lemma 9. *For any instance of the co-payments allocation problem (CAP) with $c_1 \geq 0$, and for any scaling parameter $\delta \geq 0$, there exists a modified instance with $c_1 = \delta$ such that the modified instance has the same set of optimal solutions,*

which attain the same objective value.

From Lemma 9 it follows that we can assume, without loss of generality, that $c_1 = 0$ in the rest of the analysis. Nonetheless, in order to simplify some proofs we will assume $c_1 = \delta > 0$, when convenient.

From Lemma 10 below, it follows that there exists an instance that attains the worst-case in the performance of uniform co-payments in maximizing the market consumption in this setting.

Lemma 10. *For any given number of firms in the market $n \geq 2$, there exists an instance of problem (CAP) that minimizes the ratio Q^U/Q^* .*

Additionally, Propositions 6 and 7 below allow us to reduce the family of instances of problem (CAP) that we need to consider, in order to quantify the worst case performance of uniform co-payments in maximizing the market consumption in this setting. This will be crucial to simplify the analysis.

Proposition 6. *For any given number of firms in the market $n \geq 2$, any instance of problem (CAP) that minimizes the ratio Q^U/Q^* must be such that the indexes given in Proposition 5 and Lemma 7 must satisfy $m = u = n$.*

Proposition 7. *For any given number of firms in the market $n \geq 2$, any instance of problem (CAP) that minimizes the ratio Q^U/Q^* must be such that $q_n^U = 0$. It follows that both the consistency constraint (2.27), and the upper bound for Q^U from Lemma 8, must be tight.*

Proposition 6 states that, without loss of generality, we can focus on instances of problem (CAP) such that, in the market equilibrium induced by the optimal co-payments allocation, all the firms in the market have a positive market output. Namely, such that $m = n$. Similarly, Propositions 6 and 7 state that, without loss of generality, we can focus on instances of problem (CAP) such that, in the market equilibrium induced by the uniform co-payments, the last firm in the market is exactly on the verge of start having a positive market output. Namely, such that $u = n$ and $q_n^U = 0$. From the perspective of the mathematical program we are constructing,

Propositions 6 and 7 significantly reduce the number of instances we need to consider, by fixing the values of the indexes m and u to n , as well as fixing the value of the budget B , as a function of the marginal costs c_i , and the demand parameters a , b , as in the consistency constraint (2.27). Additionally, Proposition 7 allows us to work directly with the linear expression for Q^U from Lemma 8.

2.4.1 Problem Statement

Now we are ready to define the problems, and sub-problems, that we are interested in solving, in order to identify the worst case performance of uniform co-payments in maximizing the market consumption in this setting. We will check later that, for any given number of firms in the market $n \geq 2$, the instance of problem (CAP) that minimizes the ratio Q^U/Q^* will be defined by the marginal costs c_i , and the budget B , for any given demand parameter values $a > 0$, $b > 0$. Moreover, the value of the worst case ratio will be independent of the values of a , b . Therefore, we will only consider c_i , for each $i \in \{1, \dots, n\}$, and B as variables, while we will treat $a > 0$ and $b > 0$ as parameters.

To simplify the notation let us define the function

$$\begin{aligned} & \sqrt{*}_l(B, \mathbf{c}) \\ \equiv & (n-l+1) \left(\left(a + \sum_{i=1}^{l-1} c_i - lc_l \right)^2 + \frac{l(n+1)}{n-l+1} \sum_{i=l+1}^n (c_i - c_l)^2 + \frac{4l(n+1)}{n-l+1} bB \right. \\ & \left. - \frac{l}{n-l+1} \left(\sum_{i=l+1}^n (c_i - c_l) \right) \left(2a + 2 \sum_{i=1}^{l-1} c_i + \sum_{i=l+1}^n c_i - (n+l)c_l \right) \right)^{\frac{1}{2}}. \end{aligned} \quad (2.31)$$

The problem we want to solve is introduced in Proposition 8. The proof is presented in Appendix A.2.

Proposition 8. *For any given number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, let $WC(l, n)$ be the worst case performance of uniform co-payments in maximizing the market consumption, assuming that the index $l \in \{1, \dots, n\}$ is fixed. Then, $WC(l, n)$ can be computed as the optimal objective value of*

problem (WCP_l) below

$$\min_{B, \mathbf{c}} \frac{WC(l, n) \equiv Q^U(\mathbf{c})}{Q_l^*(B, \mathbf{c})} = \frac{2(n+1)l(nc_n - \sum_{i=1}^n c_i)}{(2ln - n + l - 1)a - (n + l + 1) \sum_{i=1}^{l-1} c_i - l \sum_{i=l}^n c_i + \sqrt{*}_i(B, \mathbf{c})}$$

(2.32)

s.t. $c_1 = 0$

$$c_i \leq c_{i+1}, \text{ for each } i \in \{1, \dots, n-1\} \quad (2.33)$$

$$c_n \leq a \quad (2.34)$$

$$bB = \left(nc_n - \sum_{i=1}^n c_i \right) \left((n+1)c_n - \sum_{i=1}^n c_i - a \right) \quad (2.35)$$

$$(WCP_l) \quad bB \leq \frac{1}{4(n+1)} \left(\sum_{i=l}^n (c_i - c_{l-1}) \right) \left(2a + 2 \sum_{i=1}^{l-2} c_i + \sum_{i=l}^n c_i - (n+l-1)c_{l-1} \right) - \sum_{i=l}^n \frac{(c_i - c_{l-1})^2}{4} \quad (2.36)$$

$$bB \geq \frac{1}{4(n+1)} \left(\sum_{i=l+1}^n (c_i - c_l) \right) \left(2a + 2 \sum_{i=1}^{l-1} c_i + \sum_{i=l+1}^n c_i - (n+l)c_l \right) - \sum_{i=l+1}^n \frac{(c_i - c_l)^2}{4} \quad (2.37)$$

$$B \geq 0. \quad (2.38)$$

In Proposition 8, $WC(l, n)$ is defined restricting ourselves to instances of problem (CAP) where the first firm that receives a positive co-payment, in the optimal co-payment allocation, is exactly l . This definition is motivated by the fact that, given an index $l \in \{1, \dots, n\}$, Proposition 5 provides a closed form expression for the market consumption induced by the optimal co-payments allocation, Q^* . Note that, to simplify the notation of $WC(l, n)$, we have omitted its dependence on the demand parameters a and b . This is because the worst case ratio will be independent of their values, as we will check later. Note that constraint (2.32) follows from Lemma 9. Additionally, constraints (2.35), (2.36) and (2.37) correspond to the consistency constraints (2.27), (2.29) and (2.30), respectively.

A priori, for any given number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, it is not clear which case of $l \in \{1, \dots, n\}$ attains the worst case

performance of uniform co-payments, in maximizing the market consumption in this setting. Therefore, we are interested in solving the following problem

$$\min_{l \in \{1, \dots, n\}} WC(l, n). \quad (2.39)$$

Solving problem (2.39) potentially requires computing $WC(l, n)$, for each $l \in \{1, \dots, n\}$. That is, solving problem (WCP_l) , for each $l \in \{1, \dots, n\}$.

Additionally, in order to compute the largest *uniform* worst-case guarantee for the performance of uniform co-payments, in maximizing the market consumption, for Cournot competition with linear demand and constant marginal costs, we are interested in solving the following problem

$$\inf_{n \in \mathbb{N}, n \geq 2, l \in \{1, \dots, n\}} WC(l, n). \quad (2.40)$$

2.5 Worst-Case Performance of Uniform Co-payments

In this section we will solve both problems (2.39) and (2.40). Namely, we will compute an asymptotically tight uniform worst case guarantee for the performance of uniform co-payments in maximizing the market consumption, for Cournot competition with linear demand and constant marginal costs. Moreover, for any given number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, we will actually identify the worst case instance for the performance of uniform co-payments in closed form, which will provide a tight worst case bound.

The following subsection describes a candidate instance to be the worst case, and the value it induces for the ratio $Q^U(\mathbf{c})/Q_l^*(B, \mathbf{c})$ in problem (WCP_l) .

2.5.1 Candidate to Worst Case Instance

Consider the following instance of problem (CAP) . For any given number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, let

$$c_1 = 0, \quad (2.41)$$

$$c_i = \left(\frac{n + \sqrt{\frac{n(n+1)}{2}}}{3n + 1} \right) a \quad \text{for each } i \in \{2, \dots, n\}, \quad (2.42)$$

$$B = \left(\frac{(n-1)\sqrt{\frac{n(n+1)}{2}}}{(3n+1)^2} \right) \frac{a^2}{b}. \quad (2.43)$$

It is not hard to check that this instance is feasible for problem (WCP_2) , that is for the case $l = 2$. Namely, in this instance the first firm that receives a positive co-payment, in the optimal co-payment allocation, is firm $l = 2$. We can also check that it attains an objective value of $\frac{2 + \sqrt{2 + 2/n}}{4}$ in problem (WCP_2) . For completeness, we prove these facts in Lemmas 15 and 16 in Appendix A.2, respectively. Note that this objective value, that is our candidate to be the worst case performance of uniform co-payments, is independent of the actual values of the demand parameters a and b . This was already mentioned in the notation of $WC(l, n)$ in Section 2.4.1, which omits the dependence on these parameters.

The rest of the chapter focuses on showing that, for any given $n \geq 2$, $a > 0$, $b > 0$, the candidate instance from equations (2.41)-(2.43), effectively attains the worst case performance of uniform co-payments in maximizing the market consumption.

2.5.2 Tight Worst-Case Performance Guarantees

This is the main result in this chapter

Corollary 1. *For any given number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, the candidate instance from equations (2.41)-(2.43) is the worst case instance for the performance of uniform co-payments in maximizing the market consumption, for Cournot competition with linear demand and constant marginal costs, attaining a value of*

$$\min_{l \in \{1, \dots, n\}} WC(l, n) = \frac{2 + \sqrt{2 + 2/n}}{4}. \quad (2.44)$$

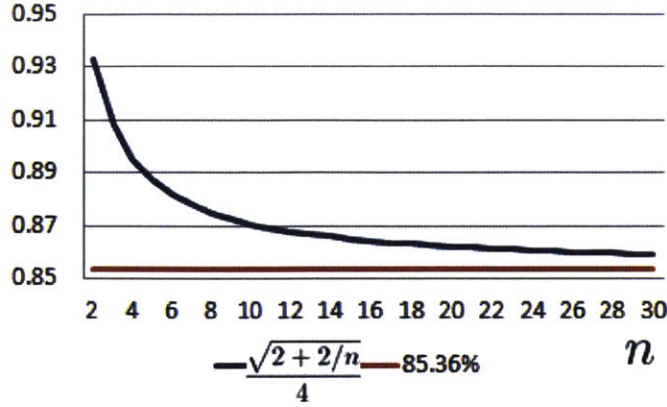


Figure 2-1: Worst-Case Guarantee for Uniform Co-Payments

Hence, an asymptotically tight bound for the performance of uniform co-payments in maximizing the market consumption in this model is

$$\inf_{n \in \mathbb{N}, n \geq 2, l \in \{1, \dots, n\}} WC(l, n) = \frac{2 + \sqrt{2}}{4} \approx 85.31\%. \quad (2.45)$$

Proof. The first result follows directly from Theorems 5 and 6 below.

The second result follows from $\frac{2 + \sqrt{2 + 2/n}}{4}$ being decreasing in n , and taking the limit as $n \rightarrow \infty$. ■

The insights from Corollary 1 are summarized in Figure 2-1. It provides a *tight* worst-case performance guarantee of $\frac{2 + \sqrt{2 + 2/n}}{4}$, for the market consumption induced by uniform co-payments, for any given $n \geq 2$, $a > 0$, $b > 0$. This bound is attained by the candidate instance from equations (2.41)-(2.43). and it decreases asymptotically to $\frac{2 + \sqrt{2}}{4} \approx 85.31\%$. Namely, the efficiency loss in maximizing the market consumption induced by implementing the much simpler uniform co-payments policy is at most 15%, for any instance of Cournot competition with linear demand and constant marginal costs. Hence, the practical advantages presented by the uniform co-payments -including ease of implementation, communication and control of the co-payments program- should be weighted against this bounded efficiency loss. Nonetheless, as shown in Figure 2-1, for any finite number of firms in the market n ,

the tight worst-case guarantee is strictly larger than the uniform bound. In particular, if the number of firms in the market is $n \in \{2, 3\}$, then uniform co-payments are guaranteed to induce more than 90% of the optimal market consumption.

Proof Structure To prove the results that imply Corollary 1, we will proceed as follows.

1. First, Theorem 5 below shows that, for any given $n \geq 2$, $a > 0$, $b > 0$, the candidate instance from equations (2.41)-(2.43) is an optimal solution of problem (WCP_2) .
2. Then, Theorem 6 shows that, for any $n \geq 2$, $a > 0$, $b > 0$, the worst case instance does not belong to the cases $l \in \{1\} \cup \{3, \dots, n\}$.

Theorem 5. *For any number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, we have that $WC(2, n) = \frac{2 + \sqrt{2 + 2/n}}{4}$. Hence, the candidate instance from equations (2.41)-(2.43) is the optimal solution to the problem (WCP_2) .*

Proof. The proof structure is the following. For any given $n \geq 2$, $a > 0$, $b > 0$, we will analyze a mathematical programming relaxation of problem (WCP_2) , denoted by $(RWCP_2)$, whose optimal solution provides a lower bound on $WC(2, n)$. We will show that solving this relaxation is equivalent to solving one of $(n - 1)$ one variable optimization problems. Then, we will show that the objective function of any of these simpler problems is lower bounded by $\frac{2 + \sqrt{2 + 2/n}}{4}$. The conclusion will follow from the fact that this objective value is attained by the candidate instance from equations (2.41)-(2.43).

All these auxiliary results are presented in Section 2.6. In particular, Lemma 11 in Section 2.6 describes the relaxation $(RWCP_2)$. From Proposition 11 in Section 2.6, it follows that solving the relaxation $(RWCP_2)$ is equivalent to solving problem $(RWCP_{2,1})$. We will show here that the candidate instance from equations (2.41)-(2.43) is the optimal solution to problem $(RWCP_{2,1})$, hence it is the optimal solution to the relaxation $(RWCP_2)$. Because this instance is in fact feasible in the original problem (WCP_2) , it follows that it is optimal for this problem as well.

From Lemma 12 in Section 2.6, it follows that the objective function of problem $(RWCP_{2,1})$ is quasiconvex. Now we show that the candidate instance from equations (2.41)-(2.43) is its unique minimizer. Any interior stationary solution must satisfy

$$\frac{d Q^U(c_n^*)/Q_{2,1}^*(c_n^*)}{dc_n} = 0.$$

After simplifying, this condition is equivalent to

$$\begin{aligned} \sqrt{*_{2,1}}(c_n^*) &= \frac{\sqrt{(n-1)(a-2c_n^*)((n-1)a-2(5n+3)c_n^*)}}{2(n-1)(3n+1)c_n^* - (n-1)^2a} \\ &= \frac{2(n-1)(3n+1)c_n^* - (n-1)^2a}{3n+1}. \end{aligned} \quad (2.46)$$

Equation (2.46) is quadratic in c_n^* , and its unique non-negative solution is $c_n^* = \left(\frac{n+\sqrt{\frac{n(n+1)}{2}}}{3n+1}\right)a \in [\frac{a}{2}, a]$, where $c_n^* \geq \frac{a}{2}$ is equivalent to $2n \geq 0$. Hence, we conclude that this is the unique minimizer of $(RWCP_{2,1})$. Namely, that the candidate instance from equations (2.41)-(2.43) is the optimal solution to problem $(RWCP_{2,1})$. ■

Let us emphasize again that, a priori, it is not clear which case of $l \in \{1, \dots, n\}$ attains the worst case performance of uniform co-payments, in maximizing market consumption in this setting. We will show now that it must be attained for the case $l = 2$. In other words, we will show that the worst case instance must be such that the first firm that receives a positive co-payment, in the optimal co-payment allocation, is the firm $l = 2$.

Theorem 6. *For any given number of firms in the market $n \geq 4$, and demand parameters $a > 0$, $b > 0$, we have that*

$$WC(2, n) \leq WC(l, n), \text{ for each } l \in \{1\} \cup \{3, \dots, n\}. \quad (2.47)$$

Proof. The proof outline is the following. Propositions 12 and 13 in Section 2.6 show equation (2.47), for the cases $l = 1$ and $l = 3$, respectively. Then, we will use the lower bound on $WC(l, n)$, for any $n \geq 2$, and for each $l \in \{2, \dots, n\}$, from Proposition 9, to show equation (2.47) for the cases $l \in \{4, \dots, n\}$.

In more details, Proposition 12 shows that, for any optimal solution to problem (WCP_1) , constraint (2.37) must be binding. Recall, from the equivalent consistency constraint (2.30), that this corresponds to y_1^* attaining its lower bound when $l = 1$, namely $y_1^* = 0$, or equivalently to y_2^* attaining its upper bound when $l = 2$, namely $y_2^* = (c_2 - c_1)/2$. In other words, it shows that the worst case instance for $l = 1$ must lie in the boundary between the cases $l = 1$ and $l = 2$.

Similarly, in the proof of Proposition 13, for any given $n \geq 3$, $a > 0$, $b > 0$, the value of $WC(3, n)$ is lower bounded by $\frac{2 + \sqrt{2 + 2/n}}{4} \geq WC(2, n)$, where the inequality follows from the fact that the right hand side is attained by the candidate instance for the case $l = 2$ from equations (2.41)-(2.43).

Additionally, Proposition 9 below provides a lower bound on $WC(l, n)$, for any $n \geq 2$, and for each $l \in \{2, \dots, n\}$.

To conclude, note that for any given $n \geq 4$, $a > 0$, $b > 0$ and for each $l \in \{4, \dots, n\}$, we have that

$$WC(l, n) \geq \frac{2nl - 2n + 2l - 2}{2nl - n + l - 1} \geq \frac{6(n+1)}{7n+3} \geq \frac{2 + \sqrt{2 + 2/n}}{4} \geq WC(2, n),$$

where the first inequality follows from Proposition 9, the second inequality follows from the left hand side being increasing in l (the numerator increases faster than the denominator), and taking $l = 4$. The third inequality holds for any $n \geq 1$. The last inequality follows from the fact that the left hand side is attained by the candidate instance from equations (2.41)-(2.43), for the case $l = 2$. This completes the proof of inequality (2.47). ■

The next result provides a parametric lower bound, based on linear programming, on the worst case performance of uniform co-payments in maximizing the market consumption in this setting, as a function of the number of firms in the market n , and the index l of the first firm that receives a positive co-payment in the optimal co-payment allocation.

Proposition 9. *For any given number of firms in the market $n \geq 2$, demand param-*

eters $a > 0$, $b > 0$, and for each $l \in \{2, \dots, n\}$, it must be the case that

$$WC(l, n) \geq \frac{2nl - 2n + 2l - 2}{2nl - n + l - 1}.$$

Proof. The proof structure is the following. We will consider the mathematical programming relaxation of problem (WCP_l) from Lemma 13 in Section 2.6, denoted by (LBP_l) . Its optimal solution provides a lower bound on $WC(l, n)$, for any $n \geq 2$, and for each $l \in \{2, \dots, n\}$. We will reformulate this relaxation as a linear program, and we will use strong duality to obtain its optimal objective value in closed form.

The relaxation (LBP_l) is a linear fractional program. Hence, from Charnes and Cooper (1962), it follows that by defining the transformation

$$x_i = \frac{c_i}{2na - 2 \sum_{i=1}^{l-2} c_i - (n-l+3)c_{l-1} - \sum_{i=l}^n c_i} \text{ for each } i \in \{1, \dots, n\}, \quad (2.48)$$

and

$$t = \frac{1}{2na - 2 \sum_{i=1}^{l-2} c_i - (n-l+3)c_{l-1} - \sum_{i=l}^n c_i}, \quad (2.49)$$

the relaxation (LBP_l) is equivalent to the following linear program

$$\begin{aligned} \min_{t, \mathbf{x}} \quad & nx_n - \sum_{i=1}^n x_i \\ \text{s.t.} \quad & 0 \leq x_2 \end{aligned} \quad (2.50)$$

$$x_i \leq x_{i+1} \text{ for each } i \in \{2, \dots, n-1\} \quad (2.51)$$

$$(LP_l) \quad x_n \leq at \quad (2.52)$$

$$(n+1)x_n - \sum_{i=1}^n x_i - at \geq 0. \quad (2.53)$$

$$2nat - 2 \sum_{i=1}^{l-2} x_i - (n-l+3)x_{l-1} - \sum_{i=l}^n x_i = 1 \quad (2.54)$$

$$t \geq 0. \quad (2.55)$$

Note that, for simplicity, and without loss of generality, we have dropped the constant $2(n+1)$ from the objective value of problem (LP_l) . Similarly, we have

dropped the dummy variable $x_1 = 0$, and replaced it with the equivalent constraint $x_2 \geq 0$. The dual of problem (LP_l) is

$$\begin{aligned}
& \max_{\lambda, \gamma, \mathbf{u}} \quad \lambda \\
& \text{s.t.} \quad u_2 + \gamma - 2\lambda \leq -1 \tag{2.56} \\
& \quad \quad -u_{i-1} + u_i + \gamma - 2\lambda = -1 \text{ for each } i \in \{3, \dots, l-2\} \tag{2.57} \\
(DLP_l) \quad & \quad \quad -u_{l-2} + u_{l-1} + \gamma - (n-l+3)\lambda = -1 \tag{2.58} \\
& \quad \quad -u_{i-1} + u_i + \gamma - \lambda = -1 \text{ for each } i \in \{l, \dots, n-1\} \tag{2.59} \\
& \quad \quad -u_{n-1} + u_n - n\gamma - \lambda = n-1 \tag{2.60} \\
& \quad \quad -au_n + a\gamma + 2na\lambda \leq 0 \tag{2.61} \\
& \quad \quad \gamma \leq 0. \tag{2.62} \\
& \quad \quad u_i \leq 0 \text{ for each } i \in \{2, \dots, n-1\}. \tag{2.63}
\end{aligned}$$

We will now show that for any $n \geq 2$, $a > 0$, $b > 0$, and for each $l \in \{2, \dots, n\}$, the solution

$$c_i = 0 \text{ for each } i \in \{1, \dots, l-1\}, \quad c_i = \frac{a}{l} \text{ for each } i \in \{l, \dots, n\}, \tag{2.64}$$

is optimal for problem (LBP_l) . It is straightforward to check that this solution is feasible for problem (LBP_l) , and that it attains an objective value of $\frac{2nl-2n+2l-2}{2nl-n+l-1}$.

From the Charnes and Cooper transformation given in equations (2.48) and (2.49), it follows that the associated solution to the linear program (LP_l) is

$$x_i = 0 \text{ for each } i \in \{1, \dots, l-1\}, \quad x_i = \frac{1}{2nl-n+l-1} \text{ for each } i \in \{l, \dots, n\}, \tag{2.65}$$

$$t = \frac{l}{(2nl-n+l-1)a}. \tag{2.66}$$

This solution is primal feasible and attains an objective value of $\frac{(l-1)}{2nl-n+l-1}$. Recall that, without loss of generality, we have dropped the constant $2(n+1)$ from the

objective function of problem (LP_l) .

On the other hand, we can also check that the following solution

$$\lambda = \frac{(l-1)}{2nl - n + l - 1}, \quad \gamma = -2n\lambda, \quad (2.67)$$

$$u_i = (l-i-1)\gamma - (n+l-2i-1)\lambda + l-i-1 \text{ for each } i \in \{2, \dots, l-2\}, \quad (2.68)$$

$$u_{l-1} = 0, \quad u_i = -(i+1)\gamma - (n-i)\lambda - i \text{ for each } i \in \{l, \dots, n-1\}, \quad u_n = 0, \quad (2.69)$$

is dual feasible for problem (DLP_l) , and it attains the same objective value $\frac{(l-1)}{2nl-n+l-1}$.

For completeness, we prove this fact in Lemma 18 in Appendix A.2.

Hence, from strong duality in linear programming, it follows that the solutions (2.65)-(2.66), and (2.67)-(2.69), are primal and dual optimal, respectively, see for example Bertsimas and Tsitsiklis (1997). Therefore, the associated solution (2.64) is optimal for problem (LBP_l) , and $\frac{2nl-2n+2l-2}{2nl-n+l-1}$ is a lower bound for $WC(l, n)$ for any $n \geq 2$, $a > 0$, $b > 0$, and for each $l \in \{2, \dots, n\}$. ■

2.6 Auxiliary Results

In this section, we present several auxiliary results that are important to show the main results in the previous section. In particular, lemma 11 below provides a relaxation of problem (WCP_2) , denoted by $(RWCP_2)$.

Lemma 11. *For any given number of firms in the market $n \geq 3$, demand parameter $a > 0$, and budget B , problem $(RWCP_2)$ below is a mathematical programming relaxation of problem (WCP_2) , whose optimal objective value provides a lower bound on*

$WC(2, n)$

$$\begin{aligned} \min_{B, c} \quad & \frac{Q^U(c)}{Q_2^*(B, c)} \\ \text{s.t.} \quad & c_1 = 0 \end{aligned} \tag{2.70}$$

$$c_i \leq c_{i+1}, \text{ for each } i \in \{1, \dots, n-1\} \tag{2.71}$$

$$(RWCP_2) \quad c_n \leq a \tag{2.72}$$

$$(n+1)c_n - \sum_{i=1}^n c_i - a \geq 0. \tag{2.73}$$

Proof. We ignore constraints (2.36) and (2.37) from problem (WCP_2) . Additionally, from $c_n \geq c_i$ for each $i \in \{1, \dots, n\}$, together with the expression for the budget B in constraint (2.35) from problem (WCP_2) , it follows that constraint (2.38) from problem (WCP_2) is equivalent to $(n+1)c_n - \sum_{i=1}^n c_i - a \geq 0$. To conclude, we replace constraints (2.38) and (2.35) from problem (WCP_2) with this linear inequality. ■

The following proposition shows that solving the relaxation $(RWCP_2)$ is equivalent to solving one of $(n-1)$ one variable optimization problems.

Proposition 10. *Without loss of generality, solving problem $(RWCP_2)$ is equivalent to solving one of the following one variable optimization problems*

$$\begin{aligned} \min_{c_n} \quad & \frac{Q_k^U(c_n)}{Q_{2,k}^*(c_n)} = \frac{4(n+1)kc_n}{(3n+1)a - 2(n-k)c_n + \sqrt{^*_{2,k}(c_n)}} \\ (RWCP_{2,k}) \quad \text{s.t.} \quad & \frac{a}{k+1} \leq c_n \end{aligned} \tag{2.74}$$

$$c_n \leq a, \tag{2.75}$$

for some index $k \in \{1, \dots, n-1\}$.

Proof. Note that for any number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, any optimal solution \mathbf{c}^* to problem $(RWCP_2)$ must satisfy that there exists an index $k \in \{1, \dots, n-1\}$ such that $c_i^* = c_1^*$, for each $i \in \{1, \dots, k\}$, and $c_i^* = c_n^*$, for each $i \in \{k+1, \dots, n\}$. The proof of this statement is identical to the first part of the proof of Proposition 19 in Appendix A.2, and it is therefore omitted.

It follows that, without loss of generality, we can focus on solutions to problem $(RWCP_2)$ with a special structure, which can be parametrized by the number of firms k with their marginal cost equal to c_1^* . Moreover, from Lemma 9, we assume, without loss of generality, that $c_1^* = 0$. Then, in this case, the function $\sqrt{*}_l(B, \mathbf{c})$ in equation (2.31) simplifies to

$$\sqrt{*}_{2,k}(c_n) \equiv ((n-1)((n-1)a^2 + 2(k+1)(4nk + n + 3k)c_n^2 - 4(2nk + n + k)ac_n))^{1/2},$$

where, from Proposition 7, without loss of generality, we have dropped the dependency on the budget B , by replacing it directly by the expression in constraint (2.35) in problem (WCP_l) .

Similarly, the functions $Q_l^*(B, \mathbf{c})$ and $Q^U(\mathbf{c})$ simplify to

$$Q_{2,k}^*(c_n) \equiv \frac{(3n+1)a - 2(n-k)c_n + \sqrt{*}_{2,k}(c_n)}{4(n+1)b}, \text{ and } Q_k^U(c_n) \equiv \frac{kc_n}{b}.$$

This completes the proof. ■

It is straightforward to check that the candidate instance from equations (2.41)-(2.43) is feasible for problem $(RWCP_{2,1})$.

Proposition 11 below allows us to focus on problem $(RWCP_{2,1})$ only, as it shows that any solution to problems $(RWCP_{2,k})$, for any index $k \in \{2, \dots, n-1\}$, must attain a larger objective function. Additionally, Lemma 12 shows that the objective function of the problems $(RWCP_{2,k})$ is quasiconvex. The proofs of these results is given in Appendix A.2.

Proposition 11. *For any given number of firms in the market $n \geq 3$, demand parameters $a > 0$, $b > 0$, solving problem $(RWCP_2)$ is equivalent to solving problem $(RWCP_{2,1})$.*

Lemma 12. *For any given number of firms in the market $n \geq 2$, demand parameters $a > 0$, $b > 0$, and for any index $k \in \{1, \dots, n-1\}$, the objective function of problem $(RWCP_{2,k})$ is quasiconvex in its feasible set.*

Propositions 12 and 13 below show that, in order to identify the worst case instance for the performance of uniform co-payments in maximizing the market consumption in this setting, without loss of generality we can ignore the cases $l = 1$ and $l = 3$, respectively. The proofs of these results are given in Appendix A.2.

Proposition 12. *For any given number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, any optimal solution to problem (WCP_1) must be such that constraint (A.26) is binding. Therefore, $WC(2, n) \leq WC(1, n)$.*

Proposition 13. *For any given number of firms in the market $n \geq 3$, and demand parameters $a > 0$, $b > 0$, it must be the case that $WC(2, n) \leq WC(3, n)$.*

Lemma 13 below provides a relaxation of problem (WCP_l) , which is a fractional linear program.

Lemma 13. *For any given number of firms in the market $n \geq 3$, and demand parameter $a > 0$, problem (LBP_l) below is a mathematical programming relaxation of problem (WCP_l) , whose optimal objective value provides a lower bound on $WC(l, n)$, for each $l \in \{2, \dots, n\}$.*

$$\begin{aligned} \min_{\mathbf{c}} \quad & \frac{2(n+1)(nc_n - \sum_{i=1}^n c_i)}{2na - 2\sum_{i=1}^{l-2} c_i - (n-l+3)c_{l-1} - \sum_{i=l}^n c_i} \\ \text{s.t.} \quad & c_1 = 0 \end{aligned} \tag{2.76}$$

$$c_i \leq c_{i+1}, \text{ for each } i \in \{1, \dots, n-1\} \tag{2.77}$$

$$(LBP_l) \quad c_n \leq a \tag{2.78}$$

$$(n+1)c_n - \sum_{i=1}^n c_i - a \geq 0. \tag{2.79}$$

Proof. By replacing the function $Q_l^*(B, \mathbf{c})$ with its upper bound from Lemma 6, in the objective function of problem (WCP_l) , we obtain a mathematical programming relaxation whose objective function does not depend on the budget B . Additionally, we ignore constraints (2.36) and (2.37) from problem (WCP_l) altogether. Finally, from $c_n \geq c_i$, for each $i \in \{1, \dots, n\}$, together with the expression for the budget B in constraint (2.35) in problem (WCP_l) , it follows that constraint (2.38) from problem

(WCP_l) is equivalent to $(n + 1)c_n - \sum_{i=1}^n c_i - a \geq 0$. We replace constraints (2.38) and (2.35) with this linear inequality. We also drop the variable B , as it does not play a role anymore. ■

2.7 Conclusions

We studied the problem faced by a central planner allocating a budget of co-payment subsidies to Cournot competitors who produce a good. We assume that the firms have constant marginal costs and face linear demand, and that the central planner's objective is to maximize the aggregated market consumption of the good. We characterized the structure of the optimal co-payment allocation policy, showing that it consists of allocating larger co-payments to less efficient firms. We argued that this policy is hard to implement in practice, and therefore we focused on the performance of the more practical, and conceptually simple, uniform co-payments allocation.

We used linear programming duality to show a lower bound on the performance of uniform co-payments in maximizing market consumption for this model. Then, we analyzed a family of non-convex optimization problems to conclude that, for any number of firms in the market $n \geq 2$, the worst case performance of uniform co-payments in maximizing market consumption is $\frac{2+\sqrt{2+2/n}}{4}$. This immediately provides an asymptotically tight bound of $\frac{2+\sqrt{2}}{4} \approx 85.31\%$. Hence, we conclude that this bounded, and relatively small, loss of efficiency should be weighted against the practical benefits of the uniform co-payments policy. Such practical benefits include that uniform co-payments are simple to communicate and control, as well as being more stable to misspecifications of the marginal costs compared to the optimal co-payments allocation.

Future research on this topic should study whether the worst case bounds for the performance of uniform co-payments presented in this chapter hold for a larger family of instances, as well as whether there are generalized worst case bounds that show that uniform co-payments have a guaranteed performance for more general market competition models.

Chapter 3

A Continuous Knapsack Problem with Separable Convex Utilities: Approximation Algorithms and Applications

3.1 Introduction

In this chapter we study a continuous knapsack problem with separable convex utilities. We show that the problem is *NP*-hard, and we provide two simple algorithms that have worst-case performance guarantees. We consider as an application a novel subsidy allocation problem in the presence of market competition, subject to a budget constraint and upper bounds on the amount allocated to each firm, where the objective is to minimize the market price of a good.

Specifically, we study a continuous knapsack problem, where the objective is to maximize the sum of separable convex utility functions. We denote this problem by (CKP). Beyond general methods for concave minimization, see for example Benson (1995), there is not much literature on this class of problems. An exception is More and Vavasis (1991), and their algorithm to find local minima. A comprehensive re-

view of the related nonlinear knapsack problem literature is presented in Bretthauer and Shetty (2002); however, in most cases, the objective function considered in this literature is concave. On the other hand, for any given tolerance $\epsilon > 0$, a fully polynomial time approximation scheme (FPTAS) is an algorithm that generates a solution which is within a factor $(1 - \epsilon)$ of being optimal, while the running time of the algorithm is polynomial in the problem size and $1/\epsilon$. Burke et al. (2008) provide a tailored FPTAS for a minimization variant of a continuous knapsack problem, in the context of allocating procurement to suppliers. The knapsack problem we study here is a maximization problem, hence the results from Burke et al. (2008) do not apply. Finally, Halman et al. (2008) develop a general purpose FPTAS for a class of stochastic dynamic programs, which applies to general nonlinear knapsack problems. In contrast, our main goal in this chapter is to study the performance of simple algorithms for problem (CKP), as well as to introduce a novel application of continuous knapsack problems into a subsidy allocation model in the presence of endogenous market competition.

The main contributions of this chapter are two-fold. First, we develop two algorithms that are computationally and conceptually simple, such that they can be used in practical applications. We show that these algorithms have good worst-case performance guarantees for problem (CKP). Moreover, we identify special settings where these simple policies are actually optimal. Second, we show that problem (CKP) characterizes a novel subsidy allocation problem, and that the simple algorithms that we develop admit a practical interpretation.

3.2 Problem formulation

Consider n items indexed by $i \in \{1, \dots, n\}$. For each i , let x_i be the non-negative quantity of item i , and let $f_i(x_i)$ be the resulting reward. Moreover, $f_i(x_i)$ is assumed to be convex. The quantity of item i cannot exceed a given upper bound u_i , and the total amount of all items is bounded by the capacity of the knapsack, denoted by B . Moreover, both B and u_i are assumed to be integers. We are interested in the

following continuous knapsack problem

$$\begin{aligned}
 & \max F(x) \equiv \sum_{i=1}^n f_i(x_i) \\
 \text{(CKP)} \quad & \text{s.t.} \quad \sum_{i=1}^n x_i \leq B \\
 & \quad \quad 0 \leq x_i \leq u_i \quad \forall i.
 \end{aligned}$$

The objective function is convex over the feasible set, which is a bounded polyhedron. Therefore, the existence of an extreme point optimal solution follows from concave minimization theory, see for example Benson (1995).

The next one is our first result

Proposition 14. *Problem (CKP) is NP-hard.*

Proof. The proof is a reduction from the subset sum problem, which is well known to be NP-complete, see Karp (1972).

Consider an arbitrary instance of the subset sum problem, where given a set of n positive integers $\{u_1, u_2, \dots, u_n\}$, and a positive integer B , the question is if there exists a subset $J \subseteq \{u_1, u_2, \dots, u_n\}$ that sums to B .

Now consider the following instance of problem (CKP): let u_i be the upper bound on x_i for each i , B be the capacity of the knapsack and $f(x_i) = x_i(x_i - u_i) + x_i$ be the convex reward. It follows that this instance of problem (CKP) can be written as

$$\begin{aligned}
 \max \quad & \sum_{i=1}^n x_i - \sum_{i=1}^n x_i(u_i - x_i) \\
 \text{s.t.} \quad & \sum_{i=1}^n x_i \leq B \\
 & \quad \quad 0 \leq x_i \leq u_i \quad \forall i \in \{1, \dots, n\}.
 \end{aligned}$$

Note that B is an upper bound on the optimal objective value of this problem. Moreover, this upper bound is attained if and only if there exists a subset $J \subseteq \{u_1, u_2, \dots, u_n\}$ that sums to B .

Hence, if we can solve problem (CKP) in polynomial time, it follows that we can solve the subset sum problem in polynomial time. ■

The proof of Proposition 14 is in the same spirit of Sahni (1974), which shows the *NP*-hardness of non-convex quadratic programming among other problems.

We now make a couple of remarks that will make the exposition clearer.

Remark 2. *There is no loss of generality in assuming that, for each $i \in \{1, \dots, n\}$, the functions $f_i(x_i)$ are positive and non decreasing.*

Specifically, we can pre-process the problem data replacing $f_i(x_i)$ by the amount $\max\{f_i(x_i), f_i(0)\}$, for each i and x_i , obtaining non decreasing functions without changing the problem. Similarly, by adding a constant $K > \min_i\{f_i(0)\}$ to each of the functions $f_i(x_i)$ we obtain positive functions.

Remark 3. *There is no loss of generality in assuming that, for each $i \in \{1, \dots, n\}$, $u_i \leq B$.*

Specifically, if any upper bound u_i is larger than the capacity B , then it follows that any feasible solution will allocate at most B to item i . Hence, we can pre-process the data and replace u_i by $\min\{B, u_i\}$, for each i , without changing the problem.

3.2.1 A simple 1/2-approximation algorithm

We next describe a 1/2-approximation algorithm for problem (CKP). Specifically, we will show that intuitive ideas perform well in this model. Namely, the best solution between (i) allocating the capacity greedily to the items with the *fastest rate of increase* in their utility function, and (ii) allocating the capacity greedily to the items with the *largest absolute increase* in their utility function, attains an objective value that is at most half the value of the optimal objective value.

This algorithm is a generalization of the well known 1/2-approximation algorithm for the 0/1 knapsack problem. The latter is attained by the best solution between greedily picking the objects by decreasing ratio of profit to size, and picking the most profitable object, see for example Williamson and Shmoys (2011).

Consider first idea (i). We denote the resulting solution by \mathbf{x}^{rate} . Essentially, \mathbf{x}^{rate} is the result of a greedy procedure with respect to $\frac{f_i(u_i)-f_i(0)}{u_i}$, which is the rate of increase in the utility function of item i , assuming that x_i is set to its upper bound.

Algorithm 1 Compute \mathbf{x}^{rate}

$\mathbf{x}^{\text{rate}} \leftarrow \vec{0}$
 Let $\lambda_i = \frac{f_i(u_i)-f_i(0)}{u_i}$, for each i
 Sort indexes by decreasing λ_i
 Find \hat{i} s.t. $\sum_{i=1}^{\hat{i}-1} u_i \leq B$ and $\sum_{i=1}^{\hat{i}} u_i > B$
 $x_i^{\text{rate}} \leftarrow u_i$, for each $i \leq \hat{i}$
 $x_{\hat{i}}^{\text{rate}} \leftarrow B - \sum_{i=1}^{\hat{i}-1} u_i$

On the other hand, consider idea (ii). We denote the resulting feasible solution by \mathbf{x}^{max} . Essentially, \mathbf{x}^{max} is the result of a greedy procedure with respect to $f_i(\min(u_i, \tilde{B}))$, which is the absolute increase in the utility function of item i , when allocating the minimum between the remaining capacity \tilde{B} , and its upper bound. In case of a tie, $f_i(0)$ is used as a tie-breaker.

Algorithm 2 Compute \mathbf{x}^{max}

$\tilde{B} \leftarrow B$
 $\mathbf{x}^{\text{max}} \leftarrow \vec{0}$
while $\tilde{B} > 0$ **do**
 Let $S_1 = \{i \mid f_i(\min(u_i, \tilde{B})) \geq f_j(\min(u_j, \tilde{B}))\}$, for each $j : \mathbf{x}_j^{\text{max}} = 0\}$
 Let $S_2 = \{i \in S_1 \mid f_i(0) \leq f_j(0), \text{ for each } j \in S_1\}$
 Select $i \in S_2$
 $\tilde{B} \leftarrow \tilde{B} - \min(u_i, \tilde{B})$
 $x_i^{\text{max}} \leftarrow \min(u_i, \tilde{B})$
end while

It is not hard to see that each algorithm, considered separately, can be made to perform arbitrarily bad. Examples drawn from a 0/1 knapsack problem are sufficient.

In order to show a worst-case performance guarantee for problem (CKP), we need an upper bound on its optimal objective value, as provided in the following proposition.

Proposition 15. Let \mathbf{x}^* be an optimal solution to problem (CKP). Algorithm 1 provides the following upper bound,

$$F(\mathbf{x}^*) \leq F(\mathbf{x}^{\text{rate}}) + f_{\hat{i}}(0) - f_{\hat{i}}(x_{\hat{i}}^{\text{rate}}) + \frac{f_{\hat{i}}(u_{\hat{i}}) - f_{\hat{i}}(0)}{u_{\hat{i}}} x_{\hat{i}}^{\text{rate}}.$$

Where $F(x) = \sum_{i=1}^n f_i(x_i)$.

Proof. Lets relax the knapsack constraint in problem (CKP) with an associated Lagrange multiplier λ , to obtain the following relaxed optimization problem,

$$\begin{aligned} \max \quad & \lambda B + \sum_{i=1}^n (f_i(x_i) - \lambda x_i) \\ \text{s.t.} \quad & 0 \leq x_i \leq u_i \quad \forall i. \end{aligned}$$

The resulting problem is separable in the variables x_i . Specifically, for each variable it maximizes a convex function over a closed interval. It follows that the optimal solution is attained at one of the extremes of the interval. For any fixed Lagrangian multiplier λ , let $G(\lambda)$ denote the optimal objective value of the relaxed problem. Namely, $G(\lambda) = \lambda B + \sum_{i=1}^n \max(f_i(0), f_i(u_i) - \lambda u_i)$. From duality theory, it follows that, for any $\lambda \geq 0$, $G(\lambda)$ is an upper bound for the optimal objective value of problem (CKP), see for example Boyd and Vandenberghe (2007). Moreover, the best possible upper bound can be computed from the following program,

$$\begin{aligned} \min_{\lambda} \quad & G(\lambda) = \lambda B + \sum_{i=1}^n \max(f_i(0), f_i(u_i) - \lambda u_i) \\ \text{s.t.} \quad & \lambda \geq 0. \end{aligned}$$

The objective function of this problem is piecewise linear and convex. Hence, it can be solved by trying out the values of λ where the slope of the objective function changes. In particular, Algorithm 1 solves this problem. The optimal Lagrange multiplier is $\lambda_{\hat{i}} = \frac{f_{\hat{i}}(u_{\hat{i}}) - f_{\hat{i}}(0)}{u_{\hat{i}}}$, where \hat{i} was defined in Algorithm 1 as being such that $\sum_{i=1}^{\hat{i}-1} u_i \leq B$ and $\sum_{i=1}^{\hat{i}} u_i > B$.

Plugging in the optimal Lagrange multiplier $\lambda_{\hat{i}}$ in $G(\lambda)$, results in the best possible

upper bound from this relaxation. Without loss of generality, set $x_i = 0$, then

$$\begin{aligned}
G(\lambda_i) &= \sum_{i=1}^{\hat{i}-1} f_i(u_i) + \sum_{i=\hat{i}}^n f_i(0) + \lambda_i \underbrace{\left(B - \sum_{i=1}^{\hat{i}-1} u_i \right)}_{=x_i^{\text{rate}}} \\
&= F(\mathbf{x}^{\text{rate}}) + f_i(0) - f_i(x_i^{\text{rate}}) \\
&\quad + \underbrace{\frac{f_i(u_i) - f_i(0)}{u_i}}_{\lambda_i} x_i^{\text{rate}}.
\end{aligned}$$

The second equality follows from adding and subtracting the term $f_i(x_i^{\text{rate}})$. \blacksquare

The next theorem is the main result in this section.

Theorem 7. *Let \mathbf{x}^* be an optimal solution to problem (CKP). Let \mathbf{x}^{rate} be the solution computed by Algorithm 1, and \mathbf{x}^{max} be the solution computed by Algorithm 2. Then*

$$\frac{\max(F(\mathbf{x}^{\text{rate}}), F(\mathbf{x}^{\text{max}}))}{F(\mathbf{x}^*)} \geq \frac{1}{2}.$$

Proof. To make the notation clearer, define $\bar{f} = \max_i \{f_i(u_i)\}$. Note that,

$$\begin{aligned}
&\frac{\max(F(\mathbf{x}^{\text{rate}}), F(\mathbf{x}^{\text{max}}))}{F(\mathbf{x}^*)} \\
&\geq \frac{\max(F(\mathbf{x}^{\text{rate}}), \bar{f})}{F(\mathbf{x}^*)} \\
&\geq \frac{\max(F(\mathbf{x}^{\text{rate}}), \bar{f})}{F(\mathbf{x}^{\text{rate}}) + \underbrace{f_i(0) - f_i(x_i^{\text{rate}})}_{\leq 0} + \frac{f_i(u_i) - f_i(0)}{u_i} x_i^{\text{rate}}} \\
&\geq \frac{\max(F(\mathbf{x}^{\text{rate}}), \bar{f})}{F(\mathbf{x}^{\text{rate}}) + \frac{f_i(u_i)}{u_i} x_i^{\text{rate}}} \\
&\geq \frac{1}{1 + \underbrace{\frac{f_i(u_i)}{\bar{f}}}_{\leq 1} \underbrace{\frac{x_i^{\text{rate}}}{u_i}}_{\leq 1}} \\
&\geq \frac{1}{2}.
\end{aligned}$$

The first inequality follows from Remark 2, and the definitions of \bar{f} and \mathbf{x}^{\max} . Specifically, they imply $\bar{f} \leq F(\mathbf{x}^{\max})$. The second inequality follows from Proposition 15, while the third inequality follows from Remark 2. The fourth inequality follows from the definition of \bar{f} . ■

To conclude this section, the following lemma identifies three cases that can be solved in polynomial time. Specifically, Algorithm 1 solves problem (CKP) exactly if the utility functions of each item are affine, namely $f_i(x_i) = a_i + b_i x_i$ for each i , for some $a_i > 0$, $b_i > 0$. Algorithm 1 also solves problem (CKP) exactly if any number of items, ordered by fastest rate of increase in their utility function, fill the knapsack exactly. Additionally, if the upper bounds are uniform, then problem (CKP) can be solved by applying Algorithm 1 n times.

Lemma 14. *If the functions $f_i(x_i)$ are affine, or if the first $(\hat{i} - 1)$ indexes sorted by decreasing value of $\lambda_i = \frac{f_i(u_i) - f_i(0)}{u_i}$ fill the knapsack exactly, for some value of \hat{i} , then \mathbf{x}^{rate} is the optimal solution to problem (CKP).*

On the other hand, if the upper bounds on the allocation to each index are uniform, namely $u_i = u$ for each i , then problem (CKP) can be solved in polynomial time.

Proof. The first statement in the lemma is a direct consequence of Proposition 15. Specifically, if $x_i^{\text{rate}} = 0$, then it follows that $F(\mathbf{x}^*) \leq F(\mathbf{x}^{\text{rate}})$, hence \mathbf{x}^{rate} is optimal. This holds in both cases in the first statement of the lemma.

Assume now that $u_i = u$ for each i . Each extreme point solution is characterized by one fractional variable x_i , which gets an allocation $(B - \lfloor \frac{B}{u} \rfloor u)$, while $\lfloor \frac{B}{u} \rfloor$ other variables get an allocation u , and all the remaining variables get no allocation. From the first statement in the lemma, it follows that we can try each variable as the fractional variable, allocating $(B - \lfloor \frac{B}{u} \rfloor u)$ to it; and then use Algorithm 1 to optimally solve the problem of allocating the remaining capacity $\lfloor \frac{B}{u} \rfloor u$, among the remaining variables. This follows because the first $\lfloor \frac{B}{u} \rfloor$ indexes sorted by decreasing value of $\lambda_i = \frac{f_i(u_i) - f_i(0)}{u_i}$ fill this modified knapsack exactly. In conclusion, in this case problem (CKP) can be solved by running Algorithm 1 n times. ■

3.2.2 An $(1 - e^{-1})$ -approximation algorithm

In this section we present an $(1 - e^{-1})$ -approximation algorithm for problem (CKP), where $(1 - e^{-1}) \approx .632$. We denote the resulting solution by \mathbf{x}^{seq} . In this algorithm we enumerate all the solutions that allocate capacity to 3 items or less. Then, for each of these solutions we allocate the remaining capacity, if any, greedily to the remaining indexes with the *fastest rate of increase* in their utility function. In that sense, Algorithm 3 is an extension of Algorithm 1. It captures that among the two simple rules we have considered, the fastest rate of increase rule is the most powerful. Specifically, it is enough to consider all solutions that allocate capacity to 3 items or less, to rule out all the cases where the largest absolute increase rule was important to define the worst-case guarantee.

To the best of our knowledge, this is the first time that these ideas have been used in a continuous optimization setting, like our continuous knapsack problem (CKP). Similar ideas have been used before in inherently discrete settings, such as solving a budgeted maximum coverage problem in Khuller et al. (1999), and maximizing a submodular set function subject to a knapsack constraint in Sviridenko (2004).

Algorithm 3 Compute \mathbf{x}^{seq}

Consider all sequences of 3 different indexes and allocate the capacity in this order

```

Let  $\tilde{B}$  be the remaining capacity
for Each sequence do
  if  $\tilde{B} > 0$  then
    Apply Algorithm 1 to the rest of the indexes with capacity  $\tilde{B}$ 
  end if
end for

```

Theorem 8. *Let \mathbf{x}^* be an optimal solution to problem (CKP). Let \mathbf{x}^{seq} be the solution computed by Algorithm 3. Then,*

$$\frac{F(\mathbf{x}^{\text{seq}})}{F(\mathbf{x}^*)} \geq (1 - e^{-1}).$$

Proof. If \mathbf{x}^* allocates all the capacity to 3 or less indexes, then we must have $\mathbf{x}^{\text{seq}} = \mathbf{x}^*$

by enumeration. Therefore, assume that \mathbf{x}^* allocates capacity to 4 or more firms. Let

$$\bar{S} \equiv \{i : x_i^* = u_i\} = \{i_1, i_2, \dots, i_{|S|}\} \quad (3.1)$$

be the set of indexes for which their allocation attains their upper bound u_i . Assume, without loss of generality, that \bar{S} is ordered such that $f_{i_1}(u_{i_1}) \geq f_{i_2}(u_{i_2}) \geq \dots \geq f_{i_{|S|}}(u_{i_{|S|}})$. Let $Y \subset \bar{S}$ be the set including the first 3 indexes in \bar{S} . Namely, $Y = \{i_1, i_2, i_3\}$. Let B_Y be the remaining capacity, after allocating the capacity to the indexes in Y in this order. Define $\hat{\mathbf{x}}$ to be such that,

$$\hat{x}_i = \begin{cases} u_i & \text{if } i \in Y \\ 0 & \text{otherwise.} \end{cases}$$

Let \mathbf{x}^{seq} be the solution generated by completing $\hat{\mathbf{x}}$ using Algorithm 1, considering all indexes except those in Y , and an initial capacity B_Y . In fact, Algorithm 3 considers \mathbf{x}^{seq} as one of its candidate solutions, therefore it outputs a solution at least as good. We will show that \mathbf{x}^{seq} has a worst-case performance guarantee of $(1 - e^{-1})$ for problem (CKP).

Define $H(\mathbf{x}) = F(\mathbf{x}) - F(\hat{\mathbf{x}})$. Assume, without loss of generality, that indexes are numbered such that $1 = i_1, 2 = i_2, 3 = i_3$, and then in decreasing order according to λ_i , for each $i \in I \setminus Y$. Let \mathbf{x}^i be such that

$$\mathbf{x}_j^i = \begin{cases} u_j & \text{if } j \leq i \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\mathbf{x}^3 = \hat{\mathbf{x}}$. Additionally, let $B_i = \sum_{j=1}^i u_j$. Let $\rho_k = \lambda_{i+1}$, for each $k = B_i + 1, \dots, B_{i+1}$, $i \geq 3$, and $\rho_k = 0$, for each $k \leq B_Y$. From Proposition 15 it follows

that, for every $i \geq 3$

$$\begin{aligned}
H(\mathbf{x}^*) &\equiv F(\mathbf{x}^*) - F(\hat{\mathbf{x}}) \\
&\leq \sum_{j=1}^i f_j(u_j) + \sum_{j=i+1}^n f_j(0) \\
&\quad + \lambda_{i+1} (B - B_i) - F(\hat{\mathbf{x}}) \\
&= F(\mathbf{x}^i) + \lambda_{i+1} (B - B_i) - F(\hat{\mathbf{x}}) \\
&= H(\mathbf{x}^i) + \lambda_{i+1} (B - B_i) \\
&\leq H(\mathbf{x}^i) + \lambda_{i+1} (B - B_Y). \tag{3.2}
\end{aligned}$$

Let \hat{i} be the last index with a positive allocation in \mathbf{x}^{seq} . Note that $\hat{i} \geq 4$, therefore $B_{\hat{i}} > B_Y$. Additionally, note that

$$H(\mathbf{x}^i) = \sum_{j=4}^i (f_j(u_j) - f_j(0)) = \sum_{j=4}^i \lambda_j u_j = \sum_{k=1}^{B_i} \rho_k \quad \forall i \geq 3. \tag{3.3}$$

Where the first equality follows from the definition of \mathbf{x}^i and $H(\mathbf{x})$. The second equality follows from the definition of λ_j . The last equality follows from the definition of ρ_k .

Hence,

$$\begin{aligned}
&\min_{i=1, \dots, \hat{i}-1} \{H(\mathbf{x}^i) + \lambda_{i+1} B_Y\} \\
&= \min_{i=1, \dots, \hat{i}} \left\{ \sum_{k=1}^{B_i} \rho_k + \rho_{B_{i+1}} B_Y \right\} \tag{3.4}
\end{aligned}$$

$$= \min_{s=1, \dots, B_{\hat{i}}} \left\{ \sum_{k=1}^{s-1} \rho_k + \rho_s B_Y \right\}. \tag{3.5}$$

The first equality follows from equation (3.3), and the second equality follows because we are only adding non-negative terms, therefore the minimizer does not change.

It follows that,

$$\begin{aligned}
\frac{H(\mathbf{x}^{\hat{i}})}{H(\mathbf{x}^*)} &\geq \frac{H(\mathbf{x}^{\hat{i}})}{\min_{1, \dots, \hat{i}} \{H(\mathbf{x}^i) + \lambda_{i+1} (B - B_Y)\}} \\
&= \frac{\sum_{k=1}^{B_i} \rho_k}{\min_{s=1, \dots, B_i} \left\{ \sum_{k=1}^{s-1} \rho_k + \rho_s B_Y \right\}} \\
&\geq 1 - \left(1 - \frac{1}{B_Y}\right)^{B_i} \\
&> 1 - e^{-\frac{B_i}{B_Y}} \\
&> 1 - e^{-1}. \tag{3.6}
\end{aligned}$$

The first inequality follows from equation (3.2). The first equality follows from equations (3.3) and (3.4). The second and third inequalities are due to Wolsey, where it is required that both B_Y and B_i are integers, see Wolsey (1982). The last inequality follows from $B_i > B_Y$.

Finally, we conclude that,

$$\begin{aligned}
F(\mathbf{x}^{\text{seq}}) &= H(\mathbf{x}^{\text{seq}}) + F(\hat{\mathbf{x}}) \\
&= F(\hat{\mathbf{x}}) + H(\mathbf{x}^{\hat{i}}) - \left(H(\mathbf{x}^{\hat{i}}) - H(\mathbf{x}^{\text{seq}})\right) \\
&= F(\hat{\mathbf{x}}) + H(\mathbf{x}^{\hat{i}}) - \left(F(\mathbf{x}^{\hat{i}}) - F(\mathbf{x}^{\text{seq}})\right) \\
&> F(\hat{\mathbf{x}}) + (1 - e^{-1}) H(\mathbf{x}^*) \\
&\quad - \left(F(\mathbf{x}^{\hat{i}}) - F(\mathbf{x}^{\text{seq}})\right) \\
&= (1 - e^{-1}) F(\mathbf{x}^*) + e^{-1} F(\hat{\mathbf{x}}) \\
&\quad - (f_i(u_i) - f_i(x_i)) \\
&> (1 - e^{-1}) F(\mathbf{x}^*) + \frac{1}{3} F(\hat{\mathbf{x}}) - f_i(u_i) \\
&> (1 - e^{-1}) F(\mathbf{x}^*).
\end{aligned}$$

The first and third equalities follow from the definition of $H(\mathbf{x})$. The first inequality follows from equation (3.6). The last inequality follows from the definition of Y , and

the order of set \bar{S} . Specifically,

$$F(\hat{\mathbf{x}}) = \sum_{i=1}^3 f_i(u_i) + \sum_{i=4}^n f_i(0) > \sum_{i=1}^3 f_i(u_i) > 3f_{\hat{i}}(u_{\hat{i}}).$$

Where the last inequality follows from $\frac{1}{3}F(\hat{\mathbf{x}}) - f_{\hat{i}}(u_{\hat{i}}) < 0$. Which follows from $\hat{i} \geq 4$, and the order of the set \bar{S} . ■

3.3 Application: Allocating technology subsidies to minimize a good's market price

We consider the problem faced by a central planner with the goal of increasing the consumption of a given good, due to the positive societal externalities that it generates. Concrete examples of such goods are vaccines and infectious disease treatments. In order to achieve this goal, she can allocate a given budget in the form of lump sum subsidies among heterogeneous competing firms that produce the good. The introduction of subsidies in the market will induce a demand increase. We assume that the firms do not have the installed capacity to serve all the induced demand, therefore capacity is scarce. We model this by assuming that the firm's marginal costs are increasing. Furthermore, in our model, it is in the best interest of each firm to invest the subsidy to improve the efficiency of its production process, reducing its marginal costs. Therefore, we refer to them as *technology subsidies*. We model the central planner's objective as minimizing the good's market price, therefore increasing its consumption.

Allocating subsidies to producers, rather than to consumers, makes sense when the coordination costs associated with paying to each consumer are larger than the additional benefits generated by impacting consumers directly. This is frequently the case when subsidizing infectious disease treatments in developing countries. One example is the budget of \$1.5 billion allocated as lump sum subsidies to producers of the pneumococcal vaccine in 2007, by the Global Alliance for Vaccines and Immunization

(GAVI) (see Snyder et al. (2011)).

To model the market equilibrium, we assume that the market is composed by $n \geq 2$ competing firms. Firms are profit maximizers, and engage in quantity competition, with a linear inverse demand function $P(Q) = \alpha - \frac{1}{\mu}Q$, where $\mu > 0$, and $Q = \sum_{i=1}^n q_i$ is the total output produced by all firms at equilibrium. Furthermore, we denote by x_i the technology investment that each firm incurs, in order to become more efficient. Specifically, we assume that firms have a linear marginal cost function of their output q_i , $MC_i = \bar{g}_i(x_i)q_i$, for each i . Note that $\bar{g}_i(x_i) > 0$ is the parameter of the marginal cost function, which captures firm i 's efficiency. Specifically, the smaller the value of $\bar{g}_i(x_i)$, the more efficient the firm is. A larger technology investment x_i reduces the value of $\bar{g}_i(x_i)$, at a cost $c_i(x_i)$, with a maximum amount that can be borrowed \bar{x}_i . The function $c_i(x_i)$ models the financing cost of firm i . Note that the firms in our model are heterogeneous. The linear demand, and linear marginal costs assumptions are a good approximation, which allow us to obtain closed-form expressions for equilibrium outcomes, and to get insights on the subsidy allocation. Similar assumptions are frequently made by researchers in order to get different insights (see for example Deo and Corbett (2009)).

To simplify the exposition, define $g_i(x_i) \equiv (\bar{g}_i(x_i) + \frac{1}{\mu})$. Adding a constant $\frac{1}{\mu}$ to the marginal cost of each firm will not make a difference in the analysis, therefore we will refer to $g_i(x_i)$ as the marginal cost functions from now on. We make the following assumption on $g_i(x_i)$.

Assumption 1. *Assume that $g_i(x_i)$ are continuous, positive and decreasing, for each $i \in \{1, \dots, n\}$. Moreover, assume $g_i''(x_i)g_i(x_i) \leq 2(g_i'(x_i))^2$ for any technology investment $x_i \geq 0$.*

Assumption 1 implies that $\frac{1}{g_i(x_i)}$ is a convex function of the technology investment x_i . To assume that $g_i(x_i)$ is positive decreasing is natural in our setting. On the one hand, it implies that a technology investment x_i cannot increase the cost of production. On the other hand, it implies that no matter how large a technology investment x_i is, the resulting marginal cost cannot be zero or negative. The latter

condition in Assumption 1 is not very restrictive. It is satisfied by any concave, positive and decreasing function $g_i(x_i)$. It is also satisfied by convex functions, such as $g_i(x_i) = k_1 e^{-k_2 x_i}$, and $g_i(x_i) = k_1 x_i^{-k_2}$ for any $k_2 \geq 1$.

Similarly, we make the following assumption on the financing cost of firm i .

Assumption 2. Assume that $c_i(x_i)$ are continuous, positive, increasing and convex in $[0, \bar{x}_i]$, for each $i \in \{1, \dots, n\}$, where $\lim_{x_i \rightarrow \bar{x}_i} c_i(x_i) = \infty$.

Moreover, assume $\frac{c_i''(x_i)}{c_i'(x_i)} \geq -\frac{2(g_i'(x_i))^2 - g_i''(x_i)g_i(x_i)}{g_i(x_i)g_i'(x_i)} \geq 0$ for any technology investment $x_i \in [0, \bar{x}_i]$.

Assumption 2 states that the cost of borrowing money increases at an increasing rate for each firm, where \bar{x}_i is the maximum amount that can be borrowed. The latter condition in Assumption 2 is technical, it ensures the existence of the market equilibrium. Intuitively, it states that the financing cost of each firm is convex enough for the profit of each firm to be quasi-concave in the technology investment x_i . Examples of pair of functions $(g_i(x_i), c_i(x_i))$, that satisfy Assumptions 1 and 2, include $(k_1 e^{-k_2 x_i}, k_3 e^{k_4 x_i})$, for any $k_4 \geq k_2 \geq 0$, and $(k_1 x_i^{-k_2}, k_3 e^{k_4 x_i})$ for any $k_4 \geq \frac{k_2 - 1}{k_2 + 1} \geq 0$, $k_2 \geq 1$.

Now we characterize the market equilibrium. Let $P(\mathbf{x})$ denote the market price under technology investments $\mathbf{x} \in R^n$. Assuming quantity competition with linear demand allows us to write the following closed form expressions.

Proposition 16. The equilibrium market price, induced by a technology investment vector \mathbf{x} , can be written as

$$P(\mathbf{x}) = \alpha\mu \left(\sum_{i=1}^n \frac{1}{g_i(x_i)} + \mu \right)^{-1}. \quad (3.7)$$

While the market output satisfies $q_i(\mathbf{x}) = P(\mathbf{x})/g_i(x_i)$, for each firm i .

Derivations of similar closed form expressions can be found in the transportation and economics literature, therefore the proof is omitted, see for example Nagurney (1999).

From Proposition 16, it follows that the market equilibrium is only a function of the technology investments x_i . Moreover, let $\Pi_i(\mathbf{x})$ be the profit obtained by firm i at the market equilibrium. Note that the revenue of firm i is the market price times its market output. Similarly, the cost of production of firm i is the integral of its linear marginal cost, from zero to its market output. Finally, we also need to consider the the financing cost of firm i . It follows that firm i 's profit, at the market equilibrium, can be written as $\Pi_i(\mathbf{x}) = P(\mathbf{x})q_i(\mathbf{x}) - g_i(x_i)q_i(\mathbf{x})^2/2 - c_i(x_i)$, where the first term is its revenue, the second term is its production cost, and the third term is its financing cost. Moreover, from Proposition 16 we conclude that $\Pi_i(\mathbf{x}) = P(\mathbf{x})^2/(2g_i(x_i)) - c_i(x_i)$. Assumption 2 allows us to get the following result.

Proposition 17. *The profit of firm i , $\Pi_i(\mathbf{x})$, is quasi-concave in $x_i \in [0, \bar{x}_i]$.*

Moreover, the function $P(\mathbf{x})^2/(2g_i(x_i))$ is quasi-concave in $x_i \geq 0$, and attains its maximum when $q_i(\mathbf{x}) = \alpha\mu/2$.

Proof. We need to check that the derivative of each function changes sign at most once. From Proposition 16 it follows that the profit obtained by firm i at the market equilibrium can be written as

$$\Pi_i(\mathbf{x}) = \frac{\alpha^2\mu^2}{2g_i(x_i)} \left(\sum_{j=1}^n \frac{1}{g_j(x_j)} + \mu \right)^{-2} - c_i(x_i).$$

Define $\gamma \equiv \sum_{j=1, j \neq i}^n \frac{1}{g_j(x_j)} + \mu$, and note it is constant with respect to x_i . Then, the partial derivative of $\Pi_i(\mathbf{x})$ with respect to x_i is proportional to

$$\gamma - \frac{1}{g_i(x_i)} + \frac{2c'_i(x_i)}{\alpha^2\mu^2 g_i(x_i)g'_i(x_i)} (1 + \gamma g_i(x_i)). \quad (3.8)$$

From Assumption 1 the second term in equation (3.8) is increasing in x_i . Therefore, it is enough for the third term in equation (3.8) to be decreasing in x_i , for the partial derivative of $\Pi_i(\mathbf{x})$ with respect to x_i to be increasing. The derivative of the third

term in equation (3.8) with respect to x_i is proportional to

$$\begin{aligned}
& \frac{g_i''(x_i)}{g_i'(x_i)} + \frac{g_i'(x_i)}{g_i(x_i)} - \frac{3\alpha\mu g_i'(x_i)}{\sqrt{2} + \alpha\mu g_i(x_i)} - \frac{c_i''(x_i)}{c_i'(x_i)} \\
& < \frac{g_i''(x_i)}{g_i'(x_i)} - \frac{2g_i'(x_i)}{g_i(x_i)} - \frac{c_i''(x_i)}{c_i'(x_i)} \\
& = -\frac{2(g_i'(x_i))^2 - g_i''(x_i)g_i(x_i)}{g_i(x_i)g_i'(x_i)} - \frac{c_i''(x_i)}{c_i'(x_i)} \\
& \leq 0.
\end{aligned}$$

The first inequality follows from $g_i'(x_i) < 0$. The last inequality follows from Assumption 2.

Hence, the partial derivative of $\Pi_i(\mathbf{x})$ with respect to x_i is increasing and $\Pi_i(\mathbf{x})$ is quasi-concave in $x_i \in [0, \bar{x}_i]$.

Similarly, the partial derivative of $P(\mathbf{x})^2/(2g_i(x_i))$ is proportional to

$$\left(\sum_{j=1, j \neq i}^n \frac{1}{g_j(x_j)} + \mu \right) g_i(x_i) - 1. \quad (3.9)$$

Which is decreasing in $x_i \geq 0$, therefore the function $P(\mathbf{x})^2/(2g_i(x_i))$ is quasi-concave in $x_i \geq 0$. Moreover, it attains its minimum when we set the expression in equation (3.9) to zero. Namely

$$\left(\sum_{j=1}^n \frac{1}{g_j(x_j)} + \mu \right) = \frac{2}{g_i(x_i)}.$$

Or equivalently $q_i(\mathbf{x}) = \frac{\alpha\mu}{2}$. ■

Proposition 17 leads to the following Theorem.

Theorem 9. *There exists a market equilibrium as a function of the technology investments x_i .*

Moreover, if the financing cost of each firm is zero there is no market equilibrium, as each firm keeps increasing its investment level x_i without bound.

Proof. The strategy set space of each player is $[0, \bar{x}_i]$, a compact and convex set. The profit function $\Pi_i(\mathbf{x})$ is continuous and quasi-concave in $x_i \in [0, \bar{x}_i]$. Hence,

the existence of a pure strategy equilibrium follows from the Debreu-Glicksberg-Fan Theorem, see for example Tirole (1988).

From Proposition 17, it follows that if the financing cost of each firm is zero, then each firm has an incentive to increase its technology investment up to the point where its market output is $q_i(\mathbf{x}) = \alpha\mu/2$. Note that this is the optimal output of a monopolist with no production costs, facing a linear inverse demand function $P(Q) = \alpha - \frac{1}{\mu}Q$. Moreover, this output is unattainable for two or more firms simultaneously. Hence, each firm keeps increasing its investment level x_i without bound. ■

Let $\hat{\mathbf{x}}$ be the equilibrium technology investment vector. For simplicity, let us rescale, without loss of generality, the investment levels such that the equilibrium technology investments are denoted by $\hat{x}_i = 0$. We consider the case where the market consumption induced by $\hat{\mathbf{x}}$ is less than what is socially optimal. In this context, the central planner intervenes the market with the objective of minimizing the market price. The central planner invests her budget B , which we assume to be integer, into technology subsidies $x_i \geq 0$ (additional technology investments beyond the equilibrium levels), for each firm i . Note that from Theorem 9 it follows that it is in the best interest of each firm to invest the technology subsidy in becoming more efficient, as this extra technology investment has no cost. We consider the case where the central planner has an integer upper bound, denoted u_i , on the amount of money that she can allocate to each firm i . These upper bounds are motivated by fairness constraints. From the closed form expression given in equation (3.7), it follows that the problem faced by the central planner can be written as,

$$\begin{aligned} \min_{\mathbf{x}} \quad & P(\mathbf{x}) = \alpha\mu \left(\sum_{i=1}^n \frac{1}{g_i(x_i)} + \mu \right)^{-1} \\ \text{(TSAP)} \quad & \text{s.t.} \quad \sum_{i=1}^n x_i \leq B \\ & 0 \leq x_i \leq u_i \quad \forall i. \end{aligned}$$

From equation (3.7) it follows that, in order to minimize $P(x)$, we can equivalently maximize the convex function $\sum_{i=1}^n \frac{1}{g(x_i)}$ over a polyhedron. It follows that, in the

absence of upper bounds u_i on the amount of money allocated to each firm i , in an optimal solution the whole budget B would be allocated to only one firm. However, this type of solution would increase the market share of the selected firm, and decrease everyone else's, resulting in a highly concentrated market. Recognizing that allocating the whole budget to only one firm can be impractical, it is natural to consider upper bounds on the technology subsidy that can be allocated to each firm.

By defining the convex function $f_i(x_i) \equiv \frac{1}{g_i(x_i)}$, for each i , it follows that the central planner's problem (TSAP) is equivalent to our continuous knapsack problem with separable convex utility functions (CKP). Moreover, any α -approximation algorithm for problem (CKP) leads to a $\frac{1}{\alpha}$ -approximation algorithm for problem (TSAP) (note that problem (TSAP) is a minimization problem, while problem (CKP) is a maximization problem). Specifically,

$$\frac{P(\mathbf{x}^{\text{alg}})}{P(\mathbf{x}^*)} = \frac{F(\mathbf{x}^*) + \mu}{F(\mathbf{x}^{\text{alg}}) + \mu} \leq \frac{F(\mathbf{x}^*)}{F(\mathbf{x}^{\text{alg}})} \leq \frac{1}{\alpha}.$$

Where the equality follows from equation (3.7) and $F(x) = \sum_{i=1}^n f_i(x_i)$. The first inequality follows from $\mu > 0$.

Therefore, the results from previous sections suggest that simple subsidy allocation policies have a good performance guarantee for problem (TSAP). In particular, Theorems 7 and 8 show that simple ideas, like allocating the subsidies greedily to the firms that can increase their efficiency *faster* (Algorithms 1 and 3), or allocating the subsidies greedily to the firms that can increase their efficiency *the most* (Algorithm 2), have a guaranteed performance for this model.

3.4 Conclusions

In this chapter we have studied a continuous knapsack problem with separable convex utilities. We have shown the *NP*-hardness of the problem, and we have presented two simple algorithms that have both worst-case performance guarantees and a practical interpretation. Moreover, we have identified special settings where these simple

algorithms actually find the optimal solution.

As an application of this problem we have considered a novel model for the allocation of lump sum subsidies to competing firms, where the objective is to minimize the market price of a good in the presence of endogenous market competition, and subject to a budget constraint and upper bounds on the amount that can be allocated to each firm. The algorithms presented in this chapter suggest that simple subsidy allocation policies have a good performance in minimizing the market price of a good for this model.

Part II

Supply Chain Procurement

Chapter 4

Optimizing Purchasing and Handling Costs for Supply Chain Procurement

4.1 Introduction.

Procurement is a fundamental area for most large companies. It encompasses issues such as multi-sourcing, supplier relationship management and procurement contracts, see for example Mieghem (2008). However, procurement decisions in practice are often made in a silo, without taking into consideration the effect that they might have on the total internal supply chain costs of the company. In this chapter, we introduce a model that incorporates the cost of handling orders at a central distribution center, into the procurement decisions. In particular, the model provides insights into how the size of procured case packs affects the purchasing costs, as well as the handling costs incurred when serving orders at the distribution center.

The supply chains of many companies in practice have several distribution centers, and each distribution center satisfies orders placed by multiple end-point locations, for many different SKUs. For each SKU, each distribution center places orders to a supplier, who offers multiple case pack sizes, at different per-unit prices. Traditionally, in most companies the case pack selection decisions are made by the procurement department, which is primarily interested in minimizing the purchasing costs. Consequently, the case pack size usually selected is simply the one that attains the cheapest

per-unit purchasing cost. On the other hand, it is in the distribution center's best interest to select the case pack size that better fits the sizes of the orders received from the end-point locations, so as to minimize its handling costs. This chapter focuses on developing an optimization framework to inform the case pack selection in procurement contracts, which balances the procurement department's purchasing costs with the distribution center's handling costs, and applying it on real data from the supply chain of a large utility company.

To make this trade-off more concrete let us consider the following example: assume that the supplier of a given SKU offers a discount, of a few cents per unit, for a case pack of 200 units, making it the cheapest case pack size in terms of its per-unit purchasing cost. Additionally, assume that the orders placed by the end-point locations are mostly between 10 to 50 units. In this case, choosing the case pack of 200 units minimizes the purchasing costs at the expense of inducing large handling costs. Specifically, the workforce at the distribution center would spend most of their time breaking case packs of 200 units, then picking the 10 to 50 units of each order individually, and re-packaging them to send them to the end-point locations. One alternative could be to choose a case pack of 10 units. This option would fit the orders much better, reducing the handling costs at the expense of the workforce at the distribution center rounding up most of the orders to a multiple of 10, which might lead to larger purchasing cost. In other words, the distribution center would be sending out more units than the amount requested at the end-point locations. We refer to the latter as waste cost, as these units are considered a loss by most companies, both from an accounting and from a practical perspective. This is a fundamental characteristic of the practical settings that motivate this research. In more details, the distribution center has generally little or no incentive to retrieve these extra units. However, the value of the total number of units unnecessarily sent to the field, aggregating over all the SKUs, can be significant. Moreover, this cost is incurred by the procurement department, as it may end up purchasing significantly more units over a year than what is actually requested at the end-points locations. At the same time, in the practical settings that motivate this research, the distribution

center would argue that the large number of extra units sent to the field is a direct result of the case pack selected by the procurement department, a decision that is made without taking into account its impact on the distribution center's operations.

Let us point out that our setting is similar to the classical one-warehouse multi-retailer (OWMR) problem, see for example Zipkin (2000). However, we incorporate the case pack size as a decision variable, and we introduce a novel way to explicitly model the handling costs incurred, when using a given case pack to serve an order. We consider a fixed inventory policy, and we assume that it induces no shortages, allowing us to simplify the problem, and to focus on the trade-off between selecting a case pack size that takes advantage of per unit discounts on the purchasing cost, and the potential mismatch between the case pack size chosen, and the order sizes that need to be filled at the distribution center, which might imply larger handling costs.

Finally, let us emphasize that the insights that we obtain are applicable to many large companies with an internal supply chain, such as large construction companies and pharmaceuticals.

4.1.1 Main Contributions

We introduce a new model that minimizes the long run average purchasing and handling costs induced by the case pack selection in procurement contracts. We prove structural results that lead to a practical method to both selecting the best case pack size per SKU, and serving orders at the distribution center. Furthermore, we test this method on real data from a large utility company, finding significant total cost reductions.

Our model brings new insights into the procurement literature, by explicitly modeling the effects of the case pack selection in procurement contracts, on the total internal supply chain costs of a company. Specifically, we study how it affects the distribution center's policy for serving orders, and the associated handling costs it induces. To achieve this, we model the most relevant activities carried out at the distribution center in order to serve the orders, including breaking case packs, and picking single units manually.

We first consider the practically relevant problem of selecting the best case pack size per SKU. For this problem, we show that a threshold policy is optimal for serving orders at the distribution center. Notably, the threshold value is independent of the discrete probability distribution over the order sizes, and of the number of single units available from broken case packs. Moreover, the threshold value has an intuitive closed form, which illustrates the interplay between the different cost parameters defined by the procurement department's case pack selection. This allows us to obtain a close form expression for the long run average purchasing and handling costs induced by choosing any given case pack size, providing a practical method to select the case pack size that induces the least total cost.

For the problem of choosing multiple case pack sizes per SKU, we use the insights derived for the single case pack size problem to show that, under some assumptions on the cost structure and case pack sizes, selecting at most three sizes provides a 2-approximation to the optimal cost. Namely, we show that, for each instance of the problem, there exists a policy that only selects at most three case pack sizes, which is guaranteed to have a total cost that is at most 2 times the optimal total cost. This is important because the optimal policy can potentially imply selecting every case pack size available from the supplier, making it highly impractical.

Finally, we test the method developed to select the best case pack size per SKU on real data from a large utility company. The numerical results suggest that the proposed method has the potential to reduce the purchasing and handling costs of a SKU by 16%. Importantly, our proposed threshold policy is simple to implement, and to communicate. In fact, the distribution centers at the utility company were already using a threshold policy, albeit with a fixed threshold of 50% of the case pack size for each SKU, simplifying the application of our method in practice. Similarly, the implementation of our method is also simple from the procurement department's perspective. It only requires to compare the easily computable long run average purchasing and handling costs induced by each case pack size available from the supplier, therefore facilitating the incorporation of the distribution center's handling costs into the procurement department decisions.

4.2 Related Literature

Our problem is related to the classical assortment problem in Pentico (1974), where the objective is to identify which set of sizes of some product should be stocked, when substitution in one direction is available at some cost. Pentico (2008) provides a recent and thorough review of this stream of literature. The main differences with respect to our model are related to the type of costs being considered. Specifically, in our model demand is generic, as opposed to being specific to a given case pack size, hence there is no substitution cost. Additionally, we incorporate the handling cost induced at the distribution center by selecting any given case pack size.

Another related research area is on the design and control of warehouse order picking, see Koster et al. (2007) for a literature review. The typical problems considered in this literature include layout design, storage assignment, routing, order batching and zoning. Let us emphasize that Koster et al. (2007) state that broken case picking is an important warehouse function. However, the decision of when to break a case pack, and when to round up and serve an order with a whole case pack, is not included in any paper in their review. This is precisely one of the decisions we consider in this chapter, specifically in how it is related to the case pack selection. As already mentioned, our problem is also related to the OWMR model, where one warehouse orders a product from a supplier to serve orders from multiple retailers. The objective is to decide the warehouse and retailers' orders so as to minimize the fixed ordering costs plus the inventory holding costs over the planning horizon, see Zipkin (2000) for a review of the classical results for this model.

Recently, Wen et al. (2012) worked with a major US retailer in considering the problem of selecting the ship-pack for each SKU, in a two-echelon distribution system. They consider a similar cost structure for the distribution center's handling costs as we do in this chapter, as well as additionally considering inventory-related costs. However, they do not consider inventory decisions in their model. Moreover, for tractability purposes they assume that the weekly demand occurs at a known constant rate, and that the inventory position of a SKU when a store places an order

follows a uniform distribution, between zero and the store's known reorder point. The implementation of their model on real data provided by a retailer suggests that, by selecting the appropriate ship-pack size per SKU, the retailer can reduce its total cost by 0.3% - 0.4%. In contrast, in this chapter we focus on the trade off between purchasing costs and handling costs, and we do not consider inventory costs, allowing us to consider a general discrete demand distribution. The common insight from both this paper and our work is that, modeling and optimizing the handling costs induced at distribution centers when dealing with different case pack sizes, can have a significant impact on the total cost incurred by companies in practice.

Finally, there is little theoretical work in operations management that considers the case pack size as a decision variable. Exceptions include Cachon and Fisher (2000), where they find a substantial supply chain cost reduction generated by smaller batch (case pack) sizes in a OWMR setting. Similarly, Kök and Fisher (2007) also find a reduction in inventory costs generated by products with a smaller case pack size, in the context of an assortment planning model with product substitution. Later, Donselaar et al. (2010) present an empirical study of the ordering behavior of retail store managers, showing that they tend to deviate more, from the order advices generated by an automated inventory system, for products with larger case packs. In contrast to this literature, our model suggests that, when taking into account the handling costs at the distribution center, a smaller case pack size is not always better, as it may lead to higher costs. For example, consider the extreme case where each case pack size is a single unit, then each order would have to be picked manually, potentially increasing the handling costs. Our model captures the trade-off between purchasing costs and handling costs in the case pack selection, which leads to not necessarily selecting the smallest case pack available.

4.3 Model

Our goal in this chapter is to develop a model that allows us to identify both the set of case pack sizes that should be selected by the procurement department, as well as

the optimal policy to serve the orders at the distribution center. We now describe the main components of the model.

Demand For each SKU, we consider an infinite horizon model, where an infinite sequence of orders received from the end-point locations must be served at the distribution center. We assume that the order sizes are uncertain, and follow a known stationary discrete distribution, taking values from the set of positive integers $\{d_1, d_2, \dots, d_n\}$, for some positive integer n . Namely, we assume that $D_t = d_i$ with probability p_i , for $i \in \{1, \dots, n\}$. The goal is to minimize the long run average expected purchasing and handling costs.

Decisions To serve each order, we assume that the distribution center has a large enough supply of case packs, of each of the sizes that have been already selected by the procurement department. Let us denote by the positive integer m , the number of case pack types selected by the procurement department, and by S_j , $j \in \{0, 1, \dots, m\}$, their size in number of units. For each order, the distribution center has to decide (i) how many whole case packs of each size to use to serve the order, and (ii) how many units to pick manually from broken case pack of each size, breaking new case packs if necessary.

Costs We consider the following stationary cost structure. For each case pack size S_j , $j \in \{0, 1, \dots, m\}$, we model the purchasing costs by a price per case pack P_j quoted by the supplier. The handling costs include a cost C_j for using a whole case pack, a cost K_j for breaking a case pack, and a per unit picking cost V for units picked manually from an opened case pack, independent of the specific size of the case pack they come from. Note that the cost K_j may include the additional costs associated with a broken case pack, including shrinkage and the increased cost of keeping an accurate inventory.

We make the following natural assumption on the cost structure

Assumption 3. *Assume that for each case pack size S_j , $j \in \{0, 1, \dots, m\}$, it is always*

more convenient to pick the whole case pack rather than opening it and picking all its units manually. Namely, assume that $C_j < K_j + S_j V$, for each $j \in \{0, 1, \dots, m\}$.

Assumption 3 is intuitive, and it was satisfied by the instances in the numerical experiments on real data from a large utility company described later in Section 4.5. Moreover, this assumption is precisely what makes the problem interesting, otherwise the optimal policy is to serve all the orders using manual picking.

Our model was originally motivated by the supply chain of a large utility company. Due to constraints imposed by its IT systems and processes, the company was interested in identifying the best *unique* case pack size per SKU. In particular, the software used at the distribution center only supports one case pack size per SKU. Therefore, different case pack sizes would be treated as different SKUs, and thus be placed at a random location within the distribution center, see for example Koster et al. (2007). As a result, having multiple case pack sizes per SKU could significantly complicate the picking process. Similarly, the processes followed by the procurement department are also designed for only one case pack size per SKU. Motivated by this practical requirements, we start by specializing our model for the case where the procurement department selects one case pack size per SKU.

4.3.1 DP Formulation for Selecting the Best Case Pack Size

In this section, we will assume that $m = 1$. Namely, that the procurement department selects a unique case pack size per SKU. Therefore, we will drop the index j for the case pack type.

We will address the problem in two steps. First, assuming that the procurement department has already selected a case pack size, we will study what is the optimal policy to serve orders at the distribution center. Second, assuming that the distribution center follows the optimal policy from the first step, we will identify what is the optimal case pack size that the procurement department should select.

From our assumptions on the cost structure, there are a couple of preliminary results that follow, which we describe next.

Preliminary Remarks From Assumption 3 it follows that, for each order size d_i , $i \in \{1, \dots, n\}$, there is nothing better that we can do with the first $\lfloor \frac{d_i}{S} \rfloor$ units of the order, than to serve them by picking whole case packs. Hence, any optimal solution has a long run average expected cost of at least

$$(P + C)\mathbb{E} \left[\left\lfloor \frac{D}{S} \right\rfloor \right] = (P + C) \sum_{i=1}^n p_i \left\lfloor \frac{d_i}{S} \right\rfloor, \quad (4.1)$$

It follows that solving the problem in step one reduces to deciding how to serve the remainder units of the orders that are not considered in equation (4.1). Therefore, let us define the a priori random remainder orders R_t by

$$R_t \equiv D_t - \left\lfloor \frac{D_t}{S} \right\rfloor S \quad \text{for each } t \in \{1, 2, \dots\},$$

where by definition we have $R_t \in \{0, 1, \dots, S - 1\}$. Note that, from our assumption on the stationary distribution of the order sizes D_t , it follows that the remainders R_t are independent and identically distributed, with the following discrete distribution, $R_t = k$, $k \in \{0, 1, \dots, S - 1\}$, with probability $q_k = \sum_{i \in \{i: d_i - \lfloor \frac{d_i}{S} \rfloor = k\}} p_i$.

The problem of deciding how to serve the remainder orders can be posed as the following infinite horizon dynamic program. Let $x_t \in \{0, 1, \dots, S - 1\}$ be the leftover units available at stage t from an opened case pack, and $r_t \in \{0, \dots, S - 1\}$ be the remainder order realized at stage t . Then, (x_t, r_t) will be the state of the dynamic program. Let $\mu_t \in \{W, O\}$ be the control at stage t , namely a function from the state of the dynamic program to the feasible set of actions. Specifically, $\mu_t = W$ denotes serving the remainder r_t using a whole case pack, and $\mu_t = O$ denotes serving the remainder r_t by picking individual units from a broken case pack, breaking a new case pack in the process if necessary. Additionally, let $\Pi = \{\mu_0, \mu_1, \dots\}$ be a policy for the infinite horizon problem.

Let $g(x_t, r_t, \mu_t)$ be the cost incurred at stage t when applying control μ_t , and being in state (x_t, r_t) . From the description of the cost structure it follows that $g(x_t, r_t, W) = P + C$, and $g(x_t, r_t, O) = Vr_t + (P + K)\mathbb{1}_{\{r_t > x_t\}}$. Similarly, the state

transitions for the leftover units available at stage t from an opened case pack, x_t , are

$$x_{t+1}(x_t, r_t) = \begin{cases} x_t & \text{if } \mu_t = W \\ x_t - r_t & \text{if } \mu_t = O \text{ and } r_t \leq x_t \\ x_t + S - r_t & \text{if } \mu_t = O \text{ and } r_t > x_t. \end{cases}$$

Namely, if the control is to serve r_t using a whole case pack, then x_t stays the same. If the control is to serve r_t by picking individual units from an opened case pack, then x_t is reduced by r_t , and S units are added if it was necessary to open a new case pack (i.e., if $r_t > x_t$). Note that we assume that at any point in time there will be at most one case pack opened. This policy of sequentially opening case packs as needed is intuitive and practical. Moreover, the analysis of the problem will show that this modeling assumption does not play a relevant role.

Finally, let $J_{\Pi}(x_0, r_0)$ be the long run average expected cost induced by policy Π , starting from the initial state (x_0, r_0) . Note that

$$J_{\Pi}(x_0, r_0) = \limsup_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \left[\sum_{t=0}^{N-1} g(X_t, R_t, \mu_t) \right].$$

Then, the problem that we are interested in solving is finding

$$J^*(x_0, r_0) = \inf_{\Pi} J_{\Pi}(x_0, r_0),$$

for any initial state (x_0, r_0) .

Alternative Cost Accounting There is an alternative method to account for the cost incurred at each stage, which significantly simplifies the analysis, as stated in the following proposition. We will focus on state dependent stationary policies.

Proposition 18. *For any state dependent stationary policy, the long run average cost attained by the original cost accounting $g(x_t, r_t, O) = Vr_t + (P + K)\mathbb{1}_{\{r_t > x_t\}}$, $g(x_t, r_t, W) = P + C$, is equivalent to the long run average cost attained by the alter-*

native cost accounting $g'(x_t, r_t, O) = \left(V + \frac{(P+K)}{S}\right) r_t$, $g'(x_t, r_t, W) = P + C$.

Moreover, minimizing the long run average cost is equivalent to minimizing the average charge per unit.

Proof. Let μ be a stationary state dependent policy. Let $C(N)$, $C'(N)$ be its cumulative cost from stage 1 up to N , using the cost accounting $g(x_t, r_t, \mu)$, and $g'(x_t, r_t, \mu)$, respectively. Additionally, let $B(N) = C(N) - C'(N)$.

Note that, for any stage N , $B(N) \leq (P + K)$. This follows because at most one case pack will be broken and charged at a time, before it is actually used for manual picking. Hence,

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{C(N)}{N} &= \limsup_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \left[\sum_{t=0}^{N-1} g(X_t, R_t, \mu_t) \right] \\ &= \limsup_{N \rightarrow \infty} \frac{1}{N} \left(\mathbb{E} \left[\sum_{t=0}^{N-1} g'(X_t, R_t, \mu_t) \right] + B(N) \right) \\ &= \limsup_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \left[\sum_{t=0}^{N-1} g'(X_t, R_t, \mu_t) \right] \\ &= \limsup_{N \rightarrow \infty} \frac{C'(N)}{N}. \end{aligned}$$

Where the third equality follows because $B(N)$ is bounded by a constant. This completes the proof of the first statement in the proposition.

Let $U(T)$ be the total number of order units satisfied by stage T . We consider the alternative cost accounting $g'(x_t, r_t, \mu)$, and every time a whole case pack is used, we allocate its cost uniformly over all the units being served. Let a_i the cost allocated to unit i in this manner, for $i \in \{1, \dots, U(T)\}$. Note that a_i has a stationary distribution, for any stationary state dependent policy. Then, we have that the long run average cost induced by μ is

$$\begin{aligned}
\limsup_{N \rightarrow \infty} \frac{C'(N)}{N} &= \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^{U(N)} a_i \\
&= \limsup_{N \rightarrow \infty} \frac{U(N)}{N} \sum_{i=1}^{U(N)} \frac{a_i}{U(N)} \\
&= \mathbb{E}[R] \limsup_{N \rightarrow \infty} \sum_{i=1}^{U(N)} \frac{a_i}{U(N)}.
\end{aligned}$$

Because $\mathbb{E}[R]$ is a constant that does not depend on the policy, we conclude that minimizing the long run average cost is equivalent to minimizing the average charge per unit. ■

The alternative cost accounting from Proposition 18 allocates the per unit long run average cost of using a broken case pack to each unit, as opposed to keeping track of the necessity of opening a new case pack at the appropriate stage. Specifically, the long run average cost of using an opened case pack is the cost of picking a single unit V , plus the total cost of purchasing and breaking a case pack $(P + K)$, divided by the total number of units in the case pack S . Intuitively, this follows from the observation that any stationary policy will use all the leftover units of an opened case pack to satisfy some order in the long run.

4.4 Structural Insights

In this section we present the structural insights we derived for the model described in Section 4.3. In particular, Section 4.4.1 shows how to efficiently solve the problem of selecting the case pack size that minimizes the long run average purchasing and handling costs, for each SKU. Additionally, Section 4.4.2 shows that, under some assumptions, selecting at most 3 case pack sizes can be sufficient for a policy to be guaranteed to induce at most twice the optimal long run average purchasing and handling costs, for each SKU. Section 4.4.2 also provides counter-examples, which show that the assumptions made are, in some sense, necessary to obtain a worst-case

performance guarantee.

4.4.1 Selecting the Best Case Pack Size per SKU

The following is the main result in this section.

Theorem 10. *An stationary threshold policy is optimal to serve any given remainder order r_i at the distribution center.*

Moreover, for any given case pack size the optimal threshold \bar{d} is

$$\bar{d} = \frac{(P + C)S}{VS + P + K}, \quad (4.2)$$

independently of the discrete probability distribution over the order sizes.

Hence, the long run average expected purchasing and handling cost induced by selecting any given case pack size is

$$(P + C)\mathbb{E}\left[\left\lceil\left\lfloor\frac{D}{S}\right\rceil\right\rceil\right] + (P + C)\mathbb{P}(R > \bar{d}) + \left(V + \frac{P + K}{S}\right)\mathbb{E}[R|R \leq \bar{d}]\mathbb{P}(R \leq \bar{d}), \quad (4.3)$$

where $R \equiv D - \lfloor \frac{D}{S} \rfloor$ is the, a priori random, remainder order.

Proof. Without loss of generality, we focus on stationary state dependent policies. From Proposition 18 it follows that we can equivalently minimize the average charge per unit. Moreover, the optimality of the threshold policy in the the theorem follows directly from the alternative cost accounting from Proposition 18. Specifically, for any given remainder order r , the charge per unit induced by the optimal control is $\min\left\{\frac{P+C}{r}, \left(V + \frac{P+K}{S}\right)\right\}$, which leads to the threshold \bar{d} in equation (4.2). Similarly, computing the expected cost induced by this optimal threshold policy gives

$$(P + C)\mathbb{P}(R > \bar{d}) + \left(V + \frac{P + K}{S}\right)\mathbb{E}[R|R \leq \bar{d}]\mathbb{P}(R \leq \bar{d}). \quad (4.4)$$

Finally, equation (4.3) in the Theorem follows from adding up equations (4.1) and (4.4). ■

Let us emphasize that the threshold \bar{d} in equation (4.2) only depends on the parameters associated to the case pack size being used, and, surprisingly, it is independent of both the distribution of R_t , and of the leftover units available from an opened case pack x_t . Moreover, the threshold \bar{d} has an intuitive closed form expression that illustrates the interplay between the different cost parameters. Specifically, we can rewrite equation (4.2) as follows

$$\frac{P + C}{\bar{d}} = V + \frac{P + K}{S}. \quad (4.5)$$

From equation (4.5) it follows that the threshold value \bar{d} balances, on the left hand side, the per unit cost of using a whole case pack to serve a remainder order of size \bar{d} , with the per unit cost of using an opened case pack on the right hand side. In particular, on the left hand side of equation (4.5) the cost of purchasing and using a whole case pack to serve an order is divided by the number of units that are being satisfied \bar{d} .

Additionally, note that in most practical settings we will likely have a non trivial threshold policy, namely we will have $\bar{d} > 1$. For this to be the case it is enough that, for any case pack with S units, it is always more convenient to open the case pack and pick all its units manually, rather than satisfying S units of an order, each one of them with a whole case pack, that is $P + K + SV < (P + C)S$. If this is not the case, then it is optimal to round-up every order, and use only whole case packs to serve them.

Theorem 10 provides us with a closed form expression for the long run average purchasing and handling costs induced by selecting any given case pack size. It follows that we can then simply select the case pack that induces the least cost. Hence, Theorem 10 provides a practical method for both serving orders at the distribution center, and selecting the best case pack size per SKU at the procurement department, respectively.

4.4.2 Selecting Multiple Case Pack Sizes per SKU

In this section we consider the problem of selecting multiple case pack sizes per SKU. From the analysis of the problem of selecting the best case pack size per SKU in Section 4.4.1, it follows that in the long run it is optimal to pick units manually from a single case pack size. Namely, it is optimal to pick units manually only from case packs of the size that induces the smallest per unit long run average cost of using an opened case pack. Without loss of generality, we will denote the index of this size by 0, therefore its per unit long run average cost of using an opened case pack is $V_0 + \frac{(P_0+C_0)}{S_0}$.

Furthermore, the problem now becomes how to serve each possible order size d_i , $i \in \{1, 2, \dots, n\}$. Namely, what number of whole case packs of each size should be used to satisfy the order? How many units should be picked manually from case packs of size S_0 ?

Without loss of generality, we will only consider the case pack sizes S_j , $j \in \{1, 2, \dots, m\}$, to be sent as whole case packs to serve the orders (duplicating S_0 if necessary), so we can effectively split the actions of using whole case packs, and picking single units. Then, for any given order size d_i , the problem we are interested in solving can be casted as the following integer program.

The decision variables we consider include the number of case packs of size S_j to satisfy the order, denoted by z_j , for each $j \in \{1, 2, \dots, m\}$, and the number of units that are going to be picked manually from case packs of size S_0 , denoted by y . We impose the constraint that the total number of units sent must be *at least* the amount ordered d_i , as it may be cost efficient to send more units than the amount requested, incurring in a waste cost, in order to reduce the handling costs. In summary, the solution to the following integer program provides a detailed policy to serve the orders

from the end-point locations at the distribution center

$$\begin{aligned}
\min_{y, z_j} \quad & \sum_{j=1}^m z_j(P_j + C_j) + \left(V_0 + \frac{(P_0 + C_0)}{S_0} \right) y \\
\text{s.t.} \quad & \sum_{j=1}^m z_j S_j + y \geq d_i \\
(IP) \quad & z_j \in \mathbb{N}^+ \quad \forall j \in \{1, \dots, m\} \\
& y \geq 0.
\end{aligned}$$

Note that in the objective function of problem (IP) we have used the alternative cost accounting from Proposition 18. Namely, we have allocated the per unit long run average cost of using an opened case pack, $V_0 + \frac{(P_0 + C_0)}{S_0}$, to each unit picked manually, as opposed to keeping track of whether a case pack had to be opened for this particular order.

Because in any optimal solution each variable z_j will take only values in $\left\{0, 1, \dots, \left\lceil \frac{d_i}{S_j} \right\rceil\right\}$, the integer program (IP) can be reformulated using $\sum_{j=1}^m \left\lceil \frac{d_i}{S_j} \right\rceil$ binary variables. Moreover, any instance of the problem with a large enough cost of picking single units V_0 , is equivalent to a 0-1 covering problem. Hence, problem (IP) is NP-hard.

Furthermore, this problem is a covering version of the 0-1 knapsack problem with a single continuous variable introduced by Marchand and Wolsey (1999). Recently Zhao and Li (2013) provide a 2-approximation algorithm for the knapsack version of the problem, while Lin et al. (2011) provide an exact exponential algorithm for it, showing good results in simulations when compared to general purpose solvers. Unfortunately, these results do not carry over directly to the covering version of the problem.

In the remainder of this section, we will assume, without loss of generality, that the case pack sizes are labeled, and the respective costs are scaled, such that

$$\frac{P_1 + C_1}{S_1} \leq \frac{P_2 + C_2}{S_2} \leq \dots \leq \frac{P_n + C_n}{S_n} \leq V_0 + \frac{P_0 + C_0}{S_0} = 1. \quad (4.6)$$

Additionally, let $l = \arg \min_j \{P_j + C_j\}$ be the index of the case pack size with least

S_j	$P_j + C_j$	$(P_j + C_j)/S_j$
500	112	0.224
250	87	0.348
100	42	0.42
50	24	0.48
20	15	0.75

Table 4.1: An Instance of the Problem of Selecting Multiple Case Pack Sizes

S_j	$D_i = 795$	$D_i = 672$...
500	1	1	
250	1	0	
100	0	1	
50	1	1	
20	0	1	
S_0	0	2	

Table 4.2: Examples of Optimal Solutions to Problem (*IP*)

total cost of purchasing it, and using it to satisfy an order. Note that, without loss of generality, S_i corresponds to the smallest case pack size available. Specifically, if there exist indexes j and k such that $P_j + C_j \leq P_k + C_k$ and $S_j \geq S_k$, then it follows that for any solution that uses some case packs of size S_k , we can replace them by case packs of size S_j obtaining a feasible solution which attains an objective value that is no worse than the original. Hence, without loss of generality, we can assume that $P_j + C_j \leq P_k + C_k$ implies $S_j \leq S_k$, for any pair of indexes j and k .

Although the large utility company that provided us with real data was not interested in selecting multiple case pack sizes per SKU, the analysis in this section allows to explore the potential benefit of considering this option. Unfortunately, it is not hard to construct instances of the problem of selecting multiple case pack sizes per SKU, for which all the available sizes from the supplier are required in the optimal solution. For example, Table 4.1 describes an instance with the same set of sizes used in the numerical experiments in Section 4.5, but with different costs that have been normalized such that $V_0 + \frac{P_0 + C_0}{S_0} = 1$. For any given order size D_i , we obtain an instance of problem (*IP*). Table 4.2 describes the optimal solutions for a couple of order sizes, including the number of whole case packs of each size used to serve the

order, as well as the number of single units picked manually from a case pack of size S_0 in the last row of the table. These examples are sufficient to show how all the sizes available from the supplier are used in some optimal solution. This characteristic makes the optimal policy potentially impractical, as many different case packs would have to be carried, for each SKU, in order to serve the orders, significantly increasing the complexity of the picking and handling operations at the distribution center, as well as the negotiations between the procurement department and the suppliers.

In this context, a relevant question is whether only selecting a limited number of case packs from the supplier is sufficient to have a constant worst-case performance guarantee. We will show in the main result of this section that, under some assumptions, selecting at most three sizes will induce a total cost that is at most twice the cost of the optimal policy. Namely, selecting at most three sizes will be sufficient to have a 2-approximation for the problem of selecting multiple case pack sizes per SKU.

Theorem 11. *For any order size $d_i \geq \frac{S_1}{2}$, simply using case packs of size S_1 gives a 2-approximation for problem (IP).*

Additionally, if $P_1 + C_1 \leq 2(P_l + C_l)$ or $S_1 \leq 2S_l$, then using only case packs of size S_0 , S_1 and S_l gives a 2-approximation for the problem of selecting multiple case pack sizes per SKU.

Proof. For any given order size d_i , let us denote the optimal objective value of problem (IP) by $\text{OPT}(d_i)$. For any order size $d_i \geq \frac{S_1}{2}$ we have that

$$2\text{OPT}(d_i) \geq 2 \left(\frac{P_1 + C_1}{S_1} \right) d_i \geq \left\lceil \frac{d_i}{S_1} \right\rceil (P_1 + C_1),$$

where the first inequality follows from equation (4.6), and the second inequality follows from the assumption that $d_i \geq \frac{S_1}{2}$. The last expression is exactly the cost incurred by only using case packs of size S_1 to serve the order d_i . Therefore, we conclude that using only case packs of size S_1 induces a cost that is at most twice the cost of the optimal solution of problem (IP).

For the second result in the theorem, note that for any order size $d_i \leq S_l$ the optimal cost of serving it is $\min\{d_i, P_l\}$, where d_i is the cost allocated to picking

the units manually, because we have normalized $V_0 + \frac{(P_0+C_0)}{S_0} = 1$ in equation (4.6). Namely, the best that we can do to serve the order is to use a threshold policy between manual picking from case packs of size S_0 , and rounding up to a case pack of size S_l . Hence, we know what the optimal policy is for any order $d_i \leq S_l$. To conclude note that if $S_1 \leq 2S_l$, then for any order size $d_i \geq S_l \geq \frac{S_1}{2}$ we get a 2-approximation from the previous analysis and we are done. Similarly, if $P_1 + C_1 \leq 2(P_l + C_l)$ then for any order $d_i \in (S_l, \frac{S_1}{2}]$ we have that

$$2\text{OPT}(d_i) \geq 2(P_l + C_l) \geq P_1 + C_1.$$

Therefore, for any order $d_i \geq S_l$ we get a 2-approximation by using only case packs of size S_1 .

Because this holds for any order size d_i , it follows that if $P_1 + C_1 \leq 2(P_l + C_l)$ or $S_1 \leq 2S_l$, then using only case packs of size S_0 , S_1 and S_l gives a 2-approximation for the problem of selecting multiple case pack sizes per SKU, for any discrete distribution of the order sizes, and for any number of case pack types available from the supplier. ■

The result in Theorem 11 is interesting as it gives sufficient conditions for a more practical policy to have a guaranteed performance compared to the optimal policy. In particular, the case pack sizes suggested to be selected are intuitive and can be motivated in practice. S_0 is the case pack size that induces the least per unit long run average cost of picking units from an opened case pack, therefore it is a size that is preferred by the distribution center to minimize the cost of their manual picking operations. Similarly, S_1 is the case pack size that minimizes the per unit cost of purchasing a case pack and using to satisfy an order, therefore it partially considers the procurement department objective of minimizing the purchasing cost, incorporating into this criteria the cost of using the whole case pack to serve orders at the distribution center. Finally, S_l is the smallest case pack size available, and the one that minimizes the absolute cost of purchasing a case pack and using to satisfy an order. In order words, S_l can often be the most convenient case pack size to round-up

an order.

Counter-examples with unbounded worst-case performance The assumptions on the cost structure and case pack sizes made in Theorem 11 are, in some sense, necessary in order to have a guaranteed performance using at most three case pack sizes. Specifically, we now provide counter-examples for the case where the assumptions in Theorem 11 are not met. It shows that the relative performance attained by selecting the natural case pack sizes S_0 and S_1 , together with at most one additional case pack size, can be arbitrarily bad when compared to the optimal solution.

For any given order size d_i , let us denote by $\text{ALG}(d_i)$ the cost incurred by serving the order using only case pack sizes S_0 and S_1 , together with at most one additional case pack size. Similarly, let us denote by $\text{OPT}(d_i)$ the cost incurred by serving d_i using the optimal policy given by solving problem (IP) .

First, assume that the additional size selected is S_l . Consider the following instance of problem (IP) , with an additional case pack size S_j , where everything is parametrized in terms of S_1 ,

- $P_1 + C_1 = \sqrt{2S_1} - \left\lfloor \sqrt{\frac{S_1}{2}} \right\rfloor \epsilon.$
- $P_l + C_l = 2 - \epsilon$, for $\epsilon > 0$ small enough, $S_l = 2.$
- $P_j + C_j = 2$, $S_j = \sqrt{2S_1}.$
- $d_i = \sqrt{2S_1}.$

It is not hard to see that this instance is well defined, in the sense that it satisfies equation (4.6), and the definition of the index m , for $\epsilon > 0$ small enough and $S_1 > 2$. Similarly, $d_i = \sqrt{2S_1} \in (S_l, \frac{S_1}{2})$ for any $S_1 > 8$, and $2(P_l + C_l) < P_1 + C_1$ for S_1 large enough, and $\epsilon > 0$ small enough. Hence, the sufficient conditions in Theorem 11 are not met for this instance. The analysis will be based on increasing S_1 in order to get an unbounded relative performance of the policy that only uses the case pack sizes S_0 , S_1 , and S_l to serve the order, when compared to the optimal policy.

The cost incurred when using only case pack sizes S_0 , S_1 , and S_l to serve the order $d_i = \sqrt{2S_1}$ is $\text{ALG}(S_1) = P_1$, as the best we can do in this case is to round-up to a case pack of size S_1 . Specifically, we have that

$$\begin{aligned} P_1 + C_1 &= \sqrt{2S_1} - \left\lfloor \sqrt{\frac{S_1}{2}} \right\rfloor \epsilon \\ &= d_i - \left\lfloor \frac{d_i}{S_l} \right\rfloor \epsilon \\ &= \left\lfloor \frac{d_i}{S_l} \right\rfloor P_l + \left(d_i - \left\lfloor \frac{d_i}{S_l} \right\rfloor S_l \right), \end{aligned}$$

where the second and third equalities follow from the definitions of d_i , S_l and P_l for this instance. The last expression is precisely the cost incurred when using case packs of size S_l , and picking units manually from opened case packs of size S_0 , to serve the order. This follows because $\left(d_i - \left\lfloor \frac{d_i}{S_l} \right\rfloor S_l \right) \leq P_l = S_l - \epsilon$ for $\epsilon > 0$ small enough. Hence, the instance is such that we are indifferent between rounding up to a case pack of size S_1 , and using case packs of size S_l together with manual picking. We also have $\text{OPT}(S_1) \leq P_j = 2$, where the inequality follows because using a case pack of size S_j is a feasible solution. Hence, we conclude that

$$\frac{\text{ALG}(S_1)}{\text{OPT}(S_1)} \geq \sqrt{\frac{S_1}{2}} - \left\lfloor \frac{S_1}{2} \right\rfloor \frac{\epsilon}{2}.$$

Therefore, the relative performance of the policy that only uses case pack sizes S_0 , S_1 , and S_l can be made arbitrarily bad by increasing S_1 .

The other case is similar. Assume that the additional case pack size selected is $S_j \neq S_l$. Consider the following instance of problem (IP) ,

- $P_1 + C_1 = \frac{S_1}{\alpha}$, for $\alpha > 1$.
- $P_j + C_j = S_j - \epsilon$, for $0 < \epsilon < \frac{\alpha - 1}{\alpha} S_j$.
- $P_l + C_l = \epsilon$, $S_l = S_j - \epsilon$.

The bounds on α and ϵ ensure that equation (4.6) holds for indexes 1 and j . Additionally, equation (4.6) for indexes 1 and m imposes the following consistency constraint

$\epsilon \geq \frac{S_j}{\alpha+1}$. Similarly, the definition of the index m imposes the consistency constraint $\epsilon > S_j - S_1$.

Consider any fixed order size d_i such that $S_l \leq d_i = k(S_j - \epsilon) \leq \frac{S_1}{2}$, for an appropriate positive integer k . Similarly, the condition $2P_l = 2\epsilon < P_1 = \frac{S_1}{\alpha}$ ensures that the sufficient conditions in Theorem 11 are not met for this instance. The cost incurred when using only case pack sizes S_0 , S_1 , and S_j to serve the order is

$$\begin{aligned} \text{ALG}(k) &= \left\lfloor \frac{d_i}{S_j} \right\rfloor P_j + \left(d_i - \left\lfloor \frac{d_i}{S_j} \right\rfloor S_j \right) \\ &= k(S_j - \epsilon) - \left\lfloor \frac{k(S_j - \epsilon)}{S_j} \right\rfloor \epsilon, \end{aligned}$$

where again the equality follows from the definitions of P_j and d_i for this instance. Additionally, $\text{OPT}(k) \leq k\epsilon$, where the inequality follows because using k case packs of size S_j is a feasible solution. Hence, it follows that

$$\frac{\text{ALG}(k)}{\text{OPT}(k)} \geq \frac{(S_j - \epsilon)}{\epsilon} - \left\lfloor \frac{k(S_j - \epsilon)}{S_j} \right\rfloor \frac{1}{k}.$$

Therefore, the relative performance of the policy that only uses case pack sizes S_0 , S_1 , and S_j can be made arbitrarily bad by decreasing ϵ , and increasing S_1 , S_j and α appropriately to satisfy the consistency constraints between these parameters.

4.5 Numerical Experiments on Real Data

In this section we present the results obtained when testing the method for serving orders at the distribution center, and for selecting the case pack size that induces the least long run average purchasing and handling costs, as described Section 4.4.1, on real data from a large utility company. Due to confidentiality reasons, all the names and specific costs have been disguised.

For the numerical experiments, we first selected the SKU carried by the company that was most frequently ordered during the year 2013, with the goal of being able to quickly estimate the cost parameters of the model. This SKU was a square washer, a

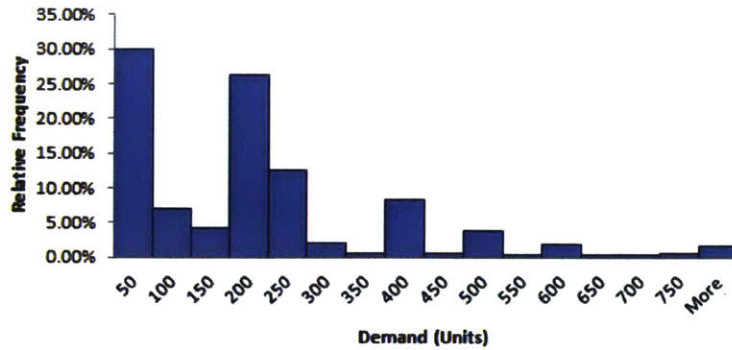


Figure 4-1: Empirical Distribution of the Order Size for the Square Washer in 2013

low-cost high-volume item frequently used in construction projects, which had more than 3,600 independent orders placed during 2013. The case pack that was being used for this item was of 250 units. Interestingly, the distribution center was already using a threshold policy to serve the orders received from the end-point locations, for essentially every SKU that they carried. However, they were using a rule of thumb for the threshold value of 50% of the case pack for each SKU, regardless of any differences among the SKUs. This significantly facilitates the implementation of our proposed method to serve orders at the distribution center, as it only requires to change the threshold used to the value given in Theorem 10, which depends only on the cost parameters.

For the probability distribution of the order sizes we used the empirical distribution of the orders placed in 2013, displayed in Figure 4-1. We considered the 5 different case pack sizes offered by the supplier, which are $\{20, 50, 100, 250, 500\}$, with their respective prices. Finally, we estimated the cost parameters of the model by measuring the time and resources necessary to carry out all the relevant activities to serve the orders at the distribution center. After evaluating the long run average purchasing and handling costs induced by each case pack size available, we obtained the results displayed in Figure 4-2, where all the costs have been normalized to 1 for confidentiality reasons. From here we concluded that the procurement department should select the case pack with 50 units for this SKU. Moreover, an interesting insight from Figure 4-2 is that the long run average purchasing and handling costs are

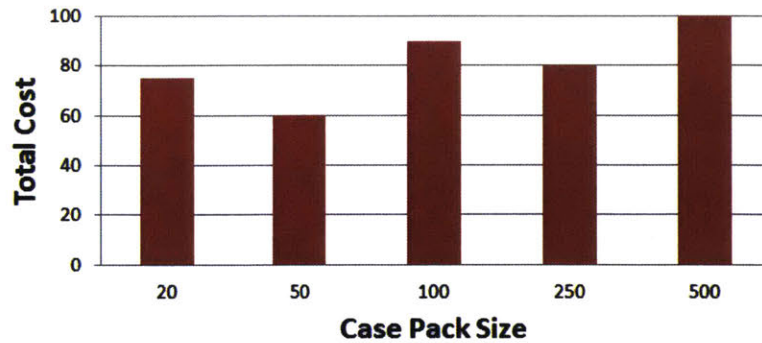


Figure 4-2: Purchasing and Handling Costs Induced per Case Pack Size

not a unimodal function of the case pack size. This makes the problem interesting and challenging. A more detailed account of the costs savings attained by this solution in our model is discussed next.

A summary of the insights provided by testing our model on real data is given in Figure 4-3, where all the costs are normalized to 1 for confidentiality reasons. In each set of results, the first bar corresponds to the total cost, namely the purchasing and handling costs, while the second and third bars correspond to the handling costs and the waste cost respectively. The latter bars are already included in the total cost, but we also display them individually as they are the quantities that the company cares the most about.

The first set of results is the base case, with a case pack size of 250 units and a threshold of 50% of the case pack size, or 125 units. The second set of results answers the question of what happens if we maintain the case pack size of 250 units, but we use the optimal threshold to serve orders at the distribution center, which in this case is 92% of the case pack size, or 230 units. Figure 4-3 shows that the total cost is reduced by 8.6% in this case, where this cost reduction comes from essentially eliminating the waste cost, at the expense of increasing the handling costs by 22.9%. In other words, the model suggests that if the company wants to maintain the same case pack size of 250 units for this SKU, then the distribution center should work harder. This may not be an attractive result for the distribution center, however we have not yet addressed which is the case pack size that the company should select.

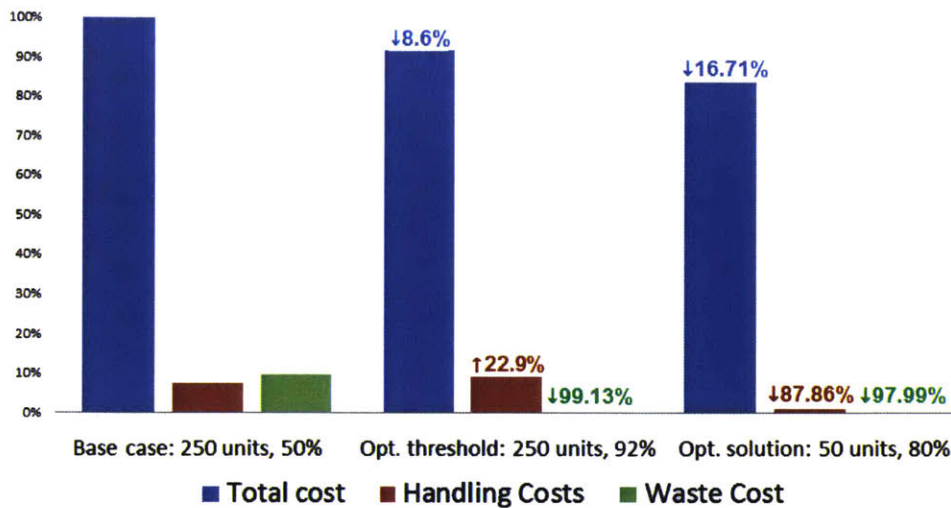


Figure 4-3: Summary of the Experimental Results

In particular, the last set of results addresses the question of what happens when we select the optimal case pack size, which as already discussed is of 50 units, *and* we use the optimal threshold policy to serve orders at the distribution center, which in this case is 80% of the case pack size, or 40 units. Figure 4-3 shows that the total cost is reduced by 16.7%, that is an additional 8% when compared to the second set of results, where this cost reduction is attained again by essentially eliminating the waste cost, but at the same time also significantly reducing the handling costs by 87.8%.

It is interesting to note that, although the model proposes a threshold of 80% of the optimal case pack size to serve orders at the distribution center, which is much larger than the rule of thumb threshold of 50%, the handling costs are so much lower in this case. This is driven by the observation that the new case pack size fits the different demand modes significantly better, so that the number of orders where manual picking is required gets drastically reduced. Similarly, note that for this SKU roughly half of the purchasing and handling costs reduction comes from the fact that the procurement department is choosing the optimal case pack size, while the rest comes from the distribution center implementing the optimal threshold policy.

These experimental results suggest that both the method proposed to serve orders

at the distribution center, as well as the method to select the best case pack size at the procurement department are easy to implement and to communicate in practice. Moreover, the results obtained on real data for the square washer suggest that the cost reductions obtained by implementing these methods can be significant.

4.6 Conclusions and Future Work

In this chapter we have introduced a novel analytical framework to incorporate the cost of handling orders at a central distribution center into the procurement decisions of a company. Specifically, our model explicitly considers the effects of the case pack selection decision in procurement contracts on the purchasing and handling costs of a company. We prove structural results for our model, which lead to a practical method to both select the best case pack size per SKU at the procurement department, and to serve orders at the distribution center.

We tested our method on real data from a large utility company, finding a purchasing and handling costs reduction for one SKU of 16%. These results suggest that the insights provided by our model can be valuable for companies in practice. Additionally, we considered the problem of choosing multiple case pack sizes per SKU. For this problem we showed that, under some assumptions, selecting only three sizes can lead to a 2-approximation with respect to the optimal cost, which in general can require to use every single case pack size offered by the supplier, making it impractical. Moreover, the three sizes that should be selected are intuitive. They correspond to the case pack size that induces the least *per unit* cost of using a whole case pack to satisfy (part of) an order, the case pack size that induces the least *total* cost of using a whole case pack to satisfy (part of) an order, and the case pack size that induces the least long run average cost of picking single units from it.

Future work on this area should include both testing the robustness of the cost reductions presented here for more SKUs, as well as developing heuristics for the considerable harder problem of selecting multiple case pack sizes per SKU.

Conclusions

This thesis introduces several new models in operations management, that are motivated by practical settings, which range from subsidy allocation problems to supply chain procurement. Part I of the thesis studies subsidy allocation problems under budget constraints and endogenous market response. It characterizes sufficient conditions for the optimality of uniform co-payments in maximizing market consumption (Chapter 1), and a very high worst case performance guarantee for a relevant model where uniform co-payments are not optimal (Chapter 2). Additionally, it suggests that simple allocation policies of lump sum subsidies have a good worst case performance guarantee in a different but related model (Chapter 3).

The main insight that we get from Part I is that simple subsidy allocation policies work surprisingly well, as long as there exists market competition among the producers. The subsidy allocation policies that we analyze are motivated by the fact that they are already being implemented by practitioners, and in this sense this thesis moves away from the traditional paradigm in economic theory, and mechanism design, of focusing solely on solutions that attain the first or second best. Through this shift in focus we learned that practical and simple policies can have an unexpected, and many times counter-intuitive, good performance for complex problems.

From a theoretical perspective, the models studied in Part I belong to the class of mathematical programs with equilibrium constraints (MPEC), which are generally very hard to solve and to analyze, both in practice and in theory. In this context, a large portion of the applied research in this area focuses on studying interesting problems which have enough structure that allows to solve them either numerically or, less frequently, analytically. This thesis provides a couple of examples where,

by careful modeling, analytical solutions can be found for a complex model under relatively mild assumptions, which provide structural insights that have an interesting practical interpretation.

Part II of the thesis focuses on supply chain procurement, and proposes a model to incorporate the handling costs incurred at a central distribution center into procurement decisions (Chapter 4). From Part II we get the insight that incorporating the effects that the procurement decisions have in the internal supply chain of a company can have a significant impact on reducing the total purchasing and handling costs. More importantly, the method we propose to incorporate these effects is conceptually simple, and the solutions provided by the model are intuitive. I cannot stress enough how important both these aspects were in convincing a large utility company to provide us with real data from their supply chain to test our model, and to strengthen a collaboration that will hopefully lead to the principles derived from our model to be applied in practice.

More generally, this thesis contributes to the growing research trend that applies the toolkit of operations management and operations research into models that have traditionally fallen beyond the scope of these areas. I strongly believe that plenty of the future meaningful contributions in operations will come from the many more under-explored research areas available, which are an almost limitless source of new and exciting problems to work on.

Appendix A

Proofs

A.1 Chapter 1

Proof of Lemma 2

Proof. Assume, without loss of generality, the first chain of inequalities (1.32). Using Equation (1.31), and given that the co-payments y_i are the same for each scenario, we conclude the second set of inequalities (1.33). From here, $h_i(q_i)$ increasing implies the third set of inequalities (1.34). Summing over all firms gives us the fourth set of inequalities (1.35). Finally, given that the co-payments y_i are the same for each scenario, from the third set of inequalities we get,

$$q_i^1 y_i \geq q_i^2 y_i \geq \dots \geq q_i^M y_i, \text{ for each } i,$$

and summing over all firms gives us the fifth set of inequalities (1.36). ■

Proof of Lemma 3

Proof. The feasible set of problem (*SUBP*) is closed and bounded. It is bounded because $q_i^s \in [0, \check{q}_i]$, for each i, s , where \check{q}_i is such that $h_i(\check{q}_i)\check{q}_i = B$. Similarly, $Q^s \in [0, \check{Q}]$, for each s , where $\check{Q} = \max_{i \in \{1, \dots, n\}} \{\check{q}_i\}$. On the other hand, it is closed because it is defined by inequalities on continuous functions. Additionally, the objective function of problem (*SUBP*) is continuous. It follows that there exists an

optimal solution.

Define the set Γ , as the set of all the optimal solutions to problem (*SUBP*). The set Γ is closed and bounded. It is bounded because it is a subset of the feasible set, which is bounded. On the other hand, denote by z^* the optimal value of the objective function in problem (*SUBP*). Then, the set Γ is closed because it is the intersection of the feasible set, which is closed, and the set $\{(\mathbf{q}^s, Q^s)_{s=1,\dots,m} : \sum_{s=1}^m Q^s p_s \geq z^*\}$, which is closed because the functions Q^s are continuous.

Define the set $X(\Gamma)$, as the set of all the gaps between the maximum marginal cost in scenario $s = 1$, and the minimum marginal cost in scenario $s = 1$, induced by any optimal solution. Namely,

$$X(\Gamma) \equiv \left\{ x \mid \exists (\mathbf{q}^s, Q^s)_{s=1,\dots,m} \in \Gamma \text{ s.t. } x = \max_{i \in \{1,\dots,n\}} \{h_i(q_i^1)\} - \min_{i \in \{1,\dots,n\}} \{h_i(q_i^1)\} \right\}.$$

The set $X(\Gamma)$ is also closed and bounded. Specifically, the maximum and the minimum of continuous functions are continuous, therefore $X(\Gamma)$ is the image of a compact set under a continuous mapping. Hence, $\hat{x} \equiv \min_{x \in X(\Gamma)} x$ is well defined. Namely, the minimum of the gaps between the maximum marginal cost in scenario $s = 1$, and the minimum marginal cost in scenario $s = 1$, induced by any optimal solution, is attained. ■

Proof of Lemma 4

Proof. The modified solution generates the same aggregated market consumption Q^s . Therefore, we only need to check that the budget constraint (1.37) for scenario s , and the non-negativity of the co-payments (1.38) for scenario s , are still satisfied.

Specifically, from $h_i(q_i) = h(g_i q_i)$ with $h(x)$ increasing it follows that $h_i(q_i^s) > h_j(q_j^s)$ implies $g_i q_i^s > g_j q_j^s$. Together with $h(x)$ convex, they imply $(h_i(q_i^s) q_i^s)' = h(g_i q_i^s) + g_i q_i^s h'(g_i q_i^s) > h(g_j q_j^s) + g_j q_j^s h'(g_j q_j^s) = (h_j(q_j^s) q_j^s)'$. It follows that the modified solution has a smaller total cost, $\sum_{j=1}^n h_j(q_j^s) q_j^s$, while generating the same aggregated market consumption Q^s . Hence, it satisfies the scenario s budget constraint (1.37).

Additionally, $(\mathbf{q}^s, Q^s)_{s=1, \dots, m}$ feasible for problem (*SUBP*), and constraint (1.38), imply $h_i(q_i^s) > h_j(q_j^s) \geq P^s(q^s)$. Therefore, $h_i(q_i^s - \epsilon^s) \geq h_j(q_j^s + \epsilon^s) \geq P^s(q^s)$ holds for $\epsilon^s > 0$ sufficiently small. Namely, the modified solution also satisfies the non-negativity of the co-payments (1.38) related to scenario s . ■

Proof of Lemma 5

Proof. First, from $h_i(q_i) = h(g_i q_i)$, it follows that the left hand side of Equation (1.44) is equivalent to $h'_i(q_i^1)\epsilon^1 \leq h'_i(q_i^s - \epsilon^s)\epsilon^s$. Moreover, from this inequality and $h(x)$ convex, it follows that $h_i(q_i^1) - h_i(q_i^1 - \epsilon^1) \leq h'_i(q_i^1)\epsilon^1 \leq h'_i(q_i^s - \epsilon^s)\epsilon^s \leq h_i(q_i^s) - h_i(q_i^s - \epsilon^s)$.

Therefore, on the one hand we have $h_i(q_i^1) - h_i(q_i^1 - \epsilon^1) \leq h_i(q_i^s) - h_i(q_i^s - \epsilon^s)$. On the other hand, from constraint (1.43) it follows that $h_i(q_i^s) - P^s(Q^s) \leq h_i(q_i^1) - P^1(Q^1)$. By adding up these two inequalities we conclude,

$$h_i(q_i^s - \epsilon^s) - P^s(Q^s) \leq h_i(q_i^1 - \epsilon^1) - P^1(Q^1).$$

Second, from $h_i(q_i) = h(g_i q_i)$, it follows that the left hand side of Equation (1.45) is equivalent to $h'_i(q_i^s + \epsilon^s)\epsilon^s = h'_i(q_i^1)\epsilon^1$. Moreover, from this inequality and $h(x)$ convex, it follows that $h_i(q_i^s + \epsilon^s) - h_i(q_i^s) \leq h'_i(q_i^s + \epsilon^s)\epsilon^s = h'_i(q_i^1)\epsilon^1 \leq h_i(q_i^1 + \epsilon^1) - h_i(q_i^1)$.

Therefore, on the one hand we have $h_i(q_i^s + \epsilon^s) - h_i(q_i^s) \leq h_i(q_i^1 + \epsilon^1) - h_i(q_i^1)$. On the other hand, from constraint (1.43) it follows that $h_i(q_i^s) - P^s(Q^s) \leq h_i(q_i^1) - P^1(Q^1)$. By adding up these two inequalities we conclude,

$$h_i(q_i^s + \epsilon^s) - P^s(Q^s) \leq h_i(q_i^1 + \epsilon^1) - P^1(Q^1).$$

■

Proof of Theorem 2

Proof. Let \hat{x} be as defined in the proof of Lemma 3. Namely, let \hat{x} be the minimum of the gaps between the maximum marginal cost in scenario $s = 1$, and the minimum

marginal cost in scenario $s = 1$, induced by any optimal solution. The statement in the Theorem is equivalent to showing $\hat{x} = 0$.

Assume by contradiction that $\hat{x} > 0$. Moreover, denote the optimal solution that induces \hat{x} by $(\hat{\mathbf{Q}}^s, \hat{Q}^s)_{s=1,\dots,m}$. Let the indexes min and max be such that, $h_{\min}(\hat{q}_{\min}^1) \leq h_i(\hat{q}_i^1)$ for each i , and $h_{\max}(\hat{q}_{\max}^1) \geq h_i(\hat{q}_i^1)$ for each i . The assumption $\hat{x} > 0$ is equivalent to $h_{\max}(\hat{q}_{\max}^1) > h_{\min}(\hat{q}_{\min}^1)$. We will show that we can construct an optimal solution $(\tilde{\mathbf{Q}}^s, \tilde{Q}^s)_{s=1,\dots,m}$, such that it induces a strictly smaller gap $\tilde{x} = \max_{i \in \{1,\dots,n\}} \{h_i(\tilde{q}_i^1)\} - \min_{i \in \{1,\dots,n\}} \{h_i(\tilde{q}_i^1)\} < \hat{x}$, contradicting the definition of \hat{x} .

Specifically, from Lemma 4, it follows that if we transfer an arbitrarily small $\epsilon^1 > 0$, from \hat{q}_{\max}^1 to \hat{q}_{\min}^1 , then all the constraints (1.37)-(1.42) related to scenario $s = 1$ are still satisfied. Therefore, this modified solution could only become infeasible due to violating the relaxed non-anticipativity constraints (1.43). We can avoid this infeasibility as follows. We will show that for an arbitrarily small $\epsilon^1 > 0$, and for each scenario $s \neq 1$, there exists $\epsilon^s \geq 0$ such that,

$$h_{\max}(\hat{q}_{\max}^s - \epsilon^s) - P^s(\hat{Q}^s) \leq h_{\max}(\hat{q}_{\max}^1 - \epsilon^1) - P^1(\hat{Q}^1), \quad (\text{A.1})$$

and

$$h_{\min}(\hat{q}_{\min}^s + \epsilon^s) - P^s(\hat{Q}^s) \leq h_{\min}(\hat{q}_{\min}^1 + \epsilon^1) - P^1(\hat{Q}^1). \quad (\text{A.2})$$

Namely, we will show that we can transfer some $\epsilon^s \geq 0$ from \hat{q}_{\max}^s to \hat{q}_{\min}^s , for each scenario $s \neq 1$, such that the modified solution satisfies constraint (1.43). Additionally, we will show that the modified solution also satisfies constraints (1.37)-(1.42), for each scenario $s \neq 1$. Hence, the modified solution is feasible for problem (*SUBP*). Moreover, it is an optimal solution, and it attains a smaller gap than \hat{x} .

From Lemma 5 it follows that, for an arbitrarily small $\epsilon^1 > 0$, and for each scenario $s \neq 1$, it is enough to show that there exists an $\epsilon^s \geq 0$ such that it satisfies the following stronger condition,

$$\frac{h'(g_{\min}(\hat{q}_{\min}^s + \epsilon^s))}{h'(g_{\min}\hat{q}_{\min}^1)} \epsilon^s = \epsilon^1 \leq \frac{h'(g_{\max}(\hat{q}_{\max}^s - \epsilon^s))}{h'(g_{\max}\hat{q}_{\max}^1)} \epsilon^s. \quad (\text{A.3})$$

Specifically, from Equation (1.44) it follows that the inequality in (A.3) implies condition (A.1). Additionally, from Equation (1.45) it follows that the equality in (A.3) implies condition (A.2).

Now we show that for an arbitrarily small $\epsilon^1 > 0$, and for each scenario $s \neq 1$, there exists an $\epsilon^s \geq 0$ such that conditions (A.1) and (A.2) are satisfied, and constraints (1.37)-(1.42) for scenario s are also satisfied. We do so by considering all possible cases. Specifically, if scenario s is such that $h_{\max}(\hat{q}_{\max}^s) - P^s(\hat{Q}^s) < h_{\max}(\hat{q}_{\max}^1) - P^1(\hat{Q}^1)$, then, for an arbitrarily small $\epsilon^1 > 0$, taking $\epsilon^s = 0$ satisfies conditions (A.1) and (A.2), and constraints (1.37)-(1.42) for scenario s , and we are done with this case.

It follows that, without loss of generality, we can focus on an scenario s such that $h_{\max}(\hat{q}_{\max}^s) - P^s(Q^s) = h_{\max}(\hat{q}_{\max}^1) - P^1(\hat{Q}^1)$. From constraint (1.41) it follows that $P^1(Q^1) \geq P^s(Q^s)$. Note that if $P^1(\hat{Q}^1) = P^s(\hat{Q}^s)$, then $h_{\max}(\hat{q}_{\max}^s) = h_{\max}(\hat{q}_{\max}^1) > h_{\min}(\hat{q}_{\min}^1) \geq h_{\min}(\hat{q}_{\min}^s)$, where the last inequality follows from Lemma 2. Therefore, the convexity of $h(x)$ implies that taking an arbitrarily small $\epsilon^s = \epsilon^1 > 0$ satisfies conditions (A.1) and (A.2). Moreover, Lemma 4 ensures that constraints (1.37)-(1.42) for scenario s are also satisfied, and we are done with this case.

Therefore, without loss of generality, assume $P^s(\hat{Q}^s) < P^1(\hat{Q}^1)$. This implies,

$$\frac{h_{\max}(\hat{q}_{\max}^s)}{h_{\min}(\hat{q}_{\min}^s)} \geq \frac{h_{\max}(\hat{q}_{\max}^1) - \overbrace{\left(P^1(\hat{Q}^1) - P^s(\hat{Q}^s)\right)}^{>0}}{h_{\min}(\hat{q}_{\min}^1) - \left(P^1(\hat{Q}^1) - P^s(\hat{Q}^s)\right)} > \frac{h_{\max}(\hat{q}_{\max}^1)}{h_{\min}(\hat{q}_{\min}^1)}.$$

The first inequality follows from $h_{\max}(\hat{q}_{\max}^s) - P^s(Q^s) = h_{\max}(\hat{q}_{\max}^1) - P^1(\hat{Q}^1)$ and constraint (1.43). The second inequality follows from $h_{\max}(\hat{q}_{\max}^1) > h_{\min}(\hat{q}_{\min}^1)$. Hence, from $h_i(q_i) = h(g_i q_i)$, the fact that $h(x)$ satisfies Property 2, and the strict inequality above, we conclude that scenario s satisfies,

$$\frac{h'(g_{\max} \hat{q}_{\max}^s)}{h'(g_{\max} \hat{q}_{\max}^1)} > \frac{h'(g_{\min} \hat{q}_{\min}^s)}{h'(g_{\min} \hat{q}_{\min}^1)}. \quad (\text{A.4})$$

From Equation (A.4) it follows that the stronger condition (A.3) is satisfied for $\epsilon^s > 0$

sufficiently small. Therefore, conditions (A.1) and (A.2) hold, and we are done with this case. This completes the analysis of all possible cases.

To summarize, we have shown that there exist $\epsilon^1 > 0$, and $\epsilon^s \geq 0$ for each scenario $s \neq 1$, such that the modified solution $(\tilde{\mathbf{q}}^s, \tilde{Q}^s)_{s=1, \dots, m}$, defined by,

$$\tilde{q}_{\min}^s = \hat{q}_{\min}^s + \epsilon^s, \text{ for each } s \in \{1, \dots, m\},$$

$$\tilde{q}_{\max}^s = \hat{q}_{\max}^s - \epsilon^s, \text{ for each } s \in \{1, \dots, m\},$$

$$\tilde{q}_i^s = \hat{q}_i^s, \text{ for each } i \notin \{\min, \max\}, s \in \{1, \dots, m\}.$$

is an optimal solution to the upper bound problem (*SUBP*). Specifically, it is feasible and it attains the same objective value as the optimal solution $(\hat{\mathbf{q}}^s, \hat{Q}^s)_{s=1, \dots, m}$. Moreover, by potentially repeating this argument for the finite number of pair of indexes $i, j \in \{1, \dots, n\}$, we conclude that its gap $\tilde{x} = \max_{i \in \{1, \dots, n\}} \{h_i(\tilde{q}_i^1)\} - \min_{i \in \{1, \dots, n\}} \{h_i(\tilde{q}_i^1)\}$ is strictly smaller than \hat{x} . This contradicts the definition of \hat{x} .

Hence, we conclude that $\hat{x} = 0$. Therefore, $h_i(\hat{q}_i^1) = h_j(\hat{q}_j^1)$, for each i, j . Or equivalently, $h_i(\hat{q}_i^1) - P^1(\hat{Q}^1) = y^1$ for each $i \in \{1, \dots, n\}$. ■

Proof of Theorem 3

Proof. We will show that there exists an optimal solution to the upper bound problem (*SUBP*) that induces uniform co-payments. Moreover, this solution is feasible for the co-payment allocation problem under market state uncertainty (*SCAP*). Therefore, uniform co-payments are optimal for problem (*SCAP*).

From Theorem 2 it follows that there exists an optimal solution to the upper bound problem (*SUBP*) $(\hat{\mathbf{q}}^s, \hat{Q}^s)_{s=1, \dots, m}$ such that $h_i(\hat{q}_i^1) - P^1(\hat{Q}^1) = y^1$ for each i . We will show first that there exists an optimal solution for problem (*SUBP*), $(\tilde{\mathbf{q}}^s, \tilde{Q}^s)_{s=1, \dots, m}$, such that $h_i(\tilde{q}_i^s) - P^s(\tilde{Q}^s) = y^s$ for each i , for each scenario $s \neq 1$, for some value $y^s > 0$. Then, we will conclude by showing that we must have $y^s = y^1$ for each s .

Plugging in y^1 in the budget constraint for scenario $s = 1$ we obtain $y^1 \leq \frac{B}{\hat{Q}^1}$. Moreover, for this solution we can decompose the upper bound problem (*SUBP*) for

each scenario $s \neq 1$, and obtain the following independent problem,

$$\begin{aligned} \min_{\mathbf{q}, Q} \quad & Qp_s \\ \text{s.t.} \quad & \sum_{j=1}^n q_j h_j(q_j) - P^s(Q)Q \leq B \end{aligned} \quad (\text{A.5})$$

$$h_i(q_i) \geq P^s(Q), \text{ for each } i \in \{1, \dots, n\} \quad (\text{A.6})$$

$$(\text{SLBP} - s) \quad \sum_{j=1}^n q_j = Q \quad (\text{A.7})$$

$$q_i \geq 0, \text{ for each } i \in \{1, \dots, n\} \quad (\text{A.8})$$

$$P^1(\hat{Q}^1) \geq P^s(Q) \quad (\text{A.9})$$

$$\hat{Q}^1 \geq Q \quad (\text{A.10})$$

$$h_i(q_i) - P^s(Q) \leq y^1, \text{ for each } i \in \{1, \dots, n\}. \quad (\text{A.11})$$

It follows that the components of the optimal solution to the upper bound problem (*SUBP*) corresponding to scenario s , $(\hat{\mathbf{q}}^s, \hat{Q}^s)$, must be an optimal solution for problem (SLBP- s) as well. Note that the budget constraint (A.5) is redundant for this problem. Specifically, we have,

$$\sum_{i=1}^n q_i h_i(q_i) - P^s(Q)Q \leq Qy^1 \leq Q \frac{B}{\hat{Q}^1} \leq B.$$

The first inequality follows from constraint (A.11), the second inequality follows from $y^1 \leq \frac{B}{\hat{Q}^1}$, and the third inequality follows from constraint (A.10). Therefore, without loss of generality, we can drop the budget constraint in scenario $s \neq 1$ (A.5).

Exactly as in Lemma 3, the feasible set of problem (SLBP- s) is closed and bounded, and its objective function is continuous. It follows that there exists an optimal solution. Now we show that there exists an optimal solution for problem (SLBP- s), $(\hat{\mathbf{q}}^s, \hat{Q}^s)$, such that $h_i(\hat{q}_i^s) - P^s(\hat{Q}^s) = y^s$, for each i , for some value $y^s > 0$. Specifically, assume by contradiction that this is not the case. It follows that there must exist indexes min and max such that $h_{\min}(\hat{q}_{\min}^s) \leq h_i(\hat{q}_i^s)$ for each i , $h_{\max}(\hat{q}_{\max}^s) \geq h_i(\hat{q}_i^s)$ for each i , and $h_{\min}(\hat{q}_{\min}^s) < h_{\max}(\hat{q}_{\max}^s)$. On the other hand, let $\tilde{\mathbf{q}}^s$ be the optimal

solution to the following optimization problem.

$$\begin{aligned} \min_{\mathbf{q}} \quad & \sum_{j=1}^n q_j h_j(q_j) \\ \text{s.t.} \quad & \sum_{j=1}^n q_j = \hat{Q}^s \\ & q_i \geq 0, \text{ for each } i \in \{1, \dots, n\} \end{aligned}$$

We show that $(\tilde{\mathbf{q}}^s, \hat{Q}^s)$ is feasible for problem (SLBP-s). Because budget constraint (A.5) is redundant, and the aggregated market consumption \hat{Q}^s is fixed, it follows that we only need to check that constraints (A.6) and (A.11) are satisfied. From $h_i(q_i) = h(g_i q_i)$, and $h(x)$ convex and increasing, it follows that the objective function of this problem is convex. The first order conditions are $(h_i(\tilde{q}_i^s) \tilde{q}_i^s)' = (h_j(\tilde{q}_j^s) \tilde{q}_j^s)'$ for each i, j . Moreover, because $h(x)$ satisfy Property 2, we conclude $h_i(\tilde{q}_i^s) = h_j(\tilde{q}_j^s)$ for each i, j .

Additionally, we claim that $h_{\min}(\hat{q}_{\min}^s) < h_i(\tilde{q}_i^s) < h_{\max}(\hat{q}_{\max}^s)$ for each i . In fact, if $h_{\max}(\hat{q}_{\max}^s) > h_{\min}(\hat{q}_{\min}^s) \geq h_i(\tilde{q}_i^s)$, for each i , then we must have, $\sum_{j=1}^n \tilde{q}_j^s < \sum_{j=1}^n \hat{q}_j^s = \hat{Q}^s$. This is a contradiction to the feasibility of solution $(\tilde{\mathbf{q}}^s, \hat{Q}^s)$. Similarly, if $h_i(\tilde{q}_i^s) \geq h_{\max}(\hat{q}_{\max}^s) > h_{\min}(\hat{q}_{\min}^s)$, for each i , then we must have, $\sum_{j=1}^n \tilde{q}_j^s > \sum_{j=1}^n \hat{q}_j^s = \hat{Q}^s$. This is a contradiction to the feasibility of solution $(\tilde{\mathbf{q}}^s, \hat{Q}^s)$. This implies, together with the feasibility of $(\hat{\mathbf{q}}^s, \hat{Q}^s)$ for problem (SLBP-s), that,

$$h_i(\tilde{q}_i^s) > h_{\min}(\hat{q}_{\min}^s) \geq P^s(\hat{Q}^s), \text{ for each } i,$$

and,

$$h_i(\tilde{q}_i^s) - P^s(\hat{Q}^s) < h_{\max}(\hat{q}_{\max}^s) - P^s(\hat{Q}^s) \leq y^1, \text{ for each } i.$$

Namely, constraints (A.6) and (A.11) are satisfied. Therefore, $(\tilde{\mathbf{q}}^s, \hat{Q}^s)$ is feasible for problem (SLBP-s). Moreover, it attains the same objective value than $(\hat{\mathbf{q}}^s, \hat{Q}^s)$, therefore it is also optimal. Finally, from $h_i(\tilde{q}_i^s) = h_j(\tilde{q}_j^s)$ for each i, j , it follows that $h_i(\tilde{q}_i^s) - P^s(\hat{Q}^s) = y^s$ for each $i \in \{1, \dots, n\}$ for some value $y^s > 0$.

Finally, we show that we must have $y^s = y^1$ for each scenario s . From $h_i(q_i) =$

$h(g_i q_i)$, it follows that, for any given value of $y^s \geq 0$, \hat{Q}^s is uniquely determined by the solution of the equation,

$$\hat{Q}^s(y^s) = \sum_{i=1}^n \frac{h^{-1}\left(P^s(\hat{Q}^s(y^s)) + y^s\right)}{g_i}.$$

It follows that, $\hat{Q}^s(y^s)$ is increasing in y^s . Assume by contradiction that $y^s < y^1$, then we can increase y^s by $\epsilon > 0$ sufficiently small, and obtain a strictly better objective value while keeping feasibility. In fact, the only constraint that might prevent this increase is the budget constraint (A.5), which is not tight. This contradicts the optimality of $(\tilde{\mathbf{q}}^s, \hat{Q}^s)$.

We have shown that there exists an optimal solution to the upper bound problem (*SUBP*) $(\tilde{\mathbf{q}}^s, \hat{Q}^s)_{s=1, \dots, m}$ such that, $h_i(\hat{q}_i^s) - P^s(\hat{Q}^s) = y^1$ for each $i \in \{1, \dots, n\}$, and for each $s \in \{1, \dots, m\}$, for some value $y^1 \geq 0$. That is, it satisfies the relaxed non-anticipativity constraints with equality. Therefore, it is feasible in the co-payment allocation problem under market state uncertainty (*SCAP*). Hence, uniform co-payments are optimal for problem (*SCAP*). ■

A.2 Chapter 2

Proof of Proposition 5

Proof. The proof follows from the *KKT* conditions of problem (*CAP*).

Problem (*CAP*) is a convex optimization problem, therefore the *KKT* conditions are necessary and sufficient for optimality. Let $\lambda \geq 0$, $\delta_i \geq 0$, γ , and $\theta_i \geq 0$ be the dual variables associated with constraints (2.8), (2.9), (2.10), and (2.11), respectively. Let (\mathbf{q}^*, Q^*) be an optimal solution to problem (*CAP*). From the market equilibrium condition (2.7) it follows that the optimal co-payments are $y_i^* = c_i + b q_i^* - a - b Q^*$.

First, note that the budget constraint (2.8) must be binding for solution (\mathbf{q}^*, Q^*) . Specifically, assume for a contradiction that the budget constraint (2.8) is not binding, then we can increase some q_i^* , and Q^* , by $\epsilon > 0$ small enough, maintain the feasibility

for problem (CAP), and obtain an strictly larger objective value, contradicting the optimality of solution (\mathbf{q}^*, Q^*) . This implies that, for any positive budget $B > 0$, at least one firm i must have a positive market output $q_i^* > 0$, and a positive co-payment $y_i^* = c_i + bq_i^* - a - bQ^* > 0$.

Second, the *KKT* conditions for problem (CAP) imply that

$$\lambda(c_i + 2bq_i^*) - b\delta_i + \gamma - \theta_i = 0, \text{ for each } i \in \{1, \dots, n\}. \quad (\text{A.12})$$

On the other hand, the complementary slackness conditions imply $\delta_i y_i^* = 0$ and $\theta_i q_i^* = 0$, for each $i \in \{1, \dots, n\}$. Combining these with equation (A.12) we conclude that

$$\text{For each } i, j, \text{ if } q_i^* > 0, y_i^* > 0 \text{ then } bq_i^* + y_i^* = -\frac{\gamma + a - bQ^*}{\lambda} \leq bq_j^* + y_j^*. \quad (\text{A.13})$$

Together with the market equilibrium condition (2.7), equation (A.13) implies the following two conditions

$$\text{For each } i, j, \text{ if } q_i^* > 0, y_i^* > 0 \text{ then } y_i^* + \frac{c_j - c_i}{2} \leq y_j^*. \quad (\text{A.14})$$

$$\text{For each } i, j, \text{ if } q_i^* > 0, y_i^* > 0 \text{ then } q_i^* - \frac{c_j - c_i}{2b} \leq q_j^*. \quad (\text{A.15})$$

Where equations (A.14) and (A.15) must hold with equality for any firm j such that $q_j^* > 0$ and $y_j^* > 0$.

Recall that we have assumed, without loss of generality, that $c_1 \leq c_2 \leq \dots \leq c_n \leq a$. Therefore, from equation (A.14) it follows that $y_i^* > 0$ implies $y_j^* > 0$ for each $j \geq i$. Similarly, from equation (A.15) it follows that $q_i^* > 0$ implies $q_j^* > 0$ for each $j \leq i$. Additionally, we have already shown that at least one firm i must have a positive market output $q_i^* > 0$, and a positive co-payment $y_i^* > 0$. Hence, without loss of generality, there exist indexes $l, m \in \{1, \dots, n\}$, $l \leq m$, such that $y_i^* = 0$ for each $i \in \{1, \dots, l-1\}$, $y_i^* > 0$ for each $i \in \{l, \dots, n\}$, and $q_i^* > 0$ for each $i \in \{1, \dots, m\}$, $q_i^* = 0$ for each $i \in \{m+1, \dots, n\}$.

This completes the proof of equations (2.12)-(2.17). Specifically, equations (2.12) and (2.17) follow from the definition of indexes l and m respectively. Equations (2.13) and (2.16) follow from taking $i = l$ in equations (A.14) and (A.15), and the fact that these equations must hold with equality for any firm j such that $q_j^* > 0$ and $y_j^* > 0$, which is the case for any $j \in \{l, \dots, m\}$. Similarly, the equality in equation (2.14) follows from the market equilibrium condition (2.7) and $q_i^* = 0$, for each $i \in \{m+1, \dots, n\}$, namely, $a - bQ^* = c_i - y_i^* = c_l + bq_l^* - y_i^*$, for each $i \in \{m+1, \dots, n\}$; while the inequality in equation (2.14) follows from equation (A.15). Finally, the equality in equation (2.15) follows from the market equilibrium condition (2.7) and $y_i^* = 0$, for each $i \in \{1, \dots, l-1\}$, namely, $a - bQ^* = c_i + bq_i^* = c_l + bq_l^* - y_i^*$, for each $i \in \{1, \dots, l-1\}$; while the inequality in equation (2.15) follows from equation (A.14).

Now we show equation (2.18). Adding up equations (2.15) to (2.17) we get

$$Q^* = mq_l^* + \sum_{i=1}^{l-1} \frac{c_l - c_i}{b} - \sum_{i=l+1}^m \frac{c_i - c_l}{2b} - \frac{(l-1)}{b} y_l^*. \quad (\text{A.16})$$

On the other hand, from the market equilibrium condition (2.7) we have that $a - bQ^* = c_l + bq_l^* - y_l^*$. Plugging in the expression for Q^* from equation (A.16), and solving for q_l^* , we obtain the expression for q_l^* given in equation (2.18).

Now we show equation (2.19). We have already argued that the budget constraint (2.8) must be binding for solution (\mathbf{q}^*, Q^*) . Note that the budget constraint (2.8) is equivalent to $\sum_{i=1}^n q_i^* y_i^* = B$. Therefore, from equations (2.12)-(2.17) it follows that the budget constraint (2.8) can be written as $\sum_{i=l}^m \left(q_i^* - \frac{c_i - c_l}{2b} \right) \left(y_i^* + \frac{c_i - c_l}{2} \right) = B$. Plugging in the expression for q_i^* given in equation (2.18) we conclude that the budget constraint (2.8) is equivalent to

$$\sum_{i=l}^m \left(\frac{a + \sum_{j=1}^m c_j}{(m+1)b} - \frac{c_l}{b} - \sum_{j=l+1}^m \frac{c_j - c_l}{2(m+1)b} + \frac{ly_l^*}{(m+1)b} - \frac{c_i - c_l}{2b} \right) \left(y_i^* + \frac{c_i - c_l}{2} \right) = B.$$

This is a quadratic equation on y_l^* , whose positive root is given by equation (2.19).

Finally, plugging in the expression for q_l^* from equation (2.18), and the expres-

sion for y_l^* from equation (2.19), in equation (A.16) and simplifying, we obtain the expression for Q^* given in equation (2.20). ■

Proof of Lemma 6

Proof. By the definition of index l , for each $i \in \{1, \dots, l-1\}$ we have $c_i + bq_i^* = a - bQ^*$, which implies $q_i^* = \frac{a}{b} - Q^* - \frac{c_i}{b}$.

On the other hand, from equation (2.15) it follows that $y_l^* \leq \frac{c_l - c_{l-1}}{2}$. Moreover, together with equation (2.13), this implies $y_i^* \leq \frac{c_i - c_{l-1}}{2}$, for each $i \in \{l, \dots, m\}$. Therefore, for each $i \in \{l, \dots, m\}$ we have $c_i + bq_i^* = a - bQ^* + y_i^* \leq a - bQ^* + \frac{c_i - c_{l-1}}{2}$, which implies $q_i^* \leq \frac{a}{b} - Q^* - \frac{c_i}{2b} - \frac{c_{l-1}}{2b}$.

Adding up these inequalities, for each $i \in \{1, \dots, m\}$, we get $Q^* \leq \frac{ma}{b} - mQ^* - \sum_{i=1}^{l-2} \frac{c_i}{b} - \sum_{i=l}^n \frac{c_i}{2b} - (m-l+1) \frac{c_{l-1}}{2}$, which is equivalent to the upper bound in equation (2.21). ■

Proof of Lemma 7

Proof. Recall that we have assumed, without loss of generality, that $c_1 \leq c_2 \leq \dots \leq c_n \leq a$. From the definition of the index u , it follows that at the market equilibrium, we must have $q_i^U = 0$, for each $i \in \{u+1, \dots, n\}$, as in equation (2.23). Additionally, from the definition of the index u it follows that, for each firm $i \in \{1, \dots, u\}$, we have that $c_i \leq a - bQ^U + \frac{B}{Q^U}$. Therefore, at the market equilibrium we must have $c_i + bq_i^U = a - bQ^U + \frac{B}{Q^U}$, which is equivalent to equation (2.22).

Finally, adding up equation (2.22) for each firm that participates in the market equilibrium, we obtain $b(u+1)(Q^U)^2 - (ua - \sum_{i=1}^u c_i)Q^U = uB$. To conclude, note that equation (2.24) is exactly the positive root of this quadratic equation. ■

Proof of Lemma 8

Proof. At the market equilibrium induced by the uniform co-payments allocation we must have $c_u + bq_u^U = a - bQ^U + \frac{B}{Q^U} = c_i + bq_i^U$, for each $i \in \{1, \dots, u\}$. This implies $\frac{c_u - c_i}{b} = q_i^U - q_u^U$, for each $i \in \{1, \dots, u\}$. Adding up over all $i \in \{1, \dots, u\}$ we get $\frac{uc_u - \sum_{i=1}^u c_i}{b} = Q^U - uq_u^U \leq Q^U$, where the inequality becomes tight when $q_u^U = 0$. ■

Proof of Lemma 9

Proof. For any instance of the co-payments allocation problem (CAP) with $c_1 \geq 0$, it is enough to consider the modified instance where the demand parameter is $\hat{a} = (a + \delta - c_1)$, and the marginal costs vector is $\hat{\mathbf{c}} = (\mathbf{c} + (\delta - c_1)\mathbf{e})$, where \mathbf{e} is a vector of ones. Specifically, any feasible solution in the original instance, (\mathbf{q}, Q) , is feasible in the modified instance, and it attains the same objective value, and viceversa. We only need to check constraints (2.8) and (2.9). For constraint (2.9), note that $\hat{c}_i + bq_i - \hat{a} + bQ = c_i + bq_i - a + bQ$. Finally, multiplying both sides by q_i , and adding up over all i , we conclude that for the right hand side of constraint (2.8) we have $\sum_{i=1}^n (\hat{c}_i q_i + bq_i^2) - (\hat{a} - bQ)Q = \sum_{i=1}^n (c_i q_i + bq_i^2) - (a - bQ)Q$. ■

Proof of Lemma 10

Proof. For any given number of firms in the market $n \geq 2$, there is a finite set of possible combinations of the indexes $l, m, u \in \{1, \dots, n\}$, $l \leq m$. For each given combination, the ratio Q^U/Q^* has a closed form expression, which is a continuous function of the problem parameters: the number of firms in the market n , the demand parameters a, b , the marginal cost of each firm c_i , and the budget B .

Moreover, from Lemma 9 we can assume, without loss of generality, that $c_1 = 0$. From this, together with $c_i \leq c_{i+1}$, for each $i \in \{1, \dots, n - 1\}$, and the consistency constraints (2.27)-(2.30) it follows that the feasible set of the problem parameters is closed and bounded. Hence, we conclude that there exists an instance of problem (CAP) that minimizes the ratio Q^U/Q^* . ■

Proof of Proposition 6

Proof. For any given number of firms in the market $n \geq 2$, consider any instance of problem (CAP) a, b, \mathbf{c} , and B that minimizes the ratio Q^U/Q^* . From Proposition 5 and Lemma 7 this instance induces indexes l, m and u respectively. Let (\mathbf{q}^U, Q^U) be the solution induced by uniform co-payments in this instance.

First, assume by contradiction that $u < m$. Recall that, by definition, u is the smallest index such that $c_i > a - bQ^U + \frac{B}{Q^U}$, for each $i \in \{u + 1, \dots, n + 1\}$. Then,

at the market equilibrium induced by uniform co-payments we have $c_u + bq_u^U = a - bQ^U + \frac{B}{Q^U} < c_m$. From this, together with $q_u^U \geq 0$, it follows that $c_u < c_m$. Let \hat{i} be the first index such that $\hat{i} \geq l$, and $c_u < c_{\hat{i}}$. Again, by the definition of the index u we must have $c_{\hat{i}} > a - bQ^U + \frac{B}{Q^U}$. It follows that we can reduce the value of $c_{\hat{i}}$, by $\epsilon > 0$ sufficiently small, without affecting the uniform co-payments solution (\mathbf{q}^U, Q^U) .

On the other hand, by reducing the value of $c_{\hat{i}}$, by $\epsilon > 0$ sufficiently small, we obtain a strictly larger value for the aggregated consumption induced by optimal co-payments. Specifically, let (\mathbf{q}^*, Q^*) be an optimal solution to problem (CAP) for the original instance. Consider the modified solution $(\hat{\mathbf{q}}, Q^* + \delta)$, where $\hat{q}_i = q_i^*$ for each $i \neq \hat{i}$, and $\hat{q}_{\hat{i}} = q_{\hat{i}}^* + \delta$, where $\delta = \frac{\epsilon - \gamma}{2b} > 0$, for some $\gamma \in (0, \epsilon)$, close enough to $\epsilon > 0$. This solution is feasible for the modified instance of problem (CAP), and attains an objective value strictly larger than Q^* .

To check the feasibility of solution $(\hat{\mathbf{q}}, Q^* + \delta)$, we only need to check constraints (2.8) and (2.9). In constraint (2.9), for each $i \neq \hat{i}$, note that the left hand side remains constant, while the right hand side strictly decreases by $b\delta = \frac{\epsilon - \gamma}{2} > 0$, for any $\gamma \in (0, \epsilon)$, and any $\epsilon > 0$, therefore these constraints are satisfied by solution $(\hat{\mathbf{q}}, Q^* + \delta)$ for the modified instance. Similarly, in constraint (2.9) for \hat{i} , note that the left hand side strictly decreases by $b\delta + \gamma$, while the right hand side strictly decreases by $b\delta$. Namely, the co-payment allocated to firm \hat{i} in the modified instance strictly decreases by $\gamma > 0$ with respect to the co-payment allocated to firm \hat{i} in the original instance. On the other hand, $l \leq \hat{i} \leq m$ implies that the co-payment allocated to firm \hat{i} in the original instance is strictly positive, therefore, for any $\gamma > 0$ small enough this constraint will still be satisfied. Finally, note that the right hand side of constraint (2.8) is equal to $B + \delta(b(Q^* - q_{\hat{i}}^*) + y_{\hat{i}}^* - \gamma) - \gamma q_{\hat{i}}^* \leq B$, where the inequality holds for any $\delta = \frac{\epsilon - \gamma}{2b} > 0$ small enough, which can be attained by some $\gamma \in (0, \epsilon)$ close enough to ϵ , for any arbitrarily small $\epsilon > 0$. For the last inequality we have also used the fact that $l \leq \hat{i} \leq m$ implies that $q_{\hat{i}}^* > 0$. Hence, constraint (2.8) is satisfied by solution $(\hat{\mathbf{q}}, Q^* + \delta)$ in the modified instance.

We have shown that if $u < m$ then there exists an index \hat{i} , with $u < \hat{i} \leq m$, such that when decreasing the value of $c_{\hat{i}}$ by $\epsilon > 0$ sufficiently small, the aggregated

consumption induced by uniform co-payments, Q^U , remains constant, while the value of the aggregated consumption induced by optimal co-payments, Q^* , strictly increases. Therefore, the relative performance of uniform co-payments strictly decreases with respect to the original instance, a contradiction. It follows that any instance of problem (CAP) that minimizes the ratio Q^U/Q^* must be the case that $l \leq m \leq u$.

Second, assume by contradiction that in the original instance we have $u > m$. Recall that, by definition, m is the largest index of a firm that participates in the market equilibrium induced by the optimal co-payments. Then, it must be the case that $c_u > c_m$. It follows that we can discard, without loss of generality, any firm with index $i \geq u + 1$, because they do not participate in the equilibria under consideration (uniform co-payments and optimal). Moreover, we will assume, without loss of generality, that $q_u^U > 0$, otherwise we can discard firm u as well and analyze the instance with $n = u - 1$ firms.

It follows that we can increase the value of c_u by $\epsilon > 0$ sufficiently small, without changing the optimal co-payments solution (\mathbf{q}^*, Q^*) . On the other hand, increasing the value of c_u by $\epsilon > 0$ sufficiently small decreases the aggregated consumption induced by the optimal co-payments Q^U . Specifically, from equation (2.24) it follows that,

$$\frac{\partial Q^U}{\partial c_i} = -\frac{Q^U}{\sqrt{\left(ua - \sum_{j=1}^u c_j\right)^2 + 4u(u+1)bB}} < 0, \text{ for each } i \in \{1, \dots, u\}.$$

We have shown that when increasing the value of c_u by $\epsilon > 0$ sufficiently small, the aggregated consumption induced by optimal co-payments, Q^* , remains constant, while the value of the aggregated consumption induced by uniform co-payments, Q^U , strictly decreases. Therefore, the relative performance of uniform co-payments strictly decreases with respect to the original instance, a contradiction. It follows that for any instance of problem (CAP) that minimizes the ratio Q^U/Q^* , it must be the case that $l \leq m = u = n$. ■

Proof of Proposition 7

Proof. For any given number of firms in the market $n \geq 2$, consider any instance of problem (CAP) a, b, c , and B that minimizes the ratio Q^U/Q^* . From Lemma 9, it follows that we can assume, without loss of generality, that $c_1 > 0$.

Let (\mathbf{q}^U, Q^U) be the solution induced by uniform co-payments in this instance. Assume for a contradiction that $q_n^U > 0$. Then, from the market equilibrium condition we get that $c_n < c_n + bq_n^U = a - bQ^U + \frac{B}{Q^U}$. From this, together with the expression for Q^U given in equation (2.24), it follows that we can increase the value of c_n and reduce the value of c_1 by the same $\epsilon > 0$ sufficiently small, without affecting the aggregated consumption induced by uniform co-payments Q^U .

On the other hand, by increasing the value of c_n and reducing the value of c_1 by the same $\epsilon > 0$ sufficiently small, we obtain a strictly larger value for the aggregated consumption induced by optimal co-payments. Specifically, let (\mathbf{q}^*, Q^*) be an optimal solution to the problem (CAP) defined by the original instance. Consider the modified solution $(\hat{\mathbf{q}}, Q^* + \gamma)$, where $\hat{q}_1 = q_1^* + \delta + \gamma$, $\hat{q}_i = q_i^*$ for each $i \in \{2, \dots, n-1\}$, and $\hat{q}_n = q_n^* - \delta$, for $\delta > 0$ and $\gamma > 0$ such that $\epsilon = b(\delta + 2\gamma)$, where δ is close enough to $\frac{\epsilon}{b} > 0$ and γ is arbitrarily smaller than δ . This solution is feasible for the modified instance of problem (CAP), and attains an objective value strictly larger than Q^* .

To check the feasibility of solution $(\hat{\mathbf{q}}, Q^* + \gamma)$ for the modified instance, we only need to check constraints (2.8) and (2.9). In constraint (2.9) for $i = 1$, note that both the left hand side and the right hand side strictly decrease by $b\gamma > 0$, therefore this constraint is still satisfied. In constraint (2.9), for each $i \in \{2, \dots, n-1\}$, note that the left hand side remains constant, while the right hand side strictly decreases by $b\gamma > 0$, therefore these constraints are satisfied by solution $(\hat{\mathbf{q}}, Q^* + \delta)$. Similarly, in constraint (2.9) for $i = n$, note that the left hand side strictly increases by $2b\gamma > 0$, while the right hand side strictly decreases by $b\gamma > 0$, therefore this constraint is still satisfied. Finally, note that the left hand side of the budget constraint (2.8) is equal to $B + \gamma(b(Q^* - q_1^*) + 2bq_n^* + y_1^* - 3b\delta) - \delta(y_n^* - y_1^*)$, which is less than the budget B , for any $\gamma > 0$ arbitrarily smaller than δ , and $\delta > 0$ close enough to $\frac{\epsilon}{b} > 0$, for any arbitrarily small $\epsilon > 0$. For the last inequality we have also used the fact that $(y_n^* - y_1^*) > 0$, which follows from the following observation: from Proposition 5 and

equation (2.12) it follows that if $l \geq 2$, then $y_1^* = 0$, therefore, $(y_n^* - y_1^*) = y_n^* > 0$. Similarly, if $l = 1$, then from equation (2.13) it follows that $(y_n^* - y_1^*) = \frac{c_n - c_1}{2} > 0$, where the last inequality follows from the fact that if $c_1 = c_n$ then all the firms are homogeneous and uniform co-payments are clearly optimal. Hence, constraint (2.8) is satisfied by solution $(\hat{\mathbf{q}}, Q^* + \gamma)$ in the modified instance.

We have shown that when increasing the value of c_n and reducing the value of c_1 by the same $\epsilon > 0$ sufficiently small, the aggregated consumption induced by uniform co-payments, Q^U , remains constant, while the value of the aggregated consumption induced by optimal co-payments Q^* strictly increases. Therefore, the relative performance of uniform co-payments strictly decreases with respect to the original instance, a contradiction. It follows that in for any instance of problem (CAP) that minimizes the ratio Q^U/Q^* it must be the case that $q_n^U = 0$. ■

Proof of Proposition 8

Proof. The formulation of problem (WCP_l) follows by defining the function

$$Q^U(\mathbf{c}) \equiv \frac{nc_n - \sum_{i=1}^n c_i}{b}, \quad (\text{A.17})$$

where this expression for the market consumption induced by uniform co-payments, Q^U , follows from Lemma 8, Proposition 6 and Proposition 7. Specifically, the right hand side of equation (A.17) is the lower bound from Lemma 8, which we can write assuming $u = n$ due to Proposition 6, and with an equality by Proposition 7.

As well as defining the function

$$\begin{aligned}
& Q_l^*(B, \mathbf{c}) \\
\equiv & \frac{1}{2l(n+1)b} \left((2ln - n + l - 1)a - (n + l + 1) \sum_{i=1}^{l-1} c_i - lc_l - l \sum_{i=l+1}^n c_i \right) \\
& + \frac{n-l+1}{2l(n+1)b} \left(\left(a + \sum_{i=1}^{l-1} c_i - lc_l \right)^2 + \frac{l(n+1)}{n-l+1} \sum_{i=l+1}^n (c_i - c_l)^2 + \frac{4l(n+1)}{n-l+1} bB \right. \\
& \left. - \frac{l}{n-l+1} \left(\sum_{i=l+1}^n (c_i - c_l) \right) \left(2a + 2 \sum_{i=1}^{l-1} c_i + \sum_{i=l+1}^n c_i - (n+l)c_l \right) \right)^{\frac{1}{2}}, \quad (\text{A.18})
\end{aligned}$$

where the expression for the market consumption induced by optimal co-payments, Q^* , follows from equation (2.20) in Proposition 5. Note that we have parametrized this expression by l , the index of the first firm that receives a positive co-payment in the optimal co-payment allocation policy. We have also assumed $m = n$ from Proposition 6.

Finally, from Lemma 9 it follows that we can assume constraint (2.32) without loss of generality. Additionally, constraints (2.35), (2.36) and (2.37) correspond to the consistency constraints (2.27), (2.29) and (2.30), respectively. Finally, from Proposition 6 it follows that, without loss of generality, we can drop the consistency constraint (2.28), because $u = n$ implies that there is no firm with index $(u + 1)$. ■

Lemma 15. *For any number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, the instance defined by equations (2.41)-(2.43) is feasible for the problem (WCP_2) .*

Proof. First, we check that constraint (2.34) is satisfied. Note, from equation (2.42), that for any $n \geq 2$, and $a \geq 0$, we have that $0 \leq c_n = \left(\frac{n + \sqrt{\frac{n(n+1)}{2}}}{3n+1} \right) a \leq \frac{2n+1}{3n+1} a < a$, where the second inequality follows from $\frac{n(n+1)}{2} \leq (n+1)^2$.

Second, we check that constraint (2.35) is satisfied. Note, from equations (2.41)

and (2.42), that for this instance the right hand side of constraint (2.35) evaluates to

$$c_n(2c_n - a) = \left(\frac{n + \sqrt{\frac{n(n+1)}{2}}}{3n + 1} \right) \left(\frac{2n + 2\sqrt{\frac{n(n+1)}{2}}}{3n + 1} - 1 \right) a^2 = \left(\frac{(n-1)\sqrt{\frac{n(n+1)}{2}}}{(3n+1)^2} \right) a^2 = bB. \quad (\text{A.19})$$

Where the second equality follows from simplifying terms. The last equality follows from recognizing the expression for the budget B for this instance, given equation (2.43). Namely, constraint (2.35) holds with equality for this instance.

Third, we check that constraint (2.36) is satisfied. Note, from equations (2.41) and (2.42), that for this instance, and for the case $l = 2$, constraint (2.36) evaluates to $bB \leq \frac{(n-1)c_n}{4(n+1)}(2a + (n-1)c_n) - (n-1)\frac{c_n^2}{4} = \frac{(n-1)}{2(n+1)}c_n(a - c_n)$, where the equality follows from simplifying terms. Plugging in the first expression for bB from equation (A.19), we get that for this instance, constraint (2.36) is equivalent to $c_n(2c_n - a) \leq \frac{(n-1)}{2(n+1)}c_n(a - c_n)$, which is equivalent to $c_n \leq \frac{(3n+1)}{(5n+3)}a$. However, from equation (2.42), it follows that

$$c_n = \left(\frac{n + \sqrt{\frac{n(n+1)}{2}}}{3n + 1} \right) a \leq \frac{(3n + 1)}{(5n + 3)} a.$$

Where the inequality is equivalent to $32n^4 + 43n^3 + 26n^2 + 9n + 2 \geq 0$, which holds for any $n \geq 1$.

Finally, we check that constraints (2.37) and (2.38) are satisfied. Note, from equation (2.42), that for this instance, and for the case $l = 2$, the right hand side of constraint (2.37) evaluates to 0, therefore constraints (2.37) and (2.38) are equivalent. Moreover, from equation (2.43) it follows that for this instance $B \geq 0$, for any $n \geq 2$.

Hence, the instance defined by equations (2.41), (2.42), and (2.43), is feasible for the problem (WCP_2) . ■

Lemma 16. *For any number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, the instance defined by equations (2.41)-(2.43) attains an objective value of $\frac{2 + \sqrt{2 + 2/n}}{4}$ for problem (WCP_2) .*

Proof. Plugging in equations (2.41) and (2.42) into equations (A.17) and (A.18), it follows that for this instance the objective function of problem (WCP_2) is equivalent

to

$$\frac{4(n+1)c_n}{(3n+1)a - 2(n-1)c_n + \sqrt{(n-1)^2(a-2c_n)^2 + 8(n-1)(n+1)bB}}. \quad (\text{A.20})$$

Now we will evaluate the square root term in equation (A.20). Plugging in equations (2.42) and (2.43) we get

$$\begin{aligned} & \sqrt{(n-1)^2 \left(a - 2 \left(\frac{n + \sqrt{\frac{n(n+1)}{2}}}{3n+1} \right) a \right)^2 + 8(n-1)(n+1) \left(\frac{(n-1)\sqrt{\frac{n(n+1)}{2}}}{(3n+1)^2} \right) a^2} \\ &= \sqrt{\frac{(n-1)^2 a^2}{(3n+1)^2} \left(n+1 + 2\sqrt{\frac{n(n+1)}{2}} \right)^2} \\ &= \frac{(n-1)a}{(3n+1)} \left(n+1 + 2\sqrt{\frac{n(n+1)}{2}} \right). \end{aligned}$$

Where the first equality follows from simplifying the expression, and the second equality follows from $n \geq 2$. Plugging in the latter expression for the square root term, together with equation (2.42), in equation (A.20), we get that for this instance the objective function of problem (WCP_2) is equivalent to

$$\begin{aligned} & \frac{4(n+1) \left(\frac{n + \sqrt{\frac{n(n+1)}{2}}}{3n+1} \right) a}{(3n+1)a - 2(n-1) \left(\frac{n + \sqrt{\frac{n(n+1)}{2}}}{3n+1} \right) a + \frac{(n-1)a}{(3n+1)} \left(n+1 + 2\sqrt{\frac{n(n+1)}{2}} \right)} \\ &= \frac{4(n+1) \left(n + \sqrt{\frac{n(n+1)}{2}} \right)}{(3n+1)^2 - (n-1)^2} \\ &= \frac{n + \sqrt{\frac{n(n+1)}{2}}}{2n} \\ &= \frac{2 + \sqrt{2 + 2/n}}{4}. \end{aligned}$$

Where the first equality follows from simplifying the expression, and the second equal-

ity follows from $(3n + 1)^2 - (n - 1)^2 = 8n(n + 1)$. ■

Proof of Proposition 11

Proof. We will show that for any $n \geq 3$, $a > 0$, $b > 0$, and for any index $k \in \{2, \dots, n - 1\}$, there is no feasible solution to problem $(RWCP_{2,k})$ that attains an objective value smaller than $\frac{2+\sqrt{2+2/n}}{4}$. The result then follows from the observation that this lower bound is attained by the candidate instance from equations (2.41)-(2.43), which is feasible for problem $(RWCP_{2,1})$

From Lemma 12 it follows that the objective function in problem $(RWCP_{2,k})$ is quasiconvex. Hence, its minimum must be attained either at one of the extremes of the feasible interval $c_n \in [\frac{a}{k+1}, a]$, or at an interior stationary point. We will analyze each one of these cases, and show that none of them attains an objective value smaller than $\frac{2+\sqrt{2+2/n}}{4}$.

(i) If $c_n = \frac{a}{k+1}$, then the objective function of problem $(RWCP_{2,k})$ evaluates to

$$\begin{aligned} \frac{Q_k^U(c_n)}{Q_{2,k}^*(c_n)} &= \frac{4(n+1)k}{(n+1)(3k+1) + \sqrt{(n-1)(n+1)(k-1)(k+1)}} \\ &\geq \frac{4k}{3k+1 + \sqrt{k^2-1}} \\ &\geq \frac{8}{7 + \sqrt{3}} \\ &\geq \frac{2 + \sqrt{2+2/n}}{4}. \end{aligned}$$

Where the first inequality follows from the left hand side being decreasing in n , and taking the limit as $n \rightarrow \infty$. The second inequality follows from the left hand side being increasing in k , and taking $k = 2$. The left hand side of the second inequality is increasing in k for any $k \geq \sqrt{2}$, thus for any index $k \in \{2, \dots, n - 1\}$, for any $n \geq 3$. Finally, the last inequality holds for any $n \geq 3$.

(ii) If $c_n = a$, then the objective function of problem $(RWCP_{2,k})$ evaluates to

$$\begin{aligned}
\frac{Q_k^U(c_n)}{Q_{2,k}^*(c_n)} &= \frac{4(n+1)k}{n+2k+1+\sqrt{(n-1)((8k^2+2k-1)n+6k^2+2k-1)}} \\
&\geq \frac{4k}{1+\sqrt{8k^2+2k-1}} \\
&\geq \frac{4k}{3k+2} \\
&\geq 1 \\
&\geq \frac{2+\sqrt{2+2/n}}{4}.
\end{aligned}$$

Where the first inequality follows from the left hand side being decreasing in n , and taking the limit as $n \rightarrow \infty$. The second inequality follows from $(3k+1)^2 \geq (8k^2+2k-1)$, for any index $k \in \{2, \dots, n-1\}$, for any $n \geq 3$. The third inequality follows from the left hand side being increasing in k , and taking $k = 2$. Finally, the last inequality holds for any $n \geq 1$.

(iii) Any interior stationary solution to problem $(RWCP_{2,k})$ must satisfy

$$\frac{d Q_k^U(c_n^*)/Q_{2,k}^*(c_n^*)}{dc_n} = 0.$$

After simplifying, this condition is equivalent to

$$\sqrt{*_{2,k}} = \frac{2(n-1)(2nk+n+k)c_n^* - (n-1)^2a}{3n+1}. \quad (\text{A.21})$$

Plugging in expression (A.21) in the objective function of problem $(RWCP_{2,k})$, it follows that any interior stationary solution must satisfy

$$\begin{aligned}
\frac{Q_k^U(c_n)}{Q_{2,k}^*(c_n)} &= \frac{(3n+1)kc_n^*}{2na+n(k-1)c_n^*} \\
&\geq \frac{3nk+k}{3nk+n} \\
&\geq \frac{6n+2}{7n} \\
&\geq \frac{2+\sqrt{2+2/n}}{4}.
\end{aligned}$$

Where the first inequality follows from the right hand side being increasing in c_n^* , and taking its lower bound $c_n^* = \frac{a}{k+1}$. The right hand side of the first equality is increasing in c_n^* if $2na > 0$, which holds for any $n \geq 1$. The second inequality follows from the left hand side being increasing in k , and taking $k = 2$. Finally, the last inequality holds for any $n \geq 1$. \blacksquare

Proof of Proposition 12

Proof. For any number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, the worst case performance of uniform co-payments for the case $l = 1$, $WC(1, n)$, can be computed as as the optimal objective value of the following problem

$$WC(1, n) \equiv \min_{B, \mathbf{c}} \frac{Q^U(\mathbf{c})}{Q_1^*(B, \mathbf{c})} = \frac{2(n+1)(nc_n - \sum_{i=1}^n c_i)}{na - \sum_{i=1}^n c_i + \sqrt{*}_1(B, \mathbf{c})}$$

s.t. $c_1 = 0$ (A.22)

$$c_i \leq c_{i+1}, \text{ for each } i \in \{1, \dots, n-1\} \quad (\text{A.23})$$

$$(WCP_1) \quad c_n \leq a \quad (\text{A.24})$$

$$bB = \left(nc_n - \sum_{i=1}^n c_i \right) \left((n+1)c_n - \sum_{i=1}^n c_i - a \right) \quad (\text{A.25})$$

$$bB \geq \frac{1}{4(n+1)} \left(\sum_{i=2}^n (c_i - c_1) \right) \left(2a + \sum_{i=2}^n c_i - (n+1)c_1 \right) - \sum_{i=2}^n \frac{(c_i - c_1)^2}{4}, \quad (\text{A.26})$$

where constraint (A.26) corresponds to constraint (2.37) in the generic problem (WCP_l) . From $l = 1$ it follows that we can drop constraint (2.36) from problem (WCP_l) , because there is no firm with index $(l-1)$. Moreover, without loss of generality we drop constraint (2.38) from problem (WCP_l) as well, because $B \geq 0$ is implied by constraints (A.25) and (A.26). Specifically, assume for a contradiction that $B < 0$, then the right hand side of constraint (A.25) must be negative as well. Using Lemma 9 to assume without loss of generality that $c_1 = 0$, this implies that

$a > (n+1)c_n - \sum_{i=2}^n c_i$. Therefore, from constraint (A.26) we conclude that

$$\begin{aligned} bB &\geq \frac{1}{4(n+1)} \left(\left(\sum_{i=2}^n c_i \right) \left(2a + \sum_{i=2}^n c_i \right) - (n+1) \sum_{i=2}^m c_i^2 \right) \\ &> \frac{1}{4(n+1)} \left((n+1)c_n \sum_{i=2}^n c_i - (n+1) \sum_{i=2}^m c_i^2 \right) \geq 0, \end{aligned}$$

a contradiction with $B < 0$. Hence, constraints (A.25) and (A.26) imply $B \geq 0$ in this case.

Let (B^*, \mathbf{c}^*) be an optimal solution to problem (WCP_1) . Lemma 17 below shows that if the k largest variables c_i^* are equal to c_n^* , with $k \in \{1, \dots, n-1\}$, then the objective function is strictly increasing in c_n . It follows that c_n^* must attain its lower bound, otherwise we could strictly improve the objective by decreasing it.

From Lemma 17 for $k = 1$ it follows that either constraint (A.26) is tight, or we must have $c_n^* = c_{n-1}^*$. If constraint (A.26) is tight, we are done. Therefore, assume that $c_n^* = c_{n-1}^*$. In fact, Lemma 17 allows us to iterate this argument for each $k \in \{2, \dots, n-2\}$, and conclude that either constraint (A.26) is tight, or we must have $c_n^* = c_i^*$ for each $i \in \{2, \dots, n\}$. Again, if constraint (A.26) is tight, we are done. Therefore, assume that $c_n^* = c_i^*$ for each $i \in \{2, \dots, n\}$. It follows that constraint (A.26) simplifies to

$$\begin{aligned} bB^* &= c_n^*(2c_n^* - a) \geq (n-1) \frac{c_n^*}{2} \left(a + (n-1) \frac{c_n^*}{2} \right) \frac{1}{n+1} - (n-1) \frac{(c_n^*)^2}{4} \\ \Leftrightarrow c_n^* &\geq \frac{(3n+1)}{(5n+3)} a > 0. \end{aligned}$$

Where the first equality follows from constraint (A.25), and $c_n^* = c_i^*$ for each $i \in \{2, \dots, n\}$. The equivalence follows from simplifying the expression.

Finally, from Lemma 17 for $k = n-1$ it follows that c_n^* must attain its lower bound, hence constraint (A.26) must be tight. This concludes the proof. \blacksquare

Lemma 17. *For any given number of firms in the market $n \geq 2$, and demand parameters $a > 0$, $b > 0$, if the k largest variables c_i are equal to c_n in problem (WCP_1) ,*

with $k \in \{1, \dots, n-1\}$, then

$$\frac{\partial Q^U(\mathbf{c})/Q_1^*(B, \mathbf{c})}{\partial c_n} > 0. \quad (\text{A.27})$$

Proof. Note that, for the case $l = 1$, the function $\sqrt{*}_l(B, \mathbf{c})$ in equation (2.31) simplifies to

$$\begin{aligned} \sqrt{*}_1(B, \mathbf{c}) = & \left(\left(na - \sum_{i=2}^n c_i - c_1 \right)^2 + n(n+1) \sum_{i=2}^n (c_i - c_1)^2 \right. \\ & \left. - (n+1) \left(\sum_{i=2}^n c_i - (n-1)c_1 \right)^2 + 4n(n+1)bB \right)^{1/2}. \end{aligned}$$

Then, from the assumption that the k largest variables c_i are equal to c_n , and taking $c_1 = 0$ without loss of generality from Lemma 9, it follows that

$$\begin{aligned} \frac{\partial Q^U(\mathbf{c})/Q_1^*(B, \mathbf{c})}{\partial c_n} = & \frac{2(n+1)(n-k)}{na - kc_n - \sum_{i=1}^{n-k} c_i + \sqrt{*}_1(B, \mathbf{c})} \\ & - \frac{2(n+1) \left((n-k)c_n - \sum_{i=1}^{n-k} c_i \right)}{\left(na - kc_n - \sum_{i=1}^{n-k} c_i + \sqrt{*}_1(B, \mathbf{c}) \right)^2} \left(-k + \frac{1}{2\sqrt{*}_1(B, \mathbf{c})} \frac{\partial \sqrt{*}_1}{\partial c_n} \right). \end{aligned}$$

From algebraic manipulations, and recognizing terms, it follows that

$$\begin{aligned} & \frac{\partial Q^U(\mathbf{c})/Q_1^*(B, \mathbf{c})}{\partial c_n} > 0 \\ \Leftrightarrow & \underbrace{2nb \left(\sum_{i=1}^{n-k} (a - c_i) \right)}_{>0} \underbrace{\left(Q_1^* - Q^U \right)}_{\geq 0} + \underbrace{\left(n \sum_{i=2}^n c_i^2 - \left(\sum_{i=2}^n c_i \right)^2 \right)}_{>0} + k \underbrace{\left(\sum_{i=1}^{n-k} (c_n - c_i) \right)}_{>0} > 0. \end{aligned}$$

This concludes the proof. ■

Lemma 18. *The solution given in equations (2.67)-(2.69) is feasible for problem (DLP_l) and attains an objective value of $\frac{(l-1)}{2nl-n+l-1}$.*

Proof. The objective value of problem (DLP_l) is λ , therefore the checking the objective value is direct and we focus on checking feasibility.

For constraint (2.61) we get $-au_n + a\gamma + 2na\lambda = 0 - 2na\lambda + 2na\lambda = 0$, thus constraint (2.61) is binding. For constraint (2.60) we get $-u_{n-1} + u_n - n\gamma - \lambda = n\gamma + \lambda + n - 1 + 0 - n\gamma - \lambda = n - 1$, thus constraint (2.61) holds.

For constraints (2.59), for $i \in \{l + 1, \dots, n - 1\}$, we get $-u_{i-1} + u_i + \gamma - \lambda = i\gamma + (n - i + 1)\lambda + i - 1 - (i + 1)\gamma - (n - i)\lambda - i + \gamma - \lambda = -1$, thus constraints (2.59) hold for $i \in \{l + 1, \dots, n - 1\}$. For constraint (2.59), for $i = l$, we get $-u_{l-1} + u_l + \gamma - \lambda = 0 - (l + 1)\gamma - (n - l)\lambda - l + \gamma - \lambda = -l2n\lambda - (n - l + 1)\lambda - l = l - 1 - l = -1$, thus constraint (2.59) holds for $i = l$.

For constraint (2.58), we get $-u_{l-2} + u_{l-1} + \gamma - (n - l + 3)\lambda = -\gamma + (n - l + 3)\lambda - 1 + 0 + \gamma - (n - l + 3)\lambda = -1$, thus constraint (2.58) holds. For constraints (2.57), for $i \in \{3, \dots, l - 2\}$, we get $-u_{i-1} + u_i + \gamma - 2\lambda = -(l - i)\gamma + (n + l - 2i + 1)\lambda - l + i + (l - i - 1)\gamma - (n + l - 2i - 1)\lambda + l - i - 1 + \gamma - 2\lambda = -1$, thus constraints (2.57) hold for $i \in \{3, \dots, l - 2\}$.

For constraint (2.56), we get $u_2 + \gamma - 2\lambda = (l - 3)\gamma - (n + l - 5)\lambda + l - 3 + \gamma - 2\lambda = (l - 2)\gamma - (n + l - 3)\lambda + l - 3 = -(l - 2)2n\lambda - (n + l - 3)\lambda + l - 3 = -\frac{(2nl - 3n + l - 3)}{2nl - n + l - 1}(l - 1) + l - 3 \leq -1$, where the last inequality is equivalent to $n - l + 1 \geq 0$, which is true for any $2 \leq l \leq n$, thus constraint (2.56) holds. For constraint (2.62), we get $\gamma = -2na\lambda = -2na\frac{(l-1)}{2nl-n+l-1} \leq 0$ for any $2 \leq l \leq n$, thus constraint (2.62) holds.

For constraints (2.63), for $i \in \{2, \dots, l - 2\}$, we get $u_i = (l - i - 1)\gamma - (n + l - 2i - 1)\lambda + l - i - 1 \leq (l - 2)\gamma - (n + l - 3)\lambda + l - 2 \leq 0$, where the first inequality follows from the expression for u_i being decreasing in i , and taking $i = 1$. The second inequality is exactly what we already showed for constraint (2.56). The expression for u_i is decreasing in i if and only if $n - l + 1 \geq 0$, which is true for any $2 \leq l \leq n$, thus constraints (2.63) hold for $i \in \{2, \dots, l - 2\}$.

Finally, for constraints (2.63), for $i \in \{l, \dots, n - 1\}$, we get $u_i = -(i + 1)\gamma - (n - i)\lambda - i \leq -(l + 1)\gamma - (n - l)\lambda - l = (l + 1)2n\lambda - (n - l)\lambda - l = \frac{(2nl + n + l)}{2nl - n + l - 1}(l - 1) - l \leq 0$, where the first inequality follows from the expression for u_i being decreasing in i , and taking $i = l$. The inequality is equivalent to $n \geq 0$. The expression for u_i is decreasing in i if $n \geq 0$, thus constraints (2.63) hold, for $i \in \{l, \dots, n - 1\}$. This completes the proof. ■

Proof of Proposition 13

Proof. For any given $n \geq 3$, $a > 0$, $b > 0$, consider any optimal solution (B^*, \mathbf{c}^*) to problem (WCP_3) . Note that if (B^*, \mathbf{c}^*) is such that constraint (2.36) is tight, then it follows that the worst case instance for the case $l = 3$ lies in the boundary between the cases $l = 3$ and $l = 2$, or equivalently $WC(2, n) \leq WC(3, n)$ and we are done. Similarly, if (B^*, \mathbf{c}^*) is such that constraint (2.37) is tight, then it follows that the worst case instance for the case $l = 3$ lies in the boundary between the cases $l = 3$ and $l = 4$ (where if $n = 3$, then the case $l = 4$ is only defined in the boundary where a fictitious firm 4 is about to start producing), or equivalently $WC(2, n) \leq WC(4, n) \leq WC(3, n)$, where the first inequality follows from the case $l = 4$ in Theorem 6, and we are done in this case as well. Hence, without loss of generality we will assume that constraints (2.36) and (2.37) are *loose* for (B^*, \mathbf{c}^*) .

Lemma 19 below shows that then (B^*, \mathbf{c}^*) must be such that $c_2^* = c_1^*$, and $c_i^* = c_n^*$, for each $i \in \{3, \dots, n\}$. Therefore, without loss of generality we focus on solutions with this structure. Moreover, from Lemma 9 we will assume, without loss of generality, that $c_1^* = 0$. It follows that problem (WCP_3) simplifies to the following one variable optimization problem.

$$\min_{c_n} \frac{Q_3^U(c_n)}{Q_3^*(c_n)} = \frac{12(n+1)c_n}{(5n+2)a - 3(n-2)c_n + ((n-2)(3c_n - a)(9(3n+2)c_n - (n-2)a))^{1/2}}$$

$$\text{s.t. } \frac{a}{3} \leq c_n \tag{A.28}$$

$$c_n \leq \frac{2(5n+2)a}{9(3n+2)}, \tag{A.29}$$

where we have dropped the dependency on the budget B by directly replacing it with the expression from constraint (2.35). Moreover, constraint (A.28) is equivalent to constraint (2.36), and constraint (A.29) is equivalent to constraint (2.37).

Now we show that for any given $n \geq 3$, $a > 0$, any optimal solution c_n^* to problem $(SWCP_3)$ must have an objective value of at least $WC(2, n)$. Recall that if at any optimal solution to problem $(SWCP_3)$ constraints (A.28) or (A.29) are tight then we

are done. It follows that, without loss of generality, we can focus on stationary points in the interior of the feasible interval. Namely, we focus on values of c_n such that

$$\frac{d(Q_3^U(c_n^*)/Q_3^*(c_n^*))}{dc_n} = 0.$$

After simplifying, this condition is equivalent to

$$((n-2)(3c_n^* - a)(9(3n+2)c_n^* - (n-2)a))^{1/2} = 3(n-2)c_n^* - \frac{(n-2)^2}{5n+2}a. \quad (\text{A.30})$$

By plugging in expression (A.30) into the objective function, it follows that any interior stationary point c_n^* must be such that its objective value has the following simplified expression

$$\frac{Q_3^U(c_n^*)}{Q_3^*(c_n^*)} = \frac{(5n+2)c_n^*}{2na}. \quad (\text{A.31})$$

Furthermore, equation (A.30) is quadratic in c_n^* and its unique non-negative solution is $c_n^* = \left(\frac{3n+\sqrt{6n(n+1)}}{3(5n+2)}\right)a$. Hence, from the right hand side of equation (A.31) it follows that any interior stationary point c_n^* attains an objective value of $\frac{3+\sqrt{6+6/n}}{6} \geq \frac{2+\sqrt{2+2/n}}{4} \geq WC(2, n)$, where the first inequality holds for any $n \geq 1$, and the second inequality follows from the fact that the right hand side is attained by the candidate instance for the case $l = 2$ from equations (2.41)-(2.43). ■

Lemma 19. *For any given number of firms in the market $n \geq 3$, and demand parameters $a > 0$, $b > 0$, any optimal solution (B^*, \mathbf{c}^*) to problem (WCP_3) for which constraints (2.36) and (2.37) are loose must be such that $c_2^* = c_1^*$, and $c_i^* = c_n^*$, for each $i \in \{3, \dots, n\}$.*

Proof. For any given $n \geq 3$, $a > 0$, $b > 0$, consider any optimal solution (B^*, \mathbf{c}^*) to problem (WCP_3) such that constraints (2.36) and (2.37) are loose. First we show that, for each index $i \in \{2, \dots, n-1\}$, it must be the case that either $c_i^* = c_1^*$ or $c_i^* = c_n^*$. Assume for a contradiction that $i \in \{2, \dots, n-1\}$ is the largest index such that $c_1^* < c_i^* < c_n^*$. Then we will show that we can transfer an arbitrarily small $\epsilon > 0$ from c_1^* to c_i^* and strictly improve this solution, while maintaining feasibility

for problem (WCP_3) , a contradiction.

We first address the feasibility of the modified solution. Recall that from Lemma 9 it follows that we can assume without loss of generality that $c_1^* = \delta > 0$, allowing the latter transfer for an $0 < \epsilon \leq \delta$. From constraint (2.35) it follows that the budget B^* remains unchanged when transferring an arbitrarily small $\epsilon > 0$ from c_1^* to c_i^* . Therefore, from constraints (2.36) and (2.37) being loose, and $c_i^* < c_n^*$, we conclude that the modified solution is feasible for problem (WCP_3) for an $\epsilon > 0$ small enough.

We now address the change in the objective function. Recall from equation (A.17) that the function $Q^U(\mathbf{c})$ remains unchanged when transferring an arbitrarily small $\epsilon > 0$ from c_1^* to c_i^* , therefore we focus on the change in the function $Q_3^*(B, \mathbf{c})$, which is

$$\frac{\partial Q_3^*(B^*, \mathbf{c}^*)}{\partial c_i} - \frac{\partial Q_3^*(B^*, \mathbf{c}^*)}{\partial c_1} = \frac{(n-2)}{2b\sqrt{*}_3(B^*, \mathbf{c}^*)}(c_i^* - c_1^*) > 0,$$

where the inequality follows from the assumption that $c_1^* < c_i^*$. Namely, we have shown that there exists a feasible solution to problem (WCP_3) which attains a strictly better objective value than (B^*, \mathbf{c}^*) , a contradiction. Hence, we conclude that, for each index $i \in \{3, \dots, n-1\}$, we must have either $c_i^* = c_1^*$ or $c_i^* = c_n^*$.

To conclude, note that assuming $l = 3$ implies $c_2^* < c_3^*$, otherwise if $c_2^* = c_3^*$ then firm 2 would get a co-payment whenever firm 3 does, contradicting the definition of the index l in Proposition 5. It follows that (B^*, \mathbf{c}^*) must have the structure given in the statement of the proposition. ■

Proof of Lemma 12

Proof. For any given $n \geq 2$, $a > 0$, $b > 0$, and for any index $k \in \{1, \dots, n-1\}$, the function $\sqrt{*}_{2,k}(c_n)$ fits the setting in Lemma 21 below with $\alpha = 2(k+1)(4nk+n+3k)$, $\beta = -4(2nk+n+k)a$ and $\gamma = (n-1)a^2$. It follows that $\sqrt{*}_{2,k}(c_n)$ is concave in the set $c_n \in [\frac{a}{k+1}, a]$ if $4\alpha\gamma - \beta^2 \leq 0$, or equivalently if $(4k^2+3k+1)n^2 + (3k+1)^2n + 5k^2 + 3k \geq 0$, which holds true for any index $k \in \{1, \dots, n-1\}$, for any $n \geq 2$. From here, together with $(3n+1)a > 2(n-k)c_n$, for any $c_n \in [\frac{a}{k+1}, a]$, $k \in \{1, \dots, n-1\}$, and $n \geq 2$, it follows that condition 2 in Lemma 20 holds.

On the other hand, condition 1 in Lemma 20 holds because the numerator in the objective function of problem $(RWCP_{2,k})$ is linear, thus convex, and $c_n \geq 0$ for any $c_n \in [\frac{a}{k+1}, a]$, $k \in \{1, \dots, n-1\}$, and $n \geq 2$. ■

Lemma 20. *Let $g : S \rightarrow \mathbb{R}$ and $h : S \rightarrow \mathbb{R}$, where S is a nonempty convex set in \mathbb{R}^n . Consider the function $f : S \rightarrow \mathbb{R}$ defined by $f(x) = g(x)/h(x)$. Then, f is quasiconvex if the following two conditions hold true:*

1. *g is convex on S , and $g(x) \geq 0$ for each $x \in S$.*
2. *h is concave on S , and $h(x) > 0$ for each $x \in S$.*

Proof. See Bazaraa et al. (2006). ■

Lemma 21. *Let $f(x) = \alpha x^2 + \beta x + \gamma$, if $\beta^2 - 4\alpha\gamma \leq 0$ and $f(x) > 0$ for each $x \in S \subset \mathbb{R}$, then $\sqrt{f(x)}$ is convex on S .*

Proof. $\frac{d^2 \sqrt{f(x)}}{dx^2} = \frac{4\alpha\gamma - \beta^2}{2(f(x))^{3/2}} \geq 0 \Leftrightarrow 4\alpha\gamma - \beta^2 \geq 0$, where the equivalence follows from $f(x) > 0$ for each $x \in S \subset \mathbb{R}$. ■

Bibliography

- A2S2. Who/ema prequalified act manufacturers, 2014. URL <http://www.a2s2.org/market-data/map-of-stakeholders/#1>.
- D. Acemoglu and A. Ozdaglar. Competition and efficiency in congested markets. *Mathematics of Operations Research*, 32(1):1 – 31, 2007.
- E. Adida and V. DeMiguel. Supply chain competition with multiple manufacturers and retailers. *Operations Research*, 59(1):156–172, 2011.
- M. Aghassi, D. Bertsimas, and G. Perakis. Solving asymmetric variational inequalities via convex optimization. *Operations research letters*, 34(5):481–490, 2006.
- AMFm Independent Evaluation Team. Independent evaluation of phase 1 of the affordable medicines facility - malaria (amfm), multi-country independent evaluation report: Final report. Technical report, The Global Fund, 2012. URL www.theglobalfund.org/documents/amfm/AMFm_2012IEPhase1FullFinal_Report_en/.
- K. Arifoglu, S. Deo, and S. Irvani. Consumption externality and yield uncertainty in the influenza vaccine supply chain: Interventions in demand and supply sides. *Management Science*, 58(6):1072 – 1091, 2012.
- K. J. Arrow, C. B. Panosian, and H. Gelband. *Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance*. National Academies Press, Washington, DC, 2004.
- G. Aydin and E. L. Porteus. Manufacturer-to-retailer versus manufacturer-to-consumer rebates in a supply chain. In *Retail Supply Chain Management*, pages 237–270. Springer, 2009.
- M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*, chapter 3. Convex Functions and Generalizations, pages 97 – 162. Wiley, 2006.
- H. P. Benson. Concave minimization: Theory, applications and algorithms. In R. Horst and P. M. Pardalos, editors, *Handbook of Global Optimization*, pages 43–148. Kluwer, 1995.

- F. Bernstein and A. Federgruen. Decentralized supply chains with competing retailers under demand uncertainty. *Management Science*, 51(1):18–29, 2005.
- D. P. Bertsekas. *Nonlinear programming*. Athena Scientific, Belmont, MA, 1999.
- D. Bertsimas and J. N. Tsitsiklis. *Introduction to linear optimization*. Athena Scientific Belmont, MA, 1997.
- I. Boulton. Background paper: Forum for africa-based pharmaceutical manufacturers and potential investors. Technical report, The Global Fund, 2011. URL http://www.theglobalfund.org/documents/amfm/AMFm_ForumForAfricaBasedPharmaceuticalManufacturersAndPotentialInvestors_Background_Paper_en/.
- S. Boyd and L. Vandenberghe. *Convex optimization*, chapter Duality, pages 215 – 288. Cambridge, 2007.
- J. A. Brander. Chapter 27 strategic trade policy. In Gene M. Grossman and Kenneth Rogoff, editors, *Handbook of International Economics*, volume 3, pages 1395 – 1455. Elsevier, 1995. doi: 10.1016/S1573-4404(05)80007-3. URL <http://www.sciencedirect.com/science/article/pii/S1573440405800073>.
- K. M. Bretthauer and B. Shetty. The nonlinear knapsack problem – algorithms and applications. *European Journal of Operational Research*, 92:459–472, 2002.
- D. L. Brito, E. Sheshinski, and M. D. Intriligator. Externalities and compulsory vaccinations. *Journal of Public Economics*, 45(1):69 – 90, 1991.
- G. J. Burke, J. Geunes, H. E. Romeijn, and A. Vakharia. Allocating procurement to capacitated suppliers with concave quantity discounts. *Operations Research Letters*, 36(1):103 – 109, 2008.
- G. P. Cachon. Supply chain coordination with contracts. In A. G. de Kok and S. C. Graves, editors, *Handbooks in Operations Research and Management Science, Vol 11: Supply Chain Management*, pages 229 – 339. Elsevier, Amsterdam, North - Holland, 2003.
- G. P. Cachon and M. L. Fisher. Supply chain inventory management and the value of shared information. *Management Science*, 46(8):1032–1048, 2000.
- A. Charnes and W. W. Cooper. Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, 9(3-4):181–196, 1962.
- J. Chemama, M. Cohen, R. Lobel, and G. Perakis. Designing subsidies with industry response dynamics: Commitment vs. flexibility. *Working Paper*, 2013.
- S.E. Chick, H. Mamani, and D. Simchi Levi. Supply chain coordination and influenza vaccination. *Operations Research*, 56(6):1493 – 1506, 2008.

- S. C. Choi. Price competition in a channel structure with a common retailer. *Marketing Science*, 10(4):271–296, 1991.
- J. Cohen, P. Dupas, and S. G. Schaner. Price subsidies, diagnostic tests, and targeting of malaria treatment. Technical report, National Bureau of Economic Research, 2014.
- M. Cohen, R. Lobel, and G. Perakis. Consumer subsidies with industry response for green technology adoption. *Under review*, 2012.
- M. Cohen, G. Perakis, and C Thraves. The effects of competition in green technology adoption. *Working Paper*, 2013.
- D. Collie. Strategic trade policy under asymmetric oligopoly. *European Journal of Political Economy*, 9:275 – 280, 1993.
- C. J. Corbett and U. S. Karmarkar. Competition and structure in serial supply chains with deterministic demand. *Management Science*, 47(7):966–978, 2001.
- J. R. Correa and N. E. Stier-Moses. Wardrop equilibria. *Wiley Encyclopedia of Operations Research and Management Science*, 2011.
- J. R. Correa, N. Figueroa, R. Lederman, and N. E. Stier Moses. Pricing with markups in industries with increasing marginal costs. *Mathematical Programming*, page Forthcoming, 2014.
- A. Cournot. *Researches into the Mathematical Principles of the Theory of Wealth*. Macmillan, New York, 1897.
- V. DeMiguel and H. Xu. A stochastic multiple-leader stackelberg model: analysis, computation, and application. *Operations Research*, 57(5):1220–1235, 2009.
- S. Deo and C. J. Corbett. Cournot competition under yield uncertainty: The case of the u.s. influenza vaccine market. *Manufacturing and Service Operations Management*, 11(4):563 – 576, 2009.
- L. Dong, P. Kouvelis, and P. Su. Global facility network design in the presence of competition. *European Journal of Operational Research*, 228(2):437–446, 2013.
- K. H. Van Donselaar, V. Gaur, T. Van Woensel, R. A. C. M. Broekmeulen, and J. C. Fransoo. Ordering behavior in retail stores and implications for automated replenishment. *Management Science*, 56(5):766–784, 2010.
- P. Dupas. Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica*, 82(1):197–228, 2014.
- A. Farahat and G. Perakis. Technical note—a comparison of bertrand and cournot profits in oligopolies with differentiated products. *Operations research*, 59(2):507–513, 2011.

- D. Fullerton and G. E. Metcalf. Chapter 26 tax incidence. In Alan J. Auerbach and Martin Feldstein, editors, *Handbook of Public Economics*, volume 4, pages 1787 – 1872. Elsevier, 2002. doi: 10.1016/S1573-4420(02)80005-2. URL <http://www.sciencedirect.com/science/article/pii/S1573442002800052>.
- M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.0 beta. <http://cvxr.com/cvx>, September 2012.
- S. Gupta and R. Loulou. Process innovation, product differentiation, and channel structure: Strategic incentives in a duopoly. *Marketing Science*, 17(4):301–316, 1998.
- J. Häckner. A note on price and quantity competition in differentiated oligopolies. *Journal of Economic Theory*, 93(2):233–239, 2000.
- N. Halman, D. Klabjan, C. Li, J. Orlin, and D. Simchi-Levi. Fully polynomial time approximation schemes for stochastic dynamic programs. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 700–709. Society for Industrial and Applied Mathematics, 2008.
- J. E. Harrington. Experimentation and learning in a differentiated-products duopoly. *Journal of Economic Theory*, 66(1):275–288, 1995.
- R. Johari, G. Y. Weintraub, and B. Van Roy. Investment and market structure in industries with congestion. *Operations Research*, 58(5):1303 – 1317, 2010.
- R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, New York, 1972.
- S. Khuller, A. Moss, and J. Naor. The budgeted maximum coverage problem. *Information Processing Letters*, 70:39–45, 1999.
- P. D. Klemperer and M. Meyer. Supply function equilibria in oligopoly under uncertainty. *Econometrica*, 57(6):1243 – 1277, 1989.
- A. G. Kök and M. L. Fisher. Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research*, 55(6):1001–1021, 2007.
- R. De Koster, T. Le-Duc, and K. J. Roodbergen. Design and control of warehouse order picking: A literature review. *European Journal of Operational Research*, 182(2):481–501, 2007.
- D. Krass, T. Nedorezov, and A. Ovchinnikov. Environmental taxes and the choice of green technology. *Production and Operations Management*, 22(5):1035–1055, 2013.
- D. Leahy and C. Montagna. Strategic trade policy with heterogeneous costs. *Bulletin of Economic Research*, 53(3):177 – 182, 2001.

- G. Lin, W. Zhu, and M. M. Ali. An exact algorithm for the 0–1 linear knapsack problem with a single continuous variable. *Journal of Global Optimization*, 50(4): 657–673, 2011.
- R. Lobel and G. Perakis. Consumer choice model for forecasting demand and designing incentives for solar technology. *Under review*, 2012.
- N. Van Long and A. Soubeyran. Cost heterogeneity, industry concentration and strategic trade policy. *Journal of International Economics*, 43:207 – 220, 1997.
- Z. Q. Luo, J. S. Pang, and D. Ralph. *Mathematical programs with equilibrium constraints*. Cambridge University Press, Cambridge, UK, 1996.
- H. Mamani, E. Adida, and D. Dey. Vaccine market coordination using subsidy. *IIE Transactions on Healthcare Systems Engineering*, 2(1):78 – 96, 2012.
- H. Marchand and L. A. Wolsey. The 0-1 knapsack problem with a single continuous variable. *Mathematical Programming*, 85(1):15–33, 1999.
- P. Marcotte and M. Patriksson. Traffic equilibrium. In C. Barnhart and G. Laporte, editors, *Handbooks in Operations Research and Management Science, Vol. 14: Transportation*, pages 623 – 713. Elsevier, Amsterdam, North - Holland, 2006.
- J. A. Van Mieghem. *Operations Strategy: Principles and Practice*. Dynamic Ideas, 2008.
- J. J. More and S. A. Vavasis. On the solution of concave knapsack problems. *Mathematical Programming*, 49(3):397–411, 1991.
- A. Nagurney. *Network Economics: A Variational Inequality Approach*. Kluwer Academic Publishers, Cambridge, MA, 1999.
- M. Patriksson and L. Wynter. Stochastic mathematical programs with equilibrium constraints. *Operations Research Letters*, 25:159 – 167, 1999.
- D. W. Pentico. The assortment problem with probabilistic demands. *Management Science*, 21(3):286–290, 1974.
- D. W. Pentico. The assortment problem: A survey. *European Journal of Operational Research*, 190(2):295–309, 2008.
- G. Raz and A. Ovchinnikov. Coordinating pricing and supply of public interest goods using rebates and subsidies. *Under review*, 2013.
- S. Sahni. Computationally related problems. *SIAM Journal on Computing*, 3(4): 262–279, 1974.
- N. Singh and X. Vives. Price and quantity competition in a differentiated duopoly. *The RAND Journal of Economics*, pages 546–554, 1984.

- C. M. Snyder, W. Begor, and E. R. Berndt. Economic perspectives on the advance market commitment for pneumococcal vaccines. *Health Affairs*, 30(8):1508 – 1517, 2011.
- H. V. Stackelberg. *The Theory of Market Economy*. Oxford University Press, London, 1952.
- M. Sviridenko. A note on maximizing a submodular set function subject to a knapsack constraint. *Operations Research Letters*, 32:41 – 43, 2004.
- T. A. Taylor and W. Xiao. Subsidizing the distribution channel: Donor funding to improve the availability of malaria drugs. *Management Science*, page Forthcoming, 2014.
- Jean Tirole. *The theory of industrial organization*. MIT press, 1988.
- Robert J Vanderbei. Loqo user’s manual–version 4.05. *Princeton University*, 2006.
- X. Vives. *Oligopoly Pricing: Old Ideas and New Tools*. The MIT press, Cambridge, MA, 2001.
- N. Wen, S. C. Graves, and Z. J. Ren. Ship-pack optimization in a two-echelon distribution system. *European Journal of Operational Research*, 220(3):777–785, 2012.
- N. J. White. Qinghaosu (artemisinin): the price of success. *Science*, 320:330 – 334, 2008.
- D. P. Williamson and D. B. Shmoys. *The design of approximation algorithms*. Cambridge University Press, New York, NY, 2011.
- L. A. Wolsey. Maximising real-valued submodular functions: primal and dual heuristics for location problems. *Mathematics of Operations Research*, 7:410–425, 1982.
- World Health Organization. World malaria report: 2013. Technical report, WHO, 2013. URL http://www.who.int/malaria/publications/world_malaria_report_2013/report/en/index.html.
- C. Zhao and X. Li. Approximation algorithms on 0–1 linear knapsack problem with a single continuous variable. *Journal of Combinatorial Optimization*, pages 1–7, 2013.
- P. H. Zipkin. *Foundations of inventory management*. McGraw-Hill New York, 2000.