Essays in Industrial Organization and Consumer Policy

Inauguraldissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

Christian Felix Michel

vorgelegt im Sommersemester 2013

Abteilungssprecher:	Prof. Dr. Martin Peitz
Referent:	Prof. Volker Nocke, PhD
Korreferent:	Prof. Philipp Schmidt-Dengler, PhD

Verteidigung: 04.06.2013

Acknowledgments

There are a lot of people from whom I benefited during the writing of this thesis. After almost four years of life and research in Mannheim, it is a good time to look back and thank those who helped me along the way.

First I would like to thank my advisors Volker Nocke and Philipp Schmidt-Dengler for their support throughout the writing of this thesis. Their curiosity, constructive criticism, and their ability to both spot important details and to the see the big picture enabled me to see many new dimensions in economic research.

I am also very thankful for the numerous conversations I had with Yuya Takahashi and Georg Duernecker. Their efforts crucially helped me in improving my empirical skills. To me they are role models for junior faculty members.

My papers and presentations also benefited from constant criticism and suggestions from André Stenzel and Philipp Wilking, who were never tired of having “another look” over a new draft.

I am very fortunate to have had the constant opportunity to discuss and interact with great colleagues and other researchers. In this case I especially want to thank Jana Friedrichsen, Johannes Koehnen, Tim Lee, Matthias Mand, Andras Niedermayer, Isabel Ruhmer, Nicholas Schuetz, Alex Shcherbakov, Naoki Wakamori, Stefan Weiergraeber, and Philipp Zahn. I am also grateful for many comments I received during my stays at Oxford and Bonn, where the work on chapters 3 and 4 started.[1]

In the end, and most importantly, I would like to thank my parents for their constant support and encouragement in pursuing my dreams.

[1]There are several people that might not have contributed research-wise to this thesis, but had a significant impact off-research. In this case I would like to thank Johannes Schoch, Florian Sarnetzki, Marcus Frömberg, Claus-Leo Stynen, Alexander Donges, Peter King, Phil Brooks, Larry Max, Asleif Phileasson and team, members of the 9th Bronnbacher year, Alexander v. Güstrow, Mike Costa, Antonio Fuso, as well as employees of City Doener and Ristaurante Da Clarissa for their support.

Contents

List of Figures

List of Tables

To Stefanie and Hubert Michel

Chapter 1

General Introduction

This thesis contains three distinct, single-authored papers that are all related to the field of industrial organization. Chapter 2 proposes an empirical framework that allows to estimate the degree of industry competition and the degree of horizontal integration of merging firms. Thus, this chapter is also linked to the field of organizational economics. Chapters 3 and 4 analyze the effects of consumer biases on market outcomes and assess the effectiveness of specific consumer protection policies. These chapters also relate to the field of behavioral economics. In chapter 5, I provide a general discussion and give an outlook on open problems in the literature. In the following, I outline the different chapters in more detail.

Identification and Estimation of supply-side behavior Modern empirical research in industrial organization very often makes use of structural econometric models. These models impose significant structure on firm and consumer behavior to enable the estimation of the underlying parameters in a single industry. This is important, because it avoids having to make inference by comparing industries with systematically different characteristics.

Empirical organizational economics on the other hand often aims to find the determinants of firm performance by using differences in characteristics across firms and industries, see for example Bloom et al. (2012). In chapter 2 of this thesis, I propose a framework that can be applied to questions both from industrial organization and organizational economics. First, I ask how fast and how strongly merging firms jointly maximize profits after a horizontal merger. This relates to the question of whether merging firms finally fully integrate and form a single unit or whether they remain operationally separate after a merger. Second, I ask what is the form of competition between firms in an industry. This is one of the core questions in industrial organization. Conceptually, I approach both questions using a structural model that makes use of both pre- and post-merger industry data.

To my knowledge, this is the first framework that makes use of such a structural model also in an organizational context. Adopting structural models in this case is complementary to the existing approaches. Unlike existing approaches, this enables to make inference on within-firm behavior by using aggregate data in a single industry. I especially focus on the cooperation between merging divisions. Variation in retail prices, quantities sold and input prices before and after the merger across all products in the industry allow me to estimate the form of divisional cooperation in a horizontal setting. I show that the difference between observed and predicted post-merger prices results in a structural cost error. I combine this error with appropriate instruments to identify the underlying supply-side parameters.

From an industry viewpoint, this research relates to a relatively old literature of estimating industry conduct. The main difficulty lies in jointly identifying marginal costs and the form of industry conduct. This is important because marginal costs are usually assumed to be unobservable to the researcher.
The literature in this context started with estimating conduct in homogeneous product cases, see for example Bresnahan (1989) for an overview. Even though some of these approaches were applied to several industries, econometric research later showed that these approaches do not identify the parameters of interest (Corts 1999). Identifying conduct parameters in differentiated product cases turns out to be even more problematic because one has to find sufficient variation to be able to identify the different conduct parameters of interest. Because of difficulties in finding such appropriate variation, the literature has largely abstained from directly estimating industry conduct in the differentiated product case.
With my framework, I try to bridge the gap between the old conduct estimation literature and the modern differentiated product demand models. Using the merger itself as variation on the supply side, I estimate the underlying industry conduct between firms. As in the organizational case, I make use of a structural cost error

to identify the supply side parameters. When using data from the ready-to-eat cereal industry, I find that between 14.3 percent and 25.6 percent of the markups are due to cooperative behavior between firms.

Exploration of consequences of consumer biases and assessment of effectiveness of consumer protection policies Besides focusing on the identification and estimation of behavior both within and between firms, this thesis also aims to contribute to the understanding of interactions between firms and consumers. The interactions between fully rational consumers and firms have already been studied for a long time. In recent years, more attention has also been given to deviations from full consumer rationality, see for example Ellison (2006) and Vickers (2004). Potential consumer biases are in some cases used by policy makers as a justification for further regulatory intervention. However, such policy interventions often have several drawbacks and have to be analyzed carefully. First, they can restrict the choice set of potentially sophisticated consumers, thereby suppressing potentially optimal choices. Second, such intervention may cause a best response from firms with respect to prices or product quality that is harmful for consumer welfare. In chapters 3 and 4 of this thesis, I study the effects of two specific consumer biases on market outcomes and the effectiveness of associated consumer protection policies.

In chapter 3, I focus on the prevalence and high profitability of extended warranties in many retail industries, especially the consumer electronics industry. In this industry, extended warranties have become a prominent tool by retailers to increase profits. This is due to the fact that many consumers buy extended warranties in combination with a new product, but do not claim it even if the product breaks down and they are still eligible for a refund or repair. Furthermore, in other instances, some consumers make a false inference about the quality of a product because of the existence of a warranty. I provide a model that can explain both effects using a single consumer bias. In the model, some "naive" consumers underestimate the costs they incur when claiming a warranty payment after a product breakdown. Producers anticipate this behavior and respond by optimally setting their contracts. In monopoly, this can lead to two non-standard outcomes. First, since naive consumers overpredict the value of an extended warranty, a monopolist can profitably offer extended warranties at a price that is higher than the average willingness to pay of a fully sophisticated consumer. Second, he can offer a low-quality product at medium prices. This is because product quality is not observable prior to the purchase of the good. Therefore, the warranty also has a signaling role. At some medium warranties, a naive consumer believes that the product is of high quality, while it is still optimal for the producer to offer a low-quality product. To my knowledge, this is the first model that formalizes exploitative contracting via misperceptions of product qualities.

Competition between firms decreases prices, but naive consumers still make mistakes in that they do not choose their utility-maximizing product. I study the consequences of a consumer protection policy in terms of a minimum warranty standard. Such a policy is able to avoid exploitation via offering low quality products at high prices. However, it is not able to prevent firms from selling exploitative extended warranty contracts.

In chapter 4, I focus on a prediction error attributable to state-dependent utility and advertising and assess the effects of mandatory cooling-off periods on welfare. The model gives a psychological intuition for the persuasive effects of advertising. Advertising is able to influence a state-dependent “projection bias” people face when having to predict their future consumption utilities. Models of persuasive advertising usually face the problem of an ambiguous welfare evaluation: such advertising usually changes the underlying utility of a consumer. An assessment of the effects of advertising might then differ using pre- and post-merger utility. Since advertising only influences the bias and not the underlying preferences, this model allows for an unambiguous welfare evaluation.

Both researchers and policy makers previously advised to implement mandatory cooling-off periods to prevent consumers from making mistakes due to “impulse buying”. I analyze such a policy in my framework. I find that its welfare effect ambiguous. While some consumers might be saved from impulse buying, others who also value a product in a high way ex-post will have a loss due to the foregone consumption utility in the first period. An assessment of this policy crucially depends on the fraction of consumers that are initially prone to impulse buying. Furthermore, in some cases an overall welfare standard and a consumer policy standard can give diverging normative predictions of such a policy.

Relationship to previous own work This thesis consists of five chapters, and includes three distinct, single-authored research papers. Chapters 1, 2, and 5 were exclusively created during my time in Mannheim. Chapter 3 extends my own work of Michel (2009). Several sections, especially the main monopoly section, simultaneous oligopoly section, and the introduction largely include similar or identical parts compared to my previous work. There are, however, several key contributions that are new to this thesis. First, the previous work neither focused on nor explicitly mentioned extended warranties, which are a key policy component of the chapter. Second, the sequential oligopoly section covering Propositions 3.3 and 3.4 is not included either, as have all graphs in this chapter. Third, the analysis of consumer protection policies and all other extensions in section 3.5 are new. Chapter 4 extends my own work of Michel (2009) and Michel (2007). Several sections of this chapter, especially the introduction, the main model, the monopoly section, the duopoly section, and the welfare section largely include similar or identical parts compared to my previous work. However, there are also crucial differences compared to my previous work. First, previous work does not focus on one of the two key questions in this chapter, namely the consequences of mandatory cooling-off policies on consumer and overall welfare. Second, section 4.6, including Propositions 4.5 - 4.7 are not included either. Furthermore, the chapter contains several new graphs, tables, and additional motivation.

Chapter 2

Identification and Estimation of Intra-Firm and Industry Competition via Ownership Change

This chapter proposes and empirically implements a framework for analyzing industry competition and the degree of joint profit maximization of merging firms in differentiated product industries. Using pre- and post-merger industry data, I am able to separate merging firms' intra-organizational pricing considerations from industry pricing considerations. The insights of the chapter shed light on a long-standing debate in the theoretical literature about the consequences of organizational integration. Moreover, I propose a novel approach to directly estimate industry conduct that relies on ownership changes and input price variation. I apply my framework using data from the ready-to-eat cereal industry, covering the 1993 Post-Nabisco merger. My results show an increasing degree of joint profit maximization of the merged entities over the first two and a half years after the merger. I find that between 14.3 and 25.6 percent of industry markups can be attributed to cooperative industry behavior, while the remaining markup is due to product differentiation of multi-product firms.[1]

[1] I would like to thank my advisors Volker Nocke, Philipp Schmidt-Dengler, and Yuya Takahashi for their guidance and support. I also benefited from conversations with Steve Berry, Pierre Dubois, Georg Duernecker, David Genesove, Alex Shcherbakov, Andre Stenzel, Andrew Sweeting, Naoki Wakamori, and Stefan Weiergraeber, and received helpful comments from seminar participants at Mannheim, Toulouse, the 2012 CEPR Applied IO Summer School, and the EARIE 2012 in Rome.

2.1 Introduction

In this chapter, I propose a framework to address two open questions in industrial organization and organizational economics, and apply it to a merger from the ready-to-eat cereal industry. First, I examine to what degree merging firms jointly maximize their profits after a horizontal merger. This bridges a gap between empirical industrial organizational models and organizational economics models. Second, I provide a way to directly identify and estimate industry conduct in differentiated product models, which has for a long time been a problem in the industrial organization literature.

Existing empirical merger models make several simplifying assumptions with respect to both industry and within firm behavior. On an organizational level, these models assume that merging firms fully internalize their profits immediately after a merger. On an industry level, the form of supply side competition is either assumed to be known, or is chosen from a discrete set of non-nested forms of competition using some selection criterion. These assumptions are usually used together with pre-merger data to predict post-merger industry prices. My framework makes use of both observable pre- and post-merger data. This allows me to relax either the assumption of full profit internalization or the assumption of known industry conduct while keeping the other. I recover pre-merger marginal costs using first-order conditions between merging firms and predict post-merger marginal costs using a cost function estimation. Different forms of industry conduct or post-merger profit internalization of merging firms, respectively, will lead to different markups charged by firms. My identification strategy searches for the form of supply side competition that best predicts post-merger industry prices. I can show that the difference between predicted and observed post-merger prices amounts to a structural cost function error term, which helps to identify the supply-side parameters of interest given proper instruments. Thus, my approach can be seen as a supply side analogue to common demand side models that use structural error terms to identify demand side and cost function parameters, see for example Berry et al. (1995).

I apply the developed techniques to data from the ready-to-eat cereal industry covering the 1993 Post-Nabisco merger. In January 1993, Philipp Morris corporation's owned Kraft foods with its Post cereal line purchased the Nabisco ready-to-eat cereal branch from RJR Nabisco. The results indicate an increasing degree of joint profit maximization among merging firms after the merger. This suggest the existence of informational or contractual frictions among merging firms shortly after the merger. With respect to industry competitiveness, I find that between 14.3 and 25.6 percent of industry markups can be attributed to cooperative industry behavior, while the remaining markup is due to product differentiation of multi-product firms.

The paper extends the existing literature along several dimensions. To my knowledge, this is the first paper to focus on estimating the degree of joint profit maximization of a merged entity. This links the empirical industrial organization literature to the theoretical organizational economics literature on intra-firm coordination and horizontal integration by allowing for frictions between different divisions of a firm. Conceptually, the approach also differs from existing empirical organizational economics models in that its focus is on behavior within a single (post-merger) organization rather than on correlations across firms and industries. Moreover, I show that using proper supply side variation, it is indeed possible to estimate industry conduct directly in differentiated product industries. Using demand side variation, this can typically not be achieved due to a lack of sufficiently many demand rotators, see for example Nevo (1998).
Following the seminal paper by Berry et al. (1995), identification of demand and cost parameters relies on orthogonality conditions between structural error terms and appropriate instruments. Unlike the existing literature, I also identify the model's underlying supply side parameters, i.e. the degree of profit internalization and industry conduct, respectively. I show that the difference between observed and predicted post-merger prices represents a structural cost error term. I set up moment conditions that rely on orthogonality conditions between this error term and cost-side instruments to identify the supply side parameters.

Modern empirical industrial organization models assume that a merged entity maximizes the joint profits of all its products, thus abstracting from agency problems within the firm. From an organizational viewpoint, several theories predict that full internalization of joint profits cannot be achieved after a merger. Incentive structures that give managers bonuses based on the performance of their own division rather

than the performance of the firm as a whole can cause different horizontal divisions to compete with each other. Fauli-Oller and Giralt (1995) analyze a headquarter's choice of the optimal incentive scheme for division managers. Whenever products of different divisions are substitutes to each other, managers bonuses will be partly based on their own division's performance. There is also a growing literature in organizational economics that focuses on the trade-offs between coordination of decision-making through a headquarter and strategic communication of division managers.[2] Other reasons for no full maximization of joint profits immediately after a merger are delays in post-merger harmonization of firm strategies due to old contractual agreements, or a lack of information concerning revenue potential right after a merger.
I focus on a single merger to assess its consequences on joint maximization of profits. This differs from conventional empirical organizational economics frameworks that focus on correlations between observable firm characteristics across different firms and industries.[3]

When estimating industry conduct, I maintain the assumption that merging firms internalize the profits after the merger. Given marginal cost estimates and price-elasticities obtained from demand side-estimation, I can predict the effects of an ownership change on prices ex-post. By varying the form of supply side competition (i.e. industry conduct), and accounting for input price changes on the cost side, I look for the form of competition that most accurately predicts the effects of the merger-induced ownership change on prices. I estimate the predicted post-merger prices and compare them with the observed post-merger prices. The differences of observed prices and predicted prices is used to form moments in order to obtain the model's underlying conduct parameters using a Generalized Methods of Moments estimator.
Previous attempts to estimate both marginal costs and industry conduct have

[2]Alonso et al. (2008) study the optimal degree of centralization when managers communicate strategically. They show that while centralization can improve horizontal outcomes, it will lead to adverse vertical effects. Dessein et al. (2011) show the existence of endogenous incentive conflicts between headquarter managers and division managers within multi-divisional firms.

[3]There is a large empirical literature in organizational economics focusing on the determinants for specific organizational structures across firms and industries, see for example Lafontaine and Slade (2012), and Bloom et al. (2012) for an overview over the literature. McElheran (2010) finds a positive correlation between delegation of IT system adoption in multi-divisional firms and local information advantages, but a negative correlation between delegation of system adoption and firm size. Thomas (2011) argues that a reduction in the brand portfolios of firms in the laundry detergent industry across different countries would lead to an increase in their profits.

mostly been made using demand side variation. Bresnahan (1982) and Lau (1982) provide identification results for estimating conduct in the homogeneous good case. In differentiated product industries, these approaches usually face two kinds of problems. The first problem is the difficulty to find a sufficient number of demand rotators. Without such rotators, these approaches are not able to identify industry conduct.[4] The second problem relates to the estimation techniques, which only estimate the economic parameters of interest accurately in special cases. Corts (1999) critically discusses the identification of conduct parameters. He argues that the estimated parameters usually differ from the "as-if conduct parameters" and therefore do not reflect the economic parameters of interest. The static conduct estimation models are furthermore not able to detect all dynamic forms of collusion. While I cannot account for the latter point due to the static character of my framework, my estimation technique can overcome the former.

There is a small literature related to the estimation of industry conduct using supply side variation. Ciliberto and Williams (2010) develop an approach that relies on multi-market contact for estimating conduct in the airline industry. Their model includes conduct parameters that can have three different values, accounting for different degrees of cooperation among profit-maximizing firms. Oliveira (2011) uses marginal profit ratios in a dynamic model to distinguish between market competition and efficient "stick-and-carrot" collusion in the airline market. Brito et al. (2012) explore the effects of three Portuguese insurance mergers on coordinated effects and efficiency. They find no indication for an increase in coordinated effects after the mergers.

The remainder of the chapter is organized as follows. Section 2.2 gives an overview over the industry and the merger. Section 2.3 introduces the baseline model and discusses the conduct estimation strategy in detail. Section 2.4 provides identification results for both estimating the degree of joint profit maximization of merging firms and estimating industry conduct, respectively. Section 2.5 presents estimation results for both techniques. Section 2.6 introduces several extensions to the baseline model as outlined above. Section 2.7 concludes with a discussion of the results.

[4] Nevo (1998) discusses advantages and disadvantages of a direct conduct estimation compared to a non-nested menu approach. He argues that in practice estimating conduct directly will be impossible due to a lack of sufficiently many distinct demand shifters.

2.2 Industry Overview and the 1993 Post-Nabisco Merger

2.2.1 The ready-to-eat cereal industry

There are several factors that make the ready-to-eat (RTE) cereal industry an excellent starting point for oligopoly analysis.[5] Economies of scale in packaging different cereals, as well as in the distribution of products, cause barriers to entry for new firms. There is a frequent introduction of new products by existing firms, which goes in line with large advertising campaigns in the beginning of a product's lifecycle.[6] The cereals differ with respect to their product characteristics, such as sugar content or package design, and target different consumer types. At the start of the period I analyze, the industry consists of 6 main nationwide manufacturers: Kellogg's, General Mills, Post, Nabisco, Quaker Oats, and Ralston Purina. Figure 2.1 shows the market shares of the different products. It is common to classify the cereals into different groups, such as adult, family, and kids cereals, see also Nevo (2001). Table 2.1 shows the classification of the different cereal brands in my dataset into different segments. Kellogg's as the firm with the biggest market share has a strong presence in all segments. General Mills is mainly present in the family and kids segments, whereas Post and Nabisco have their main strengths in the adult segments.

On a retail level, cereals are primarily distributed via supermarkets. Supermarket promotions via price reductions or quantity discounts are a further tool used to increase quantities sold for a period of time. Many retailers also own private labels that compete in their stores with the nationwide manufacturers. I use scanner data from the first quarter of 1991 until the fourth quarter of 1995 from the Dominick's Finer Food database. My main dataset for the conduct estimation includes 28 brands from the 6 different nationwide firms. The scanner data involves 35 stores from the Chicago Metropolitan area, see Figure 2.2 for a geographic map of the stores. In particular, the dataset includes data on product prices, quantities sold, data on promotions, as well as 1990 census data yielding demographic variables for the different store locations. I use additional input price data from the Thomson Reuters Datastream database. Even

[5] This industry has already been studied extensively, see for example Schmalensee (1978), Nevo (2000), and Corts (1995). Although Corts presents a detailed industry description, to my knowledge the dynamic aspects on the supply side have not been investigated in detail.

[6] See for example Hitsch (2006) for a study of the determinants of successful brand introductions.

though I also observe data on Dominick's private label cereal, I do not include it in my conduct estimation. There are two reasons for this. First, I want to focus on the degree of competition between firms that are operating nationwide. Because a private label is only present for one retailer, and in my case a locally operating retailer, it will have different underlying objectives than the nationwide operating manufacturers. Second, a private label firm belongs to its retailer, thus leading to a joint maximization of profits upstream and downstream. This would require additional assumptions to be compatible with estimating "upstream" industry conduct.

In my data, I also observe the retailer's average acquisition costs for each product at a given time. This variable reflects the inventory-weighted average of the fraction of the retail price that was paid to the producer. From this variable I can compute a proxy for the average retail margin for a given period.[7] Figure 2.3 shows the development of the retail margin proxy over time for the different firms in the dataset. There are several interesting features. The retail margin varies significantly across the different firms. On average retail margins are highest for Ralston, the firm with the smallest market share, followed by Kellogg's, the firm with the highest market share. Thus, there is no clear relationship between retail margin and firm size, suggesting that there is no higher bargaining power for Kellogg's.[8] Another interesting fact is that the retail margin drops significantly around the time of the merger, from over 15% to single digit figures for several firms, including the merging firms. It is not clear whether this drop is due to the merger, which would imply some form of renegotiation between manufacturers and retailer in the period, or whether it is instead a pure coincidence.

2.2.2 The 1993 Post-Nabisco merger

Between 1990 and 1992, prices steadily increased in the industry, see Figure 2.4 for the price development per firm. On November 12, 1992, Kraft Foods made an offer to purchase RJR Nabisco's ready-to-eat cereal line. The acquisition was cleared by the FTC on January 4, 1993. On February 10th, 1993, the New York State attorney however sued for a divestiture of the Nabisco assets, which was finally turned down

[7] Dominick's uses the following formula for the average acquisition costs (AAC): AAC(t+1) = (Inventory bought in t) Price paid(t) + (Inventory, end of t-l-sales(t)) AAC(t).

[8] Another potential source for bargaining power not modeled here is bargaining power in form of leading to more premium shelf spaces.

3 weeks later.[9] Figure 2.5 shows the merging firms' price development over all stores. Table 2.2 shows the price development per product for the first four quarters after the merger. Average prices for the merging firms increase over time, which is in line with a unilateral effects merger model. The price development of other firms is more heterogeneous. Kellogg's decreases part of its brands prices while increasing some of its other prices. Prices for Ralston also go up. Prices for most of General Mills products slightly decrease. This can be attributed to both a change in General Mills high management in 1993, in which the company responded to soaring market shares, and to the fact that General Mills was mostly present in the kids and family segment that was not affected as much by the merger. Overall industry behavior remained stable. Between 1993 and March 1995, industry-wide prices for branded RTE cereal increased moderately. Starting from the second quarter of 1995, I observe a downward trend in industry prices across different firms. In March 1995, two US congressmen started a public campaign to reduce cereal prices, which received a huge media attention. There is evidence that industry prices decreased after this campaign.[10] For this reason, I only consider the period until the first quarter of 1995 for most estimations.

Exogeneity of merger From an estimation standpoint, it is important to discuss concerns and potential effects of merger endogeneity. After the 1988 leveraged buyout of RJR Nabisco, the ownership group accumulated a substantial pile of debt. There is a popular claim that company divestitures were used to reduce the overall debt level. Merger endogeneity would only bias the results if the merger had led to unknown synergies, or if an anticipation of the merger by firms in the industry had led to a change in the competition between firms. There are no factory closures within the first two years of the merger. Therefore, synergies in factory production are unlikely to be achieved. Within the first two years after the merger, I also do not observe a fundamental change in industry pricing due to the merger, which is backed by anecdotal evidence. Thus, there is a relative steadiness in industry behavior in the short run.

[9] See Rubinfeld (2000) for a detailed description.

[10] See Cotterill and Franklin (1999) for a detailed analysis. In April 1996, Post decreased the prices for its products nationwide by 20%, thereby permanently increasing its markets share. This was followed by significant price cuts two months later by General Mills and Kellogg's. Overall, margin over production cost fell by 12% in 1996 due to these actions.

2.3 Empirical Model

My approach has three basic steps. In the first step, I estimate industry demand using a discrete choice model to back out price elasticities. In the second step, I recover marginal costs using first order conditions, which I subsequently use to estimate a cost function. In the third step, I predict post-merger prices, which I then use to estimate either the degree of profit internalization among merging firms or industry conduct. In the industry conduct case the last two steps are repeated in an iterative process.

2.3.1 Demand side

My demand specification is closely related to Nevo (2001). There is a total number of J brands in the market. Denote the number of individual consumers in every market by I, and denote the number of markets by T, where a market is defined as a time-store combination. Using a Random Coefficient Logit model, individual i's indirect utility of consuming product j at market t can be written as:

$$u_{ijt} = x_j\tilde{\beta}_i + \tilde{\alpha}_i p_{jt} + \xi_{jt} + \epsilon_{ijt}, \;\; i = 1, .., I; j = 1, .., J; t = 1, .., T. \tag{2.1}$$

x_j denotes a K-dimensional vector of firm j's observable brand characteristics, p_{jt} denotes the price of product j at market t, and ξ_{jt} the brand-specific mean valuation at market t that is unobservable to the researcher but observable to the firms. Finally, ϵ_{ijt} is an idiosyncratic error term. The coefficients $\tilde{\beta}$ and $\tilde{\alpha}$ are indiviual specific coefficients. These coefficients depend on their mean valuations, on demographics in each region, D_i and their associated coefficients Π, as well as on an unobserved vector of shocks, v_i that is interacted with a scaling matrix Σ:

$$\begin{pmatrix} \tilde{\alpha}_i \\ \tilde{\beta}_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i, \quad v_i \sim \; N(0, I_{K+1}). \tag{2.2}$$

Because not all of the potential consumers purchase a good in each period, I also require an outside good. The indirect utility of not purchasing any product and thus consuming the outside good can be written as:

$$u_{i0t} = \xi_0 + \pi_0 D_i + \sigma_0 v_{i0} + \epsilon_{i0t}.$$

As is common in the literature, I normalize ξ_0 to zero.
Denote the vector of all demand side parameters by γ^D. This vector can be decomposed into a vector of parameters obtained from the linear part of the estimation, $\gamma_1^D = (\alpha, \beta)$, and a vector of parameters obtained from the nonlinear part of the estimation, $\gamma_2^D = (vec(\Pi), vec(\Sigma))$, respectively.
The indirect utility of consuming a product can be decomposed into a mean utility part δ_{jt} and a mean-zero random component $\mu_{ijt} + \epsilon_{ijt}$ that takes into account heterogeneity from demographics and captures other shocks. The decomposed indirect utility can be expressed as

$$u_{ijt} = \delta_{jt}(x_j, p_{jt}, \xi_{jt}, \gamma_1^D) + \mu_{ijt}(x_j, p_{jt}, v_i, D_i; \gamma_2^D)$$

$$\delta_{jt} = x_j\beta - \alpha p_{jt} + \xi_{jt}, \;\; \mu_{ijt} = [p_{jt}, x_j]' * (\Pi D_i + \Sigma v_i), \tag{2.3}$$

where $[p_{jt}, x_j]$ is a $(K+1) \times 1$ vector.
Consumers either buy one unit of a single product or take the outside good. They will choose the option which yields the highest indirect utility. Using these assumptions, this characterizes the set A_{jt} of unobservables that yield the highest utility for a specific choice j:

$$A_{jt}(x_{.t}, p_{.t}, \delta_{.t}, \gamma_2^D) = \{(D_i, v_i, \epsilon_{it}) | u_{ijt} \geq u_{ilt} \forall l \in \{0, .., J\}\},$$

where dotted indices indicate vectors over all J brands. The market shares predicted by the model can then be obtained via integrating over the different shocks, using population moment functions $P^*(\cdot)$:

$$s_j(x_{.t}, p_{.t}, \delta_{.t}, \gamma_2^D) = \int_{A_{jt}} dP_\epsilon^*(\epsilon) dP_v^*(v) dP_D^*(D). \tag{2.4}$$

There are several possibilities to estimate the model that depend on different distributional assumptions. The most general case is a Random Coefficients Logit model. Its main advantage is a very flexible form of substitution patterns. This is desirable because it enables a detailed analysis of the substitution patterns between different brands that does not rely on any model structure. To be able to integrate out the market shares, one needs to make distributional assumptions with respect to the unobservable variables $(D_i, v_i, \epsilon_{ijt})$ and then estimate the model using Generalized Methods of Moments.

2.3.2 Industry technology

The J brands in the industry are produced by $N \leq J$ firms. Each brand can only be produced by one firm. An important part of the model is the representation of marginal cost. As is common in the literature, I assume that marginal costs can be decomposed into cost factors that are observed to the researcher as well as a component that is unobserved to the researcher. I use a linear relationship between marginal costs and the observable cost component. This reflects a relatively weak substitutability of input production factors over the medium- and short-run in the RTE cereal industry. Henceforth, I will omit the market index t in my exposition for notational simplicity. Denote the vector of brand $j's$ observed cost drivers by w_j, and j's unobserved cost component by ω_j. The marginal cost for brand j can be written as:

$$mc_j = w_j \gamma^S + \omega_j \tag{2.5}$$

where γ^S denotes a vector of marginal cost parameters.[11]

2.3.3 Pre-merger industry competition

Each firm f owns a portfolio of brands $\mathbb{F}_f$. I further allow a firm's objective function to potentially depend on other firm's profits. Denote by θ_{ij} the degree to which brand i takes into account brand j's profits when setting its optimal price. This implicitly defines a pre-merger ownership matrix Θ with the entries $\Theta_{jr} = \theta_{jr}$. Each of the parameters within Θ are normalized to lie in between 0 and 1, where 0 implies no internalization of profits, and 1 implies full internalization of profits. Note that since only relative weights matter for the first order condition, this is a normalization without loss of generality. Not allowing for negative conduct parameters also implies that a firm does not derive a positive utility from "ruining" another firm. This leads to a matrix

[11]The baseline specification implies that marginal costs are constant for different output levels. This is a relatively strict assumption, which can be relaxed by introducing scale effects. Denote q_i firm $i's$ total units sold in a period. If one assumes scale effects, i.e. decreasing marginal costs in total production together with a Cobb-Douglas cost function, then this can be written as: $mc_j = \tau log(q_j) + w_j \gamma^S + \omega_j$, where τ is the scale parameter.

of the form

$$\Theta = \begin{pmatrix} 1 & \theta_{12} & .. & \theta_{1J} \\ \theta_{21} & 1 & .. & \theta_{2J} \\ .. & .. & .. & .. \\ \theta_{J1} & \theta_{J2} & .. & 1 \end{pmatrix}.$$

Given Θ, the objective function for product j can be written as:

$$\Pi_j = (p_j - mc_j)s_j\overline{M} + \sum_{r \neq j} \theta_{jr}(p_r - mc_r)s_r\overline{M}, \tag{2.6}$$

where s_r denotes the market share of brand r, and $\overline{M}$ the market size.

The first order condition for product j's objective function with respect to its own price can be written as:

$$s_j(p) + \sum_{r=1}^{J} \theta_{jr}(p_r - mc_r)\frac{\partial s_r}{\partial p_j} = 0. \tag{2.7}$$

I make the assumption that pre-merger, each firm fully internalizes the profits of all of its brands, which implies $\theta_{ij} = 1$ if $i, j \in \mathbb{F}_f$. The marginal costs of all brands in the industry are unobserved to the researcher but common knowledge among firms.

Define $\Omega_{jr} \equiv -\theta_{jr} * \frac{\partial s_r}{\partial p_j}$. Having estimated the demand parameters γ^D, one can already infer the marginal costs of production mc conditional on the form of the ownership matrix Θ:

$$mc(\gamma^D, \Theta, p, x) = p^{pre} - \Omega^{-1}(\gamma^D, \Theta)s^{pre}. \tag{2.8}$$

Substituting the recovered pre-merger marginal cost $mc(\gamma^D, \Theta, p, x)$ into equation (2.5) yields:

$$mc(\gamma^D, \Theta, p, x) = w\gamma^S + \omega \tag{2.9}$$

I estimate the cost function parameter γ^S in equation (2.9) using a two staged least squared estimation to account for the unobserved cost component ω. I discuss instrumentation in Section 2.4. I then use observed post-merger input price drivers, w^{post}, to linearly project the post-merger input price component of marginal cost. Product $j's$ predicted input price component post-merger, $\hat{mc}_j$ can then be written as:

$$\hat{mc}_j^{post}(\gamma^D, \gamma^S, \Theta, p, x, w) = w_j^{post}\gamma^S. \tag{2.10}$$

It is important to note that this linear projection will not contain the unobserved cost component ω_j. Henceforth, for notational simplicity I will drop the observable factors w, x, and p when referring to marginal costs.

The key differences between estimating the degree of profit internalization of merging firms and estimating industry conduct between firms lies in the treatment of the pre-merger ownership matrix Θ and on the assumptions with respect to the merging firms' post-merger profit internalizations.

2.3.4 Post-merger prices when estimating the degree of profit internalization between merging firms

When estimating the degree of profit internalization post-merger, I assume that the form of pre-merger industry competition, Θ, is known to the researcher. I assume that a merger will involve a change in the pricing strategies of the merged entity. Denote by $\tilde{\theta}$ the degree of joint profit maximization between the merging firms, which I assume to be unobserved to the researcher, but common knowledge among firms in the industry. I assume that non-merging firms will not change their competitive behavior after a merger, but will adapt to the change in the merging firms' pricing. Under these assumptions, given the degree of joint profit maximization $\tilde{\theta}$, the model's post-merger prices given the parameters γ^D and γ^S can be written as[12]:

$$\hat{p}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}) = \hat{mc}^{post}(\gamma^D, \gamma^S, \Theta) + \Omega^{-1}(\gamma^D, \Theta, \tilde{\theta})s^{post} + \omega^{post}_{\tilde{\theta}}. \tag{2.11}$$

Here, $\omega^{post}_{\tilde{\theta}}$ reflects the unobserved cost component in the cost function post-merger. Thus, the unobserved error will not be accounted for in the linear projection of the post-merger marginal cost in equation (2.10), but remains a separate term in the post merger equation.

2.3.5 Post-merger prices when estimating industry conduct

When estimating industry conduct, I assume that neither the form of pre-merger conduct, Θ, nor post-merger industry conduct are known to the researcher. A key identifying assumption is that even though the researcher does not know the underlying

[12]Even if the degree of joint profit maximization, $\tilde{\theta}$, is part of a post-merger industry conduct matrix, I will explicitly state it in the model. This is to clearly distinguish the case of estimating the joint profit-maximization parameters from the case when estimating the industry conduct matrix Θ.

form of conduct, he knows exactly how the merger will affect industry conduct. There are two channels for this, namely post-merger profit internalization of merging firms, and how competitors consider the merged entity post-merger. I assume that merging firms will fully internalize their profits after a merger. Non-merging firms compete with each other in the same fashion. The change between merging and non-merging firms after the merger is also known to the researcher, which is summed up in the following assumption.

Assumption 2.1 (Conduct between merging and non-merging firms)**.** *Let f, g be two distinct merging firms, and h a non-merging firm. Let θ_{ik}^{pre} and θ_{ik}^{post} denote the pre- and post-merger conduct parameters between firm i and k, respectively. Then, $\forall i \in \mathbb{F}_f$, $\forall j \in \mathbb{F}_g$, $\forall k \in \mathbb{F}_h$, one of the following three cases holds regarding the conduct between a merging and a non-merging firm:*
a. $\theta_{ik}^{post} = \theta_{ik}^{pre}$; $\theta_{jk}^{post} = \theta_{jk}^{pre}$ *(no change in conduct);*
b. $\theta_{ik}^{post} = \theta_{ik}^{pre}$; $\theta_{jk}^{post} = \theta_{ik}^{pre}$ *(acquiring firm standard).*
c. $\theta_{ik}^{post} = \theta_{jk}^{pre}$; $\theta_{jk}^{post} = \theta_{jk}^{pre}$ *(target firm standard);*

Under Assumption 2.1a, the merger does not change how competitors consider the two merging firms after the merger. In this case, the merging firms will fully internalize the profits after the merger. Under Assumption 2.1b, the fully merged entity is considered and behaves as the acquirer did pre-merger. Assumption 2.1c implies the reverse, i.e. that the merged entity behaves as the target. I do not have to pre-specify the values of the conduct parameters, but just the way in which the parameters change. Other change patterns can also be accounted for, as long as the change in conduct between merging and non-merging firms is known post-merger. Let the underlying form of pre-merger conduct, Θ, be element of the set B. The post-merger conduct can then be expressed via a transition function $b : B \to B$ which maps the pre-merger conduct matrix Θ into the post-merger conduct matrix $\Theta^{post} = b(\Theta)$. This is because all merger-induced ownership transformations are known to the researcher. Under these assumptions, for a specific form of pre-merger industry conduct Θ and transition function b, the predicted post-merger prices given the estimated parameters γ^D and γ^S can be written as:

$$\hat{p}^{post}(\gamma^D, \gamma^S, \Theta, \Omega^{post}, b(\Theta)) = \hat{mc}(\gamma^D, \gamma^S, \Theta) + \Omega^{-1}(\gamma^D, b(\Theta))s^{post} + \omega_{\Theta}^{post}. \tag{2.12}$$

As for the case of estimating the degree of post-merger profit internalization, the unobserved cost component ω_{Θ}^{post} is not accounted for in the marginal cost term $\hat{mc}$, but is rather a distinct component in the pricing equation. Appendix A illustrates different cases of pre- and post merger industry conduct. It is worth mentioning the key difference to conjectural variation models. I treat the conduct parameters Θ as part of the firms' underlying objective functions rather than as behavioral responses with respect to the competitors' price setting behaviors. I then form moment conditions to recover the underlying "level" conduct parameters using a Generalized Method of Moments estimator.

2.4 Identification

In this section, I will present identification results for the different stages of the model, and in particular the two supply side estimation approaches. Section 2.4.1 discusses identification restrictions for the demand and cost function estimation. Section 2.4.2 discusses the identification restrictions when estimating merging firms' profit internalization parameters. Section 2.4.3 focuses on identification restrictions when estimating industry conduct.

My estimation method requires identification of three sets of parameters: demand-side parameters, γ^D, cost parameters γ^S, and the supply side parameters, which amount either to the degree of profit internalization $\tilde{\theta}$, or to the industry conduct Θ. The correlation between price and both unobserved brand and cost characteristics requires instrumentation for each brand in the demand and pricing equations, respectively.

2.4.1 Model identification of demand and cost parameters

Identification of demand parameters

Denote by $\xi(\gamma^D, x, p)$ the structural error term vector that consists of the market-specific unobservable brand valuations for all brands. Regarding the demand side, I assume that when being assessed at the true demand parameter values γ_0^D, this error term is uncorrelated with respect to a M_ξ-dimensional set of exogenous demand side instruments, Z_ξ. This leads to the identification restriction:

$$E[Z_\xi' \xi(\gamma_0^D, x, p)] = 0. \tag{2.13}$$

Note that I implicitly assume that the demand can be estimated independently of the marginal cost and supply side parameters, respectively. The orthogonality conditions would be violated if industry conduct or a change in production costs, not prices, would influence consumer choice through the unobserved brand-specific component.[13] I assume that the observable product characteristics of the different goods are exogenous, and therefore do not respond to changes in industry pricing. Also accounting for potential brand replacement or additional brand introductions would make traction of the full model nearly impossible. Because of the inherent endogeneity between price and unobserved brand characteristics, I need to find adequate instruments for the demand estimation. I use two different sets of instruments to do so.

My first set of instruments relies on production cost shifters. The economic assumption is that input cost variation should be correlated with variation in prices, but not with consumers' preferences for unobservable product characteristics. I use both cost factors that affect all products in similar fashion, such as labor costs, packaging, and transportation, as well as factors that differ among products, such as interactions between product characteristics and input prices for wheat, sugar, and corn. My second set of instruments is the ownership change itself. As argued above, a merger should cause a change in industry prices. Similar to a cost shift, one can assume that the merger affects prices, but not the demand characteristics. This assumption would be violated if the merger caused a change in brand value which would affect the ξ's of the merging firms. Because the actual brand names of the cereals involved did not change after the merger, such an effect seems unlikely.

Identification of cost parameters

Conditional on a specific form of pre-merger industry conduct Θ, I can back out the marginal cost via a first order condition and then regress them on observable product characteristics combined with input prices.[14] This allows me to predict the input cost component $\hat{mc}$ of the post merger marginal costs using post-merger input price data and the estimated parameters. I make the implicit assumption that firms

[13] This assumption would be violated if the merger caused a change in the perceived "brand values" of the merged entities, which would affect the ξ components in the demand equation.

[14] Explicitly using input prices to estimate a cost function is a difference to previous studies in the ready-to-eat cereal industry.

cannot substitute between different input goods. The recipes and production processes for a specific product in the ready-to-eat cereal industry remain constant over time, such that this assumption is likely to hold in the medium and short term. My identifying assumption concerning the marginal cost component pre-merger is that the structural error term vector representing unobserved cost characteristics ω^{pre} is uncorrelated to a M_ω-dimensional set of exogenous instruments $Z_{\omega^{pre}}$:

$$E[Z'_{\omega^{pre}}\omega_j^{pre}(\gamma^D, \gamma_0^S, \Theta)] = 0, \tag{2.14}$$

where γ_0^S reflects the true parameter value for γ^S. Together with the change in ownership, the marginal cost estimates will influence the predicted post-merger prices in the market.

Berry et al. (1995) argue that the computation of the optimal set of instruments when only conditional moment conditions are available is very difficult and numerically complex. As a less computationally demanding approximation, they use polynomials resulting from first order basis functions of the product characteristics. The validity of these basis functions as instruments relies on exchangeability assumptions of firms' own characteristics with respect to permutations in the order of competitors' product characteristics. Because I allow for the possibilty of collusion among firms, this changes the structure of potential Nash equilibria. The brand specific unobservable marginal cost component ω may be correlated with unobservable product characteristics. Therefore it is essential to look for instruments that are correlated with marginal costs, but not with the structural cost error. To account for the effects of unobserved cost drivers on prices, I use first order basis functions of the own brand characteristics, own firm characteristics, and competitors' characteristics.

2.4.2 Model identification of post-merger profit internalization parameter

Setting equation (2.11) equal to the observed post-merger prices p^{post}, and solving for the unobserved post-merger cost-component vector $\omega_{\tilde{\theta}}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta})$, yields:

$$\omega_{\tilde{\theta}}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}) = p^{post} - \hat{mc}^{post}(\gamma^D, \gamma^S, \Theta) - \Omega^{-1}(\gamma^D, \Theta, \tilde{\theta})s^{post}. \tag{2.15}$$

As an identification restriction for the degree of joint profit maximization, I use orthogonality conditions between the residual of observed and predicted post-merger prices, which results in the structural error $\omega_{\tilde{\theta},j}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta})$ for a product j, and a $M_{\tilde{\theta}}$-dimensional matrix of instruments $Z_{\tilde{\theta}}$.

The model consists of a system of J equations for the different products whose prices are functions of the profit internalization parameters, $\tilde{\theta}$. The main identification task is to find meaningful moments that allow to identify the parameters. Using the difference between the predicted and observed post-merger prices of all brands would result in only one moment, which would render estimation of more than one parameter infeasible. I instead make use of orthogonality restrictions to generate two additional sets of moments. First, because I treat product characteristics with respect to demand, x, as exogenous with respect to firm behavior in the short run, I can use them as instruments. This is analogous to the identification of the production cost. Second, an increase in consumer income will have a positive demand effect at a given price. If such an income shock does not translate in higher labor costs, then the shock should be uncorrelated with the unobserved post-merger cost component vector ω^{post}. I use regional income data and local consumer price indexes as additional instruments. This leads to the identification restriction

$$E[Z_{\tilde{\theta}}' \omega_{\tilde{\theta}}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta})] = 0. \tag{2.16}$$

Making use of the structural error term $\omega_{\tilde{\theta}}^{post}$ to construct orthogonality conditions to identify the supply side parameters does not rely on the linear cost function specification and only requires separability between the observed parameters γ^S and the unobserved cost component $\omega_{\tilde{\theta}}^{post}$. Appendix B derives the structural error term under the assumption of a logarithmic cost function.

2.4.3 Model identification of industry conduct parameters

Comparing to the case when estimating industry conduct, the main difference is that one estimates the industry ownership matrix Θ instead of the internalization parameters $\tilde{\theta}$. Equating equation (2.12) with the observed post-merger prices p^{post}, and solving for the unobserved post-merger cost error vector $\omega_{\Theta}^{post}(\gamma^D, \gamma^S, \Theta; b())$ yields:

$$\omega_{\Theta}^{post}(\gamma^D, \gamma^S, \Theta; b(\cdot)) = p^{post} - \hat{mc}^{post}(\gamma^D, \gamma^S, \Theta) - \Omega^{-1}(\gamma^D, \Theta, b(\Theta))s^{post}. \tag{2.17}$$

Besides the conventional first- and second-order polynomials of the observed product characteristics x, I use additional cross-firm polynomials which indicate the proximity between different firms' product portfolios. From a theoretical viewpoint, the proximity of two firms' brand portfolios should at least partly determine the potential profits from collusion between those firms. Furthermore, this should also determine the maximum degree of sustainable collusion between them. Denote x_{sj} as the s^{th} component of product j's observed product characteristic vector x_j. Denote by $h(x_{sj}, f)$ a $J \times 1$ row vector whose entries consists of the average of brand characteristic x_s between firm f and the average brand characteristics of each product's firm. Thus, the entries of this vector are 0 whenever a brand belongs to firm f. In case a brand belongs to firm $g \neq f$, the entry will be $\frac{1}{J_f}\sum_{i \in \mathbb{F}_f} x_{si} - \frac{1}{J_g}\sum_{j \in \mathbb{F}_g} x_{sj}$, where J_f denotes the number of brands of firm f. Under the assumption that cooperation between firms also depends on the proximity of the brand portfolios, given the correct form of conduct, the unobserved cost components of a firm should on average be uncorrelated with the differences in average brand characteristics $h(x_{s.}, .)$. For a given product characteristic s, this yields the additional moment restrictions

$$E[\omega_{\Theta}^{post}(\gamma^D, \gamma^S, \Theta; b(\cdot))h(x_{s.,}, f)] = 0, \forall f \ \in \{1, .., N\}.$$

Per characteristics used, this will results in $N - 1$ additional moment restrictions. As in the profit-internalization case, I will also use data on disposable income as an instrument for the unobserved cost component.
Stacking the different instruments into the M_{Θ}-dimensional instrument matrix Z_{Θ}, this yields the following identification restrictions:

$$E[Z'_{\Theta}\omega_{\Theta}^{post}(\gamma^D, \gamma^S, \Theta; b(\cdot))] = 0. \tag{2.18}$$

One key assumption is that industry conduct is known among firms. Relaxing this assumption would cause two problems. First, this would make the assumption on symmetric behavior between two different firms harder to sustain. Second, I would have to specify beliefs of the different firms regarding other firms' behavior, which could not be identified.

Relationship to Corts' (1999) Critique Previous research has used a conjectural variation approach in order to identify industry conduct, see for example Bresnahan (1989). In these models, a firm forms a "conjecture" about the responses of their competitors towards an increase in its own quantity. In this context, a conjecture can be seen as a reduced-form game theoretic best response function in symmetric quantity setting games. Corts (1999) critically discusses the identification of conjectural variation parameters. He shows that a conjectural variation parameter only estimates the marginal responsiveness of the marginal cost function with respect to changes in a demand shifter. As a researcher, one is however interested in the average slope of the marginal cost function instead of the marginal slope. My approach differs significantly from the conjectural variations approach and is not subject to this critique. In my framework, each firm sets prices for its portfolio of brands instead of quantities. Instead of forming conjectures about other brands' reactions, each firm's underlying objective function includes preferences for profits of other firms, thus allowing for cooperation among different firms. The preference parameters with respect to other firms' profits are essentially the conduct parameters I am interested in. I assume that these conduct parameters, as well as the marginal costs of all brands, are common knowledge in the industry, but not observed by the researcher. Using first order conditions of all brands' objective functions, my identification strategy allows to estimate both marginal cost parameters and the level conduct parameters. These amount to the "as-if conduct parameters" in Corts (1999).
Corts' also criticizes the static game character of conventional conduct estimation models. My approach is not fully exempt from this critique. I partially account for industry dynamics by modeling the merger-induced industry change. Nonetheless, my static approach may not detect certain dynamic collusion patterns. One big advantage of a static approach is a higher degree of tractability. Modeling repeated games makes identification of conduct even more difficult due to a larger set of potential dynamic equilibrium strategies. With my approach, I am also able to identify patterns of full collusion as well as patterns of collusion between only a subset of firms.

Rank conditions for industry conduct

This section provides identification results for different specifications when estimating continuous conduct parameters "directly". This is opposed to the menu approach, which selects among different non-nested models without estimating conduct parameters.

Recall the assumptions made on firms' own-profit maximization. As in standard unilateral merger models, I also assume that a merger does not change the behavior between non-merging firms. There are furthermore some global assumptions that reduce the parameter space which I will discuss in detail.

I only consider cases in which a firm treats all brands of a specific competitor's firm in the same way. This excludes the possibility that single brands of different firms collude while others play against each other competitively. From a pure rank condition perspective the number of parameters I would have to estimate when accounting for brand-specific collusion between firms would easily exceed the number brands in the market. This makes it impossible to identify the parameters.

Bilateral symmetry between firms One way to reduce the number of parameters to be estimated is to restrict the model to cases in which all brands of two firms play against each other in the same way. As a consequence, all brands have the same cross-conduct parameters for all of their brand pairs. This still allows for partial collusion between two firms, but does not allow for more elaborate strategies, such as for example collusion only between some brands of two firms. In terms of the parameter space, this reduces the number of cross-conduct parameters to $\frac{N(N-1)}{2}$.

Proposition 2.1 (Necessary conditions for identification under bilateral symmetry between firms)**.** *Suppose Assumption* 2.1 *holds, and that for distinct firms* f, g, $\theta_{ij} = \theta_{ik} = \theta_{ji} = \theta_{ki}\ \forall i \in \mathbb{F}_f, \forall\ j,k \in \mathbb{F}_g$. *Then industry conduct is identified only if the number of firms is sufficiently small compared to the number of products, i.e. if* $\frac{N(N-1)}{2} \leq J$.

Proof: See Appendix E.

Same responsiveness to all cross-firm brands Another possibility is a case in which each firm behaves in the same way to all of its competitors.

The advantage of this specification is that it reduces the number of parameters to only N different cross-conduct parameters. However, there are also several problems associated with the assumption. First, it is again no longer possible to detect partial collusion between a subset of firms in the industry. Second, there is a consistency problem with respect to a mutual responsiveness: Under this assumption, it can be possible that firm 1 is acting collusively with firm 2, and firm 2 on the other hand acts competitively towards firm 1, something which is hard to justify from an economic perspective.

Proposition 2.2 (Necessary conditions for identification under same responsiveness to all cross-firm brands)**.** *Suppose Assumption* 2.1 *holds, and that for distinct firms* f, g, h , $\theta_{ij} = \theta_{ik}$ $\forall i \in \mathbb{F}_f, \forall\ j \in \mathbb{F}_g, \forall\ k \in \mathbb{F}_h$. *Then rank conditions are met only if* $N \leq J$.

Proof: See Appendix E.

It is easy to see that the necessary rank conditions hold trivially. It can still be the case however that there are two or more identical conduct equations, which would violate identification.

Same responsiveness between all firms The most restrictive specification assumes that the cross-conduct parameters are identical for all brands in the market. The biggest advantage is that this returns a single cross-conduct parameter instead of a complicated matrix, and thus always meets the rank conditions. One disadvantage is that very often this parameter will severely restricts the set of estimable economic models. For example, one will not be able to test for partial collusion in the market, or for differences in competitive behavior between different firms.

Proposition 2.3 (Necessary conditions for identification under same responsiveness between all firms)**.** *Suppose for distinct firms* f, g, h , $\theta_{ij} = \theta_{ji} = \theta_{ik} = \theta_{jk} = \theta_{kj}$ $\forall i \in \mathbb{F}_f, \forall\ j \in \mathbb{F}_g, \forall\ k \in \mathbb{F}_h$. *Then the rank condition for industry conduct is always met.*

Proof: See Appendix E.

Overall, the direct approach requires more structure and a larger parameter space than conventional selection methods. This is because explicitly accounting for conduct parameters requires more degrees of freedom. Therefore, I will provide results specifically tailored for the different assumptions provided in the beginning of this section. Clearly, the most important trade-off is the one between the allowed flexibility of industry conduct and the number of parameters that have to be estimated. Table 2.3 sums up the necessary rank conditions for all cases.

Identifying industry conduct via product entry or exit Besides using a merger as an identification strategy for estimating industry conduct, one can also think about using other structural changes. Concerning product entry, there is the problem of comparing competition with and without the entrant. While one can still make the assumption that entry does not change how existing brands compete with each other, one has to define how a new product will interact with the existing products.

Unlike product entry, using product exit as an identification strategy is still feasible. However, one has to ask why a product will exit. One reason can be that it is just not profitable, which will then probably imply that its impact on the market is relatively low. Therefore, a reduction of the brand space would not result in a big shift for firms strategies. Another possibility would be that a brand is profitable on its own, but it would be more profitable for a multi-brand firm to exit the product out of the market. This would result in an endogeneity problem when estimating conduct using product exit.

2.5 Estimation

2.5.1 Demand estimation

I use the technique of Nevo (2001) to recover the structural demand side parameters γ^D and the unobservable error term $\xi(\gamma^D, x, p)$. Using Nevo's estimation strategy on the demand side allows me to estimate all the demand side parameters independently of the supply side. I solve for the mean utility level across all brands at market t, $\delta_{.t}$, as to match the empirical market shares $s_{jt}(x_{.t}, p_{.t}, \xi_{.t}, \gamma^D)$ from equation (2.13) with the actual market shares s_{jt} observed in the data. Following equation (2.13) the objective is to find:

$$\hat{\gamma}^D = \arg\min_{\gamma^D} \xi(\gamma^D, x, p)' Z_\xi \tilde{A}_\xi^{-1} Z_\xi' \xi(\gamma^D, x, p); \tag{2.19}$$

where $\tilde{A}_\xi^{-1}$ is an estimate of the asymptotically efficient covariance $E[Z_\xi' \xi(\tilde{\gamma}^D, x, p)\xi(\tilde{\gamma}^D, x, p)' Z_\xi]$, given demand parameters $\tilde{\gamma}^D$ obtained from the first-stage GMM estimation.

Defining the market size is an important assumption, for it has implications on the different market shares and also on the differences between markets. I assume that the market size is correlated with store specific characteristics. I compute the market size of a specific store as a function of the average total sales of all supermarket products sold in this store.[15]

Estimates Table 2.4 shows results for a Random Coefficients Logit demand model. In this specification I include random coefficients for price, a constant, and sogginess of cereal. I furthermore include coefficients for sugar content, content of refined grains, segmentation dummies, a time dummy as well as firm dummies. The inclusion of firm dummies reflects controlling for firm-specific valuations, i.e. accounting for a fixed firm value rather than a brand value. Furthermore, I use demographic data on mean income, income standard deviation, household size and on number of small children to interact them with the random coefficients. The results show a negative relationship between income and price sensitivity, which is consistent with higher markups in high income neighborhoods. Price sensitivity also interacts negatively with the number of small children, which might account for their responsiveness to advertising. However, both demographic interaction coefficients are not statistically significant. Appendix C shows details about the estimation routine and other computational issues.

As a robustness check, I also estimate different variants of a multinomial Logit model. Table 2.5 shows demand side estimation results for several specifications of the multinomial Logit model. I use input prices and prices of other zones together with the ownership change as instruments for the sales price. When also including firm dummies yields a more elastic demand curve than specification (6) without instruments, however, the price coefficients are relatively close to each other. A bigger difference occurs between specifications that include and do not include firm dummies. This can be seen by comparing specifications (1) and (2) to (3) and (4). Overall, all of the price coefficients are lower in absolute magnitude than the mean coefficient of a Random Coefficient Logit estimation. This suggest that the random coefficient model is able to capture some of the consumer heterogeneity through interacting demographic variables which increases the price coefficient in absolute terms.

Demand Elasticities Individual market shares depend on the mean utility as well as on the random and demographic components. Product j's market share for individual

[15]Because revenue from ready-to-eat cereal only amounts to a very small fraction of the total revenue generated in a store, the endogeneity between the market size variable and the cereal prices is negligible.

i at market t can be written as:

$$s_{ijt} = \frac{exp(\delta_{jt} + \mu_{ijt})}{\sum_{k=0}^{J} exp(\delta_{kt} + \mu_{ikt})} \tag{2.20}$$

Integrating over the whole distribution of individuals yields the aggregated market shares from the model. The cross-price elasticity between goods j and k at market t, η_{jkt}, can be written as

$$\eta_{jkt} = \begin{cases} \frac{-p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) dP_D(D) dP_v(v) & j = k \\ \frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} dP_D(D) dP_v(v) & j \neq k \end{cases} \tag{2.21}$$

When using the random coefficients model, one needs to compute the individual market shares using the model structure in equation (2.20). Table 2.6 and Table 2.7 show the mean elasticities over all markets for the baseline random coefficient Logit specification. The own-price elasticities are highly negative for all firms, with the exception of Kellogg's Just Right Fruit, which has an absolute own-price elasticity lower than one. One potential reason for this can be an increased popularity of adult cereals over the period in my dataset, such that both price and demand went up for this cereal at the same time.

There is furthermore significant variation in different brands' substitution patterns, which is related to the type of cereal. Adult cereals, such as Kellogg's Just Right and Kellogg's Nutri Grain, exhibit much lower cross-price elasticities than kids cereals, as for example Kellogg's Fruit Loops or General Mills Honey Nut Cheerios with coefficients higher than .1. Overall, the substitution patterns are relatively close to previous industry estimates, see for example Nevo (2001).

2.5.2 Post-merger profit internalization

I first outline each step of the estimation algorithm when estimating the degree of joint profit maximization.

Estimation algorithm

1. **Estimate demand parameters γ^D:** In a first step, I estimate the demand parameters without having to specify the supply side.

2. **Recover pre-merger marginal costs using first order conditions, estimate cost function, and predict post-merger prices** Under the assumption of a known pre-merger industry ownership matrix Θ, I use equations (2.5) and (2.8) to back out pre-merger marginal costs. Using variation in the input costs w over time, I then predict the post- merger input price component of marginal costs, $\hat{mc}$.

3. **Pick degree of profit-internalization $\tilde{\theta}$, predict post-merger prices, and compute appropriate moments** I predict the markup firms charge conditional on a specific $\tilde{\theta}$ value. Together with the estimated post-merger input price component of marginal costs from step 2., I then predict the post-merger prices.

4. **Repeat** 3. **until GMM criterion is minimized** I recover the post-merger unobserved cost component $\omega_{\tilde{\theta}}^{post}$ and interact it with the instruments $Z_{\tilde{\theta}}$. I estimate the model using Generalized Method of Moments (GMM) to find the parameters that minimize the weighted moment criterion.

I assume that at a given point in time, marginal costs are constant across all stores. All stores are within the same metropolitan area and are operated by the same retailer. Therefore, the only channels through which marginal costs could differ across stores are either a difference in the retail margin across stores, or a difference in distribution costs. I do not find evidence for structural differences regarding the retail margin in the data. Differences in the distribution costs also do not seem likely because of the relative proximity of the stores.[16]

In the second step of my estimation procedure, I use the marginal costs that were backed out conditional on a specific form of industry conduct and estimate the marginal cost equation (2.8) via minimizing the following objective function:

$$\hat{\gamma}^S = \arg\min_{\gamma^S} \omega^{pre}(\gamma^D \gamma^S, \Theta)' Z_\omega A^{-1}_{\omega^{pre}} Z'_\omega \omega^{pre}(\gamma^D, \gamma^S, \Theta), \tag{2.22}$$

where $A^{-1}_\omega = Z'_\omega Z_\omega$, therefore this amounts to a linear GMM estimator.

[16]There may be differences in store-specific fixed costs due to differences in rents or wages between the store locations. Such effect would not translate in marginal cost differences, but may be a channel for cost synergies after a merger.

The brand specific unobserved marginal cost component ω_j^{pre} may be correlated with unobservable product characteristics. Therefore it is essential to look for instruments that are correlated with marginal costs, but not with the error term. To account for the effects of unobserved cost drivers on prices, I use first order basis functions of the own brand characteristics, own firm characteristics, and competitors' characteristics. This relies on an exchangeability argument of product characteristics when facing a unique Nash equilibrium, see for example Berry et al. (1995).

Having obtained the demand side coefficients γ^D and the cost parameters γ^S for the given form of pre-merger industry conduct Θ, I estimate the degree of profit-internalization $\tilde{\theta}$ by minimizing the GMM objective function:

$$\hat{\tilde{\theta}} = \arg\min_{\tilde{\theta}} \omega_{\theta}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}) Z_{\tilde{\theta}} \tilde{W}_{\tilde{\tilde{\theta}}}^{-1} Z_{\tilde{\theta}}' \omega_{\tilde{\theta}}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}), \tag{2.23}$$

where $\tilde{W}_{\tilde{\theta}}(\tilde{\tilde{\theta}}, \tilde{\gamma}^S)$ is a consistent estimate of the covariance matrix $E[Z_{\tilde{\theta}}' \omega_{\tilde{\theta}}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \Theta, \tilde{\tilde{\theta}}) \omega_{\tilde{\theta}}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \Theta, \tilde{\tilde{\theta}})' Z_{\tilde{\theta}}]$ for a given first-stage parameter vector $\tilde{\tilde{\theta}}$. The moments consist of the empirical residuals $\omega_{\tilde{\theta}}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta})$ interacted with the specific instruments $Z_{\tilde{\theta}}$, as described in Section 2.4.

There are several advantages from using the actual post-merger prices instead of simulating a post-merger price equilibrium. First, when simulating for a new price equilibrium, one needs to make an assumption regarding competition in the market. Already without estimating industry conduct, this is computationally demanding. Furthermore, it does not make use of all the available post-merger data, i.e. market shares and prices. Second, by using post-merger price simulation, one also risks averaging out specific competitive patterns and introduces a simulation error.[17]

Estimates Table 2.8 shows the cost function estimates for four symmetric forms of industry competition, ranging from multi-brand Nash competition, i.e. $\Theta = 0$, to full collusion, i.e. $\Theta = 1$. The median marginal costs implied by the model lie between \$.114

[17]One important issue concerns the standard errors. Because of the sequential character of the estimation routine, I have to account for the demand estimation error when estimating standard errors for marginal costs and the supply side parameters, respectively. I account for these effects by using a second estimation routine. After having obtained the parameter estimates of the estimation algorithm, I estimate a sequential model that estimates all of the parameters simultaneously. This has the advantage of increasing efficiency of the estimation, as well as yielding consistent standard errors. Its disadvantage is the computational power required for this estimation. So far, I do not account for the estimation error when computing standard errors, which can potentially cause a bias in these estimates.

per serving for multi-product Nash pricing and \$.072 for full collusion, implying median markups between 40.8 and 63.1 percent. The cost function estimations show that especially the influence of wheat on overall marginal costs is decreasing in the degree of industry competition, Θ. Figure 2.6 shows the development of profit internalization parameters over time for four different forms of industry competition, ranging from multi-brand Nash competition to partial collusion, i.e. $\Theta = .5$, between all firms pre-merger.

The results indicate an increasing degree of profit internalization over the first six quarters for all three industry specifications. The parameter values are the highest for the Nash specification and are decreasing in the degree of industry cooperation. In the last two quarters of 1994, there is a sharp drop in the profit internalization, which is followed by a sharp increase over the last year.

Overall, except for the drop in joint profit internalization in 1994, the results are consistent with an increase in profit internalization over time. However, the point estimates as stated in Table 2.9 for this estimation technique are not statistically significant. This is due to a relatively flat GMM objective function at the optimum parameter values.

2.5.3 Industry conduct estimation

When estimating industry conduct, I have to iterate the processes of recovering pre-merger marginal costs, predicting post-merger marginal cost using a cost function estimation, and computing the industry conduct moments. This is because my object of interest, i.e. the pre-merger conduct matrix Θ, influences the implied marginal cost in the industry. Overall, the above steps can be decomposed into two parts. I use a nested two-step routine on the supply side. In the first step, I back out marginal cost conditional on a specific form of conduct as the outer loop. In the second step, I recover the supply side parameters by regressing the backed out marginal costs on the observable cost characteristics while controlling for unobserved brand characteristics. First, I will outline the conduct estimation routine.

1. **Estimate demand parameters** Using the instruments discussed above, I estimate the demand parameters, without having to specify supply-side competition.

2. **Pick Θ given the identification restrictions**

3. **Infer marginal costs and predict post-merger prices for given choice of Θ, and compute appropriate moments** Having estimated the demand side parameters, I can infer the marginal costs of production conditional on the form of conduct Θ using proper instruments. Using post-merger input cost data and the estimated cost-parameters, I can predict post-merger marginal costs. Given the conduct matrix Θ and the estimated demand parameters γ_D, I can then predict post-merger prices given Θ.

4. **Repeat steps 2-3 until GMM criterion is minimized** I recover the post-merger unobserved cost component ω_{Θ}^{post} and interact it with the instruments Z_{Θ}. I estimate the model using Generalized Method of Moments (GMM) to find the conduct parameters that minimize the weighted moment criterion.

In the second step of my estimation procedure, I use the marginal costs that were backed out conditional on a specific form of industry conduct and estimate the cost function from equation (2.8) via minimizing the following GMM objective function:

$$\hat{\gamma}^S = \arg\min_{\gamma^S} \omega^{pre}(\gamma^D, \gamma^S, \Theta)' Z_\omega \tilde{A}^{-1}_{\omega^{pre}} Z'_\omega \omega^{pre}(\gamma^D, \gamma^S, \Theta). \tag{2.24}$$

$\tilde{A}^{-1}_{\omega}$ is a consistent estimate of the covariance $E[Z'_\omega \omega^{pre}(\hat{\gamma}^D, \tilde{\gamma}^S, \Theta)\omega^{pre}(\hat{\gamma}^D, \tilde{\gamma}^S, \Theta)' Z_\omega]$ for a given first-stage parameter vector $\tilde{\gamma}^S$.

Having obtained the demand side coefficients γ^D and the cost parameters γ^S for any form of industry conduct, I estimate the conduct parameters Θ by minimizing the GMM objective function. The moments consist of the empirical residuals $\omega_{\Theta}^{post}(\gamma^D, \gamma^S, \Theta; b(\cdot))$ interacted with the specific instruments Z_Θ, as described in section 3.

Then the GMM objective in can be written as:

$$\hat{\Theta} = \arg\min_{\Theta} \omega_{\Theta}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \Theta; b(\cdot)) Z'_\Theta \tilde{W}_\Theta^{-1} Z'_\Theta \omega_{\Theta}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \Theta; b(\cdot)), \tag{2.25}$$

where $\tilde{W}(\tilde{\Theta}, \gamma^S)$ is a consistent estimate of the covariance $E[Z'_\Theta \omega_{\Theta}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \tilde{\Theta}; b(\cdot))\omega_{\Theta}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \tilde{\Theta}; b(\cdot))' Z_\Theta]$ for a given first stage conduct matrix $\tilde{\Theta}$. Here, $\hat{\gamma}^S$ denotes the cost estimates from the second step that are conditional on a specific form of conduct.

Estimates Table 2.10 shows estimation results when estimating a single industry conduct parameter. The obtained parameter value is 0.708. It is interesting to compare the implied price cost-margins to those from a multi-product Nash pricing supply side model. Under multi-product Nash pricing, all of the markup can be attributed to product differentiation, and not to cooperative effects. When estimating a single conduct parameter, 25.6 percent of the price-cost margin is attributable to cooperative behavior between firms.

Table 2.11 shows the results when imposing symmetry in a firm's behavior towards all of its rivals. One can see that the two biggest players, General Mills and Kellogg's, act in the most cooperative behavior, while smaller companies act more competitively. According to this specification, General Mills acts fully cooperatively, with a parameter value of 0.98. Under this specification, 17.4 percent of the markups are attributable to cooperative behavior.

To account for even more heterogeneity with respect to behavior across firms, Table 2.12 shows the conduct estimation results under the assumption of bilateral brand symmetry. The parameter estimates show a lot of heterogeneity in the parameter values, however, under this specification, none of the parameters are statistically significant. The implied median price-cost margins from the estimation are 14.3 percent higher than the median multi-brand Nash price cost margins. This is lower than under a single conduct parameter specification, reflecting the heterogeneity across different firm pairs.

2.6 Extensions

In this section I present extensions to my basic framework that address several merger related issues.

2.6.1 Supply side selection methods

The menu approach selects the best fit among a discrete set ("menu") of supply side models, for example multi-brand Bertrand-Nash competition or full collusion among all firms. This approach does not include any explicit conduct parameters, but rather fully pre-imposes the form of competition, which often relaxes identification problems. In practice, there are two popular ways to select among different non-nested industry specifications. However, both have significant weaknesses.

The first method compares the marginal cost estimates of the different supply side specifications with cost estimates from other sources, such as accounting data, see for example Nevo (2001). At first sight this seems to be an intuitive way to select the most appropriate specification from the data. This approach, however, has several disadvantages. First, outside cost estimates are not always available, or do not have a reliable economic interpretation. Second, such data is often only available on a very aggregate level, which makes a detailed industry introspective nearly impossible.[18] Third, if different specifications yield similar cost estimates, it is not clear how one can use these results for a reliable model selection.

The second method uses forms of non-nested selection test in combination with pre-merger data to look for the supply specification that is closest to the true data generating process, see for example Vuong (1989) or Rivers and Vuong (2002). In vertical relations frameworks, such tests are relatively successful for selecting among different non-nested models, see for example Bonnet and Dubois (2010) for a detailed exposition. When using only pre-merger data in a horizontal framework, however, in practice there is the problem that such tests have relatively low predictive power, due to only very limited variation between the different pre-merger model specifications.

I exploit changes in ownership as well as variation in the input cost data to select among different non-nested horizontal models.Using pre-and post merger data in combination with non-nested supply side models provides a tractable in-sample test. Testing can be done using a J-Test or using a variant of the Rivers and Vuong (2002) test. Appendix *D* shows the estimation details for applying the Rivers and Vuong (2002) approach.

Table 2.13 shows the results of using the Rivers and Vuong (2002) approach for testing different non-nested hypotheses against each other. I use five non-nested specifications in which each firm play symmetrically against each other, with values 0, .25, .5, .75, and 1. The results show that the non-nested test clearly rejects hypotheses of low industry cooperation against the hypotheses of high cooperation among firms. Overall, this approach would select a fully collusive model.

[18] Nevo (2001) states that his data is not sufficiently detailed to test for collusion among a subset of firms.

2.6.2 Direct estimation of synergies

From an antitrust viewpoint, the magnitude of synergies plays a key role for the welfare and consumer welfare effects of horizontal mergers, see for example Farrell and Shapiro (1990) and Nocke and Whinston (2010).[19] To my knowledge there is no approach that uses a differentiated goods framework to estimate the magnitude of merger related marginal cost synergies directly. I propose the following estimation method. Assume that industry conduct is known in an industry pre-merger and post-merger, and merging firms fully internalize their profits. When accounting for the change in price elasticities and the change in conduct after the merger, I can back out marginal costs both pre-merger and post merger via the vector of first-order conditions. When conduct and demand is known, the only systematic change can occur with respect to marginal costs. I will use information on the timing of the merger to assess the impact of the merger on marginal costs of the merging firms, which in economic terms reflects cost synergies.

Denote Θ^{pre} and Θ^{post} the known pre- and post-merger industry conduct, respectively. Then, using equation 2.8, I can back out the pre-merger and post-merger marginal cost from the model:

$$mc^{pre} = p^{pre} - \Omega^{-1}(\Theta^{pre})s^{pre}$$

$$mc^{post} = p^{post} - \Omega^{-1}(\Theta^{post})s^{post}$$

Define mc^{all} as the marginal cost vector both pre- and post-merger: $mc^{all} \equiv [mc^{pre}; mc^{post}]$.

I will now propose three different specifications one can use to estimate for synergies between merging firms given pre- and post-merger data.

1. Synergies in observable cost characteristics If one assumes that synergies affect all observable brand characteristics in the same way, but do not affect unobserved brand characteristics, then one can estimate the following equation:

$$mc_{jt} = (1 + \kappa \mathbb{1}_{merge,j})(\gamma^S w_j) + \omega_j + \epsilon_{jt}, \tag{2.26}$$

[19] One example for an industry with significant synergies is the beer industry. After the 2005 Coors-Molson merger, the company stated that it made $ 66 million worth merger related synergies in its first year as joined entity.

where κ represents the change in the observable brand characteristics on input prices, and $\mathbb{1}_{merge}$ is an indicator function equal to one if the brand belongs to one of the merging firms in the post-merger periods.

2. Synergies in unobservable cost characteristics If one assumes that synergies will affect only the unobserved brand specific component, then one can estimate the following equation:

$$mc_{jt} = \gamma^S w_j + \mathbb{1}^{pre} \omega_j^{pre} + (1 - \mathbb{1}^{pre}) \omega_j^{post} + \epsilon_{jt}, \quad (2.27)$$

where $\mathbb{1}^{pre}$ is an indicator function equal to one in pre-merger periods.

3. Synergies in output A third possibility to account for synergies is to test for returns to scale in total firm output. This can account for increasing returns to scale in distribution cost or advertising. In this case, the marginal cost for brand j of firm f can be written as

$$mc_{jt} = \tau \sum_{i \in \mathbb{F}_f} log(q_i) + \gamma^S w_j + \omega_j + \epsilon_{jt}, \quad (2.28)$$

where q_i denotes the total quantity sold of brand i. Since I only observe output in one metropolitan area and brand, I have to assume that my data is representative for the average output over all retailers in the industry.

2.7 Conclusion

This chapter proposes a framework to estimate the degree of joint profit maximization between merging firms and the form of industry conduct in the ready-to-eat cereal industry. The merger-induced ownership change serves as an important variation to identify firm behavior in the industry.

The availability of pre- and post-merger industry data allows me to estimate the degree of joint profit maximization rather than to assume it. The results shed light on the question of cooperation within a firm after a merger. The empirical descriptive findings show a partial pricing adjustment by the merging firms immediately after the merger. Furthermore, the structural estimation suggest an overall increase in the joint profit maximization over the first 10 quarters. The results are in line with informational and contractual frictions in a post-merger integration period.

The merged firm's pricing potential also crucially depends on the form of industry conduct. The biggest difference of my conduct estimation approach compared to other approaches lies in exploiting supply-side shifts using both pre- and post-merger industry data and thereby inferring the underlying degree of competitiveness. I do not have to rely on aggregate outside data or on relatively weak selection methods to determine the industry conduct. The estimation results suggest that markups in the industry are above those predicted under multi-product Nash pricing. I find that between 14.3 and 25.6 percent of the estimated markups can be attributed to cooperative industry behavior.

The proposed methods require sufficient variation in the price movement across different products or a variation in input prices. This might not be achieved in all horizontal merger cases. However, in the growing literature on ex-post merger evaluations there are already examples with seemingly sufficient variation, for example painkillers (Bjoernerstedt and Verboven 2012), motor oil and syrup (Weinberg and Hosken 2012), and cars (Yoshimoto 2011). Up until now, this literature mainly compares the predictions of different demand models under the assumption of multi-product Nash competition. Both methods can be applied to all of these mergers. Besides additional information about competition and merging firms' behavior within these industries, using such data would also give interesting information about behavior across industries. From an organizational perspective, this can yield insights about the effects of different managerial firm structures on post-merger behavior, and about differences in the potential to maximize joint profits. From an industry perspective, this can also provide information about the relationship between competition and market power across different industries.

2.A Rank conditions examples

In this section I present further examples that highlight the effects of the assumptions made above. The main question will be under which circumstances marginal costs and industry conduct will be jointly identified in a model.

3 firms, brands 1 and 2 belong to same firm Consider an industry that consists of 4 brands, where brands 1 and 2 belong to the same firm. For simplicity, assume in this example that marginal costs are constant for each firm. Furthermore, denote by p_i, mc_i, s_i the price, marginal costs and market share of firm i, respectively. θ_{ij} describes the degree to which brand i takes into account the profits of brand j when making its decision. In the example, the maximization problem of brand 1 thus yields

$$\max_{p_1}(p_1 - mc_1)s_1(p) + (p_2 - mc_2)s_2(p) + \theta_{13}(p_3 - mc_3)s_3(p) + \theta_{14}(p_4 - mc_4)s_4(p)$$

The first-order condition for brand 1 with respect to its price then yields

$$(p_1 - mc_1)\frac{\partial s_1}{\partial p_1} + s_1 + (p_2 - mc_2)\frac{\partial s_2}{\partial p_1} + \theta_{13}(p_3 - mc_3)\frac{\partial s_3}{\partial p_1} + \theta_{14}(p_4 - mc_4)\frac{\partial s_4}{\partial p_1} = 0$$

There is a change in the ownership matrix pre- and post-merger if firms 2 and 3 merge. When making the additional assumption that each firm maximizes the profits of all of its brands, and merging firms fully internalize their profits, the associated pre- and post-merger conduct matrices can be written as

$$\Theta = \begin{pmatrix} 1 & 1 & \theta_{13} & \theta_{14} \\ 1 & 1 & \theta_{23} & \theta_{24} \\ \theta_{31} & \theta_{32} & 1 & \theta_{34} \\ \theta_{41} & \theta_{42} & \theta_{43} & 1 \end{pmatrix}; \qquad \Theta^{post} = b(\Theta) = \begin{pmatrix} 1 & 1 & \theta_{13} & \theta_{14} \\ 1 & 1 & \theta_{23} & \theta_{24} \\ \theta_{31} & \theta_{32} & 1 & 1 \\ \theta_{41} & \theta_{42} & 1 & 1 \end{pmatrix}$$

From firm 1's first order condition, conditional on the form of industry conduct, firms will adapt their prices after an ownership change. In the above example, without symmetry, there are 10 parameters to estimate, with only 4 equations, such that the rank conditions are never met for identification. I introduce different assumption on firm supply to reduce the number of parameters to be estimated.

Bilateral symmetry between firms Instead of bilateral brand symmetry, a stricter assumption is that for all brands of two distinct firms, each brand will take the other firms' brands into account in the same fashion when making its pricing decision. Pre-merger and post-merger conduct can be written as

$$\Theta = \begin{pmatrix} 1 & 1 & \theta^a & \theta^b \\ 1 & 1 & \theta^a & \theta^b \\ \theta^a & \theta^a & 1 & \theta^c \\ \theta^b & \theta^b & \theta^c & 1 \end{pmatrix}; \qquad \Theta^{post} = b(\Theta) = \begin{pmatrix} 1 & 1 & \theta^a & \theta^b \\ 1 & 1 & \theta^a & \theta^b \\ \theta^a & \theta^a & 1 & 1 \\ \theta^b & \theta^b & 1 & 1 \end{pmatrix}$$

This leads to a number of 3 parameters to estimate, with 4 available equations, such that the system is identified in absence of multi-collinearity.

Symmetry among all cross-firm brands When assuming that all brands take the brands of all other firms into account in the same way, this results in the following pre- and post-merger conduct:

$$\Theta = \begin{pmatrix} 1 & 1 & \theta^a & \theta^a \\ 1 & 1 & \theta^a & \theta^a \\ \theta^a & \theta^a & 1 & \theta^a \\ \theta^a & \theta^a & \theta^a & 1 \end{pmatrix}; \qquad \Theta^{post} = b(\Theta) = \begin{pmatrix} 1 & 1 & \theta^a & \theta^a \\ 1 & 1 & \theta^a & \theta^a \\ \theta^a & \theta^a & 1 & 1 \\ \theta^a & \theta^a & 1 & 1 \end{pmatrix}$$

There is only one conduct parameter to estimate and in 4 equations.

Post-merger internalization of profits When estimating the internalization of post merger profits among merging firms, the form of industry competition. If one assumes multi-product Nash pricing in the 3 firm, 4 brand case, then pre- and post-merger industry conduct can be written as

$$\Theta = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \qquad \Theta^{post} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & \tilde{\theta} \\ 0 & 0 & \tilde{\theta} & 1 \end{pmatrix}$$

Here, $\tilde{\theta}$ represents the degree of profit-internalization between firms 2 and 3.

2.B Derivation of structural production error term using logarithmic cost function

Instead of the linear cost specification from equation (2.9), I now assume a logarithmic cost specification.

$$log(mc(\gamma^D, \Theta, p, x)) = w\gamma^S + \omega. \tag{2.29}$$

Combining the above equation with the recovered marginal cost estimates from equation (2.5), and solving for the pre-merger structural error term ω^{pre} yields

$$\omega_j^{pre}(\hat{\gamma}^D, \gamma^S, \Theta) = log\left(\frac{mc_j(\gamma^D, \gamma^S, \Theta, p, x, w)}{e^{\gamma^S w_j}}\right). \tag{2.30}$$

The post-merger structural error when estimating post-merger profit internalization can be derived as

$$\omega_{\tilde{\theta}^{post}}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}) = log\left(\frac{p^{post} - \Omega^{-1}(\gamma^D, \Theta, \tilde{\theta})s^{post}}{\hat{mc}^{post}(\gamma^D, \gamma^S, \Theta, p, x, w)}\right). \tag{2.31}$$

The structural error term can be obtained in an analogous fashion. The computed error terms can then again be combined with instruments discussed in section 3.

2.C Computational details

For the Random Coefficient Logit model, I use the derivative-based SOLVOPT algorithm. In practice, this algorithm has shown to provide more accurate estimation results in terms of a lower GMM objective function, see for example Knittel and Metaxoglou (2011). I also compute theoretical and numerical gradients at the equilibrium, which both are very close to zero, the highest gradient being of magnitude 0.003. In my estimation routine, I draw 50 individuals per store, and combine the draws with demographic store characteristics.

For my GMM supply-side routine, I use a basic finite-descent accelerated random search (ARS) algorithm, as proposed by Appel et al. (2004). For each estimation step, I use 1500 starting points.

2.D Rivers and Vuong approach

Formal Rivers and Vuong approach Consider a non-nested conduct specification h, with conduct matrix Θ^h at a time.[20] Note that this can include differences in the conduct matrix between pre- and post-merger. Under this hypothesis, one can recover the marginal cost using equation (2.5): $mc = p - \Omega^{-1}(\Theta^h)s$. When estimating a cost function, I decompose the structural unobserved error term into a fixed brand-specific unobserved component and a random shock component: $\omega^h_{jt} = \omega^h_j + \epsilon^h_{jt}$. Using my marginal cost specification 2.9 leads to

$$mc^h_{jt} = w'_{jt}\gamma^S + \omega^h_j + \epsilon^h_{jt}. \tag{2.32}$$

I assume that the random shock component uncorrelated to the observable input characteristics and to the unobserved brand-specific component:$E[\epsilon^h_{jt}|\omega^h_j, w_{jt}] = 0$.
The test looks for the cost equation with best statistical fit using the observable input characteristics from the cost function. Note that these characteristics are brand-specific, but do not change with the different conduct specifications. If one tests between two different models, h, and h', one obtains the following equations for the industry prices.

$$p^h_{jt} = \Omega^{-1}(\Theta^h)s + w'_{jt}\gamma^S_h + \omega^h_j + \epsilon^h_{jt}$$

$$p^{h'}_{jt} = \Omega^{-1}(\Theta^{h'})s + w'_{jt}\gamma^S_{h'} + \omega^{h'}_j + \epsilon^{h'}_{jt}$$

For a given specification h, denote $Q^h_n(\gamma^S, \omega)$ the lack-of-fit criterion given the parameters γ^S and ω. For any specification h, one now aims to find the parameter values that minimize the lack of fit criterion Q^h:

$$\min_{\gamma^S_h, \omega^h_j} Q^h_n(\gamma^S_h, \omega^h_j) = \min_{\gamma^S_h, \omega^h_j} \frac{1}{n} \sum_{j,t} (\epsilon^h_{jt})^2 = \frac{1}{n} \sum_{j,t} [p_{jt} - \Omega^{-1}(\Theta^h)s - \omega^h_j - w'_{jt}\gamma^S_h]^2$$

This does not require any of the specifications to be the correct model. Denote by $\overline{Q}^h_n(\overline{\gamma}^S_h, \overline{\omega}^h_j)$ the expected lack-of-fit criteria for specification h.
Given specifications h and h', there are three different hypothesis one has to test against each other, which have the following asymptotic properties.

[20] This Appendix uses similar structure and notation as Bonnet and Dubois (2010).

H_0 : h and $h^{'}$ are asymptotically equivalent if $lim_{n\to\infty}\{\overline{Q}_n^h(\overline{\gamma}_h^S, \overline{\omega}_j^h) - \overline{Q}_n^{h'}(\overline{\gamma}_{h'}^S, \overline{\omega}_j^{h'})\} = 0$.
H_1 : h asymptotically better fit than h' if $lim_{n\to\infty}\{\overline{Q}_n^h(\overline{\gamma}_h^S, \overline{\omega}_j^h) - \overline{Q}_n^{h'}(\overline{\gamma}_{h'}^S, \overline{\omega}_j^{h'})\} < 0$.
H_2 : h' asymptotically better fit than h if $lim_{n\to\infty}\{\overline{Q}_n^h(\overline{\gamma}_h^S, \overline{\omega}_j^h) - \overline{Q}_n^{h'}(\overline{\gamma}_{h'}^S, \overline{\omega}_j^{h'})\} > 0$.
Denote by T_n the test statistic that accounts for the variation in the lack-of-fit criteria for the different hypotheses.

$$T_n = \frac{\sqrt{n}}{\hat{\sigma}_n^{hh'}}\{Q_n^h(\hat{\gamma}_h^S, \hat{\omega}_j^h) - Q_n^{h'}(\hat{\gamma}_{h'}^S, \hat{\omega}_j^{h'})\}, \tag{2.33}$$

where $\sigma_n^{hh'}$ is the variance of the difference of the estimated lack-of-fit criteria. . Rivers and Vuong show that if two models are strictly non-nested, the asymptotic distribution of T_n is standard normal. Thus, one has to compare sample values of T_n with the critical values of a standard normal distribution.

2.E Proofs

Proof of Proposition 2.1

Proof. The demand parameters γ^D can be estimated from equations 10 and 13, respectively. Regarding the supply side, there are J estimable equations, one equation per brand post-merger. Because each firm has one conduct parameter for each competitor, this leads to an overall number of $N(N-1)$ parameters. The bilateral symmetry assumption reduces this number to $\frac{N(N-1)}{2}$. This leads to J equations with $\frac{N(N-1)}{2}$ parameters. The model is only identified if there are at least as many equations as parameters, i.e. if $\frac{N(N-1)}{2} \leq J$. □

Proof of Proposition 2.2

Proof. The demand parameters γ^D can be estimated from equations 10 and 13, respectively. Regarding the supply side, there are J estimable equations, one equation per brand pre-merger, and one equation per brand post-merger.Because each firm has one conduct parameter for all firms, this leads to an overall number of N conduct parameters. This leads to J equations with N parameters. The model is only identified if there are at least as many equations as parameters, i.e. if $N \leq J$. □

Proof of Proposition 2.3

Proof. Using the same reasoning as in the proof for Proposition 2.2, there are J equations and one parameter to estimate, so that the result trivially holds. □

2.F Graphs and tables

Adult enhanced	Adult simple	Family	Kids
PO Raisin Bran	NAB Orig. Shrd. Wheat	GM Wheaties	PO Honeycomb
GM Raisin Nut Bran	NAB Spoon Size Shrd.	GM Cheerios	GM Apple-Cinn. Cheerios
KE Raisin Bran	PO Grape Nuts Cereal	KE Corn Flakes	GM Honey Nut Cheerios
KE Nutri Grain	GM Total Corn Flakes	KE Crispix	GM Lucky Charms
QO 100% Natural	KE Special K	RA Corn Chex	GM Trix
	KE Just Right Fruit Nut	RA Wheat Chex	KE Froot Loops
		RA Rice Chex	KE Frosted Flakes
			KE Corn Pops
			KE Smacks
			QO Cap'n Crunch

Table 2.1: Product segmentation

Brand Name	% 92Q3-93Q1	% 92Q3-93Q3	% 92Q3-94Q1
NAB Orig Shred Wheat	0.03	0.12	0.11
NAB Sp.Size Shrd Wheat	0.01	0.03	0.09
PO Grape Nuts Cereal	0.04	0.01	0.04
PO Raisin Bran	0.07	0.04	0.23
PO Honeycomb	0.08	0.08	0.10
GM Raisin Nut Bran	-0.02	0.03	0.12
GM Apple-Cin Cheerios	-0.02	-0.10	-0.05
GM Wheaties	-0.01	-0.12	-0.17
GM Cheerios	-0.09	-0.14	-0.09
GM Honey Nut Cheerios	-0.02	-0.07	-0.17
GM Lucky Charms	-0.07	-0.11	-0.05
GM TOTAL Corn Flakes	0.00	0.06	-0.09
GM Trix	-0.22	-0.12	0.07
KE Froot Loops	0.00	-0.14	-0.12
KE Special K	0.04	0.02	-0.09
KE Frosted Flakes	0.01	0.03	-0.13
KE Corn Pops	0.05	-0.23	-0.28
KE Raisin Bran	0.06	-0.06	-0.04
KE Corn Flakes	-0.02	-0.03	-0.24
KE Smacks	0.06	-0.01	0.13
KE Crispix	0.04	0.08	0.15
KE Just Right FruitNut	-0.24	0.02	0.18
KE Nutri Grain	0.03	0.01	0.06
RA Corn Chex	0.03	0.06	0.05
RA Wheat Chex	0.03	0.05	0.05
RA Rice Chex	0.02	0.05	0.06
QO 100% Natural Cereal	0.03	0.05	-0.14
QO Cap'n Crunch	-0.27	0.04	-0.12

Table 2.2: Product specific price development

Note: Column 1 shows of quantity-weighted average percentage deflated price change between quarter 3, 1992, and quarter 1, 1993. Columns 2 and 3 show the price developments between 1992, quarter 2 and 1993, quarter 3, and 1994, quarter 2, respectively.

Conduct specification	Necessary condition
Bilateral symm btw. firms	$\frac{N(N-1)}{2} \leq J$
Same resp. to cross-firm br.	$N \leq J$
Same resp. btw. all firms	$J \geq 1$ (always met)
Menu approach	$J \geq 0$ (always met)

Table 2.3: Identification conditions for different specifications

Variable	Mean	Std. Dev.	Interaction Small Child	Interaction Income	Interaction Household Size	Interaction St.Dev. Income
price	-34.97	1.69	-387.90	-93.38	–	44.37
	(7.69)	(3.80)	(333.60)	(191.10)	–	(32.35)
const	-1.42	-.14	–	–	12.30	–
	(.55)	(.60)	–	–	(7.87)	–
mushy	.12	.01	6.33	1.36	–	–
	(.40)	(.23)	(19.14)	(12.14)	–	–
sugar	-.03					
	(.00)					
refined grains	.46					
	(.03)					
quarter trend	.02					
	(.00)					
adult segment	1.25					
	(.06)					
kids segment	.47					
	(.04)					

Table 2.4: Demand side estimates γ^S for Random Coefficient Logit model

Note: Num Obs: 19600. Interactions with demographics from US 1990 Census data around each store. Omitted category: Firm dummy variables to account for firm-specific brand valuations.

	(1)	(2)	(3)	(4)	(5)	(6)
Price	-11.5485***	-11.6427***	-7.3154***	-10.2634***	-13.8593***	-11.2802***
	(0.2287)	(0.4481)	(0.2736)	(0.494)	(0.2376)	(0.1254)
Constant	-2.2791***	-3.2689***	-2.3519***	-1.8492***	-3.1581***	-2.147***
	(0.0487)	(0.1515)	(0.0602)	(0.1051)	(0.0878)	(0.0358)
Sugar	-0.0344***	-0.051***	-0.0143***	-0.0189***	-0.0528***	-0.0414***
	(0.0013)	(0.0022)	(0.0014)	(0.0021)	(0.0013)	(0.0007)
Fat	-0.0369***	0.0346***	0.0155***	0.0153***	0.0184***	0.0073***
	(0.004)	(0.0067)	(0.003)	(0.0052)	(0.0038)	(0.0025)
Refined grains	0.4682***	0.3481***	0.2636***	0.2679***	0.4377***	0.4383***
	(0.0207)	(0.0349)	(0.0277)	(0.0395)	(0.0198)	(0.0134)
Sales	0.7569***	3.4081***	2.3924***	5.0058***	1.5109***	0.1584***
	(0.067)	(0.2007)	(0.0918)	(0.215)	(0.073)	(0.0102)
Quarter dummy	-0.0001***	-0.0492***	-0.0304***	-0.0647***		
	(0.0016)	(0.0041)	(0.0023)	(0.0049)		
Kids Cereal	1.0309***	1.698***	0.7072***	0.9768***	2.7253***	1.581***
	(0.054)	(0.0816)	(0.0528)	(0.0784)	(0.1024)	(0.0338)
Adult Cereal	0.3237***	0.6644***	0.4202***	0.6491***	1.6531***	0.2053***
	(0.0301)	(0.0488)	(0.0327)	(0.0502)	(0.1019)	(0.0385)
Input price IV	Yes	No	Yes	No	No	No
Zone IV	No	Yes	No	Yes	Yes	No
Ownership IV	Yes	Yes	Yes	Yes	Yes	No
Firm dummies	Yes	Yes	No	No	Yes	No

Table 2.5: Demand estimation results for different Logit specifications

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Num Obs: 19600.

Table 2.6: Mean Elasticities Random Coefficient Logit Model Part 1

	NAB ShW	NAB SSW	PO GNu	PO RBr	PO RNB	GM Hon	GM ACh	GM Whe	GM Che	GM HNC	GM LCh	GM TCF	GM Tri	KE FLo
NAB Shred Wheat	-3.835	0.038	0.043	0.056	0.071	0.022	0.037	0.036	0.113	0.122	0.094	0.055	0.056	0.123
NAB Sp Shr Whea	0.012	-3.366	0.045	0.057	0.069	0.022	0.039	0.035	0.117	0.125	0.090	0.049	0.054	0.115
PO Grape Nuts	0.010	0.035	-2.630	0.059	0.063	0.023	0.042	0.032	0.121	0.129	0.083	0.039	0.048	0.101
PO Raisin Bran	0.011	0.035	0.046	-2.910	0.063	0.023	0.042	0.032	0.123	0.130	0.084	0.040	0.049	0.102
PO Honeycomb	0.012	0.037	0.044	0.057	-3.713	0.022	0.039	0.035	0.116	0.125	0.092	0.051	0.055	0.118
GM RaisNut Bran	0.011	0.035	0.046	0.060	0.064	-4.232	0.044	0.033	0.129	0.137	0.087	0.041	0.051	0.105
GM AppCin Cheer	0.010	0.033	0.046	0.060	0.060	0.024	-2.899	0.031	0.126	0.133	0.080	0.036	0.046	0.095
GM Wheaties	0.012	0.037	0.044	0.057	0.070	0.023	0.039	-3.956	0.118	0.127	0.093	0.051	0.055	0.118
GM Cheerios	0.010	0.034	0.046	0.060	0.063	0.024	0.044	0.032	-3.495	0.134	0.084	0.039	0.049	0.102
GM HonNut Cheer	0.011	0.035	0.046	0.060	0.064	0.024	0.043	0.033	0.126	-3.613	0.086	0.040	0.050	0.104
GM Luck Charms	0.012	0.037	0.044	0.057	0.070	0.023	0.039	0.035	0.118	0.127	-4.167	0.052	0.056	0.119
GM TCorn Flakes	0.014	0.039	0.040	0.052	0.075	0.021	0.033	0.038	0.106	0.116	0.100	-4.580	0.061	0.134
GM Trix	0.012	0.038	0.044	0.056	0.071	0.023	0.038	0.036	0.117	0.126	0.095	0.053	-4.329	0.122
KE Froot Loops	0.013	0.038	0.043	0.056	0.072	0.022	0.037	0.036	0.114	0.124	0.096	0.056	0.057	-4.203
KE Special K	0.013	0.038	0.042	0.055	0.073	0.022	0.036	0.037	0.112	0.121	0.096	0.058	0.058	0.127
KE Frost Flakes	0.012	0.038	0.044	0.056	0.070	0.022	0.038	0.035	0.115	0.124	0.093	0.053	0.055	0.119
KE Corn Pops	0.013	0.038	0.043	0.055	0.072	0.021	0.036	0.036	0.112	0.121	0.095	0.057	0.057	0.125
KE Raisin Bran	0.011	0.036	0.046	0.059	0.066	0.023	0.041	0.033	0.121	0.129	0.087	0.044	0.051	0.108
KE Corn Flakes	0.011	0.035	0.046	0.059	0.063	0.023	0.042	0.032	0.120	0.128	0.083	0.040	0.048	0.101
KE Smacks	0.012	0.038	0.044	0.056	0.071	0.022	0.038	0.036	0.115	0.124	0.093	0.053	0.056	0.120
KE Crispix	0.013	0.039	0.041	0.053	0.074	0.021	0.034	0.037	0.107	0.117	0.098	0.062	0.059	0.131
KE JRight Fruit	0.008	0.028	0.045	0.057	0.049	0.022	0.045	0.025	0.119	0.123	0.063	0.023	0.036	0.071
KE Nutri Grain	0.011	0.036	0.045	0.059	0.067	0.025	0.042	0.034	0.127	0.136	0.091	0.046	0.054	0.113
RA Corn Chex	0.013	0.038	0.042	0.055	0.072	0.021	0.036	0.036	0.111	0.120	0.096	0.058	0.058	0.126
RA Wheat Chex	0.012	0.037	0.045	0.058	0.067	0.023	0.040	0.034	0.119	0.127	0.089	0.047	0.053	0.112
RA Rice Chex	0.013	0.039	0.042	0.054	0.073	0.021	0.035	0.037	0.109	0.119	0.097	0.060	0.059	0.129
QO 100 Natural	0.007	0.028	0.046	0.061	0.051	0.029	0.051	0.027	0.143	0.149	0.072	0.024	0.040	0.077
QO Capn Crunch	0.011	0.035	0.046	0.059	0.064	0.024	0.043	0.033	0.126	0.134	0.086	0.041	0.050	0.105

Table 2.7: Mean Elasticities Random Coefficient Logit Model Part 2

	KE SpK	KE FFl	KE CPo	KE RBr	KE CFl	KE Sma	KE Cri	KE JRi	KE NGr	RA CCh	RA WCh	RA RCh	QO QO1	QO CCr
NAB Shred Wheat	0.105	0.142	0.097	0.079	0.066	0.014	0.077	0.000	0.005	0.028	0.016	0.038	0.023	0.065
NAB Sp Shr Whea	0.097	0.136	0.090	0.079	0.068	0.013	0.069	0.000	0.005	0.026	0.016	0.035	0.026	0.067
PO Grape Nuts	0.083	0.123	0.078	0.078	0.070	0.012	0.057	0.001	0.005	0.022	0.016	0.029	0.034	0.068
PO Raisin Bran	0.084	0.123	0.079	0.078	0.070	0.012	0.057	0.001	0.005	0.022	0.016	0.029	0.035	0.069
PO Honeycomb	0.100	0.138	0.093	0.079	0.067	0.014	0.072	0.000	0.005	0.026	0.016	0.036	0.026	0.067
GM RaisNut Bran	0.085	0.124	0.079	0.080	0.069	0.012	0.058	0.001	0.005	0.022	0.016	0.030	0.042	0.073
GM AppCin Cheer	0.077	0.117	0.073	0.078	0.070	0.011	0.052	0.001	0.005	0.020	0.015	0.027	0.041	0.070
GM Wheaties	0.099	0.138	0.092	0.079	0.067	0.014	0.072	0.000	0.005	0.026	0.016	0.036	0.027	0.067
GM Cheerios	0.083	0.122	0.078	0.079	0.070	0.012	0.056	0.001	0.005	0.022	0.016	0.029	0.040	0.071
GM HonNut Cheer	0.085	0.124	0.079	0.079	0.069	0.012	0.058	0.001	0.005	0.022	0.016	0.030	0.039	0.071
GM Luck Charms	0.100	0.138	0.093	0.080	0.067	0.014	0.072	0.000	0.005	0.026	0.016	0.036	0.028	0.068
GM TCorn Flakes	0.117	0.151	0.107	0.077	0.062	0.015	0.088	0.000	0.005	0.031	0.017	0.043	0.018	0.062
GM Trix	0.103	0.140	0.095	0.080	0.066	0.014	0.074	0.000	0.005	0.027	0.016	0.037	0.027	0.067
KE Froot Loops	0.106	0.143	0.098	0.079	0.065	0.014	0.077	0.000	0.005	0.028	0.016	0.038	0.024	0.066
KE Special K	-4.148	0.145	0.100	0.079	0.065	0.014	0.080	0.000	0.005	0.029	0.017	0.039	0.022	0.065
KE Frost Flakes	0.101	-3.519	0.094	0.079	0.067	0.014	0.073	0.000	0.005	0.027	0.016	0.036	0.024	0.066
KE Corn Pops	0.107	0.144	-3.871	0.079	0.065	0.014	0.079	0.000	0.005	0.028	0.016	0.039	0.022	0.065
KE Raisin Bran	0.090	0.129	0.084	-3.314	0.069	0.013	0.063	0.000	0.005	0.024	0.016	0.032	0.032	0.068
KE Corn Flakes	0.083	0.123	0.079	0.078	-2.401	0.012	0.057	0.001	0.005	0.022	0.016	0.029	0.032	0.067
KE Smacks	0.102	0.140	0.095	0.079	0.067	-3.904	0.074	0.000	0.005	0.027	0.016	0.037	0.025	0.066
KE Crispix	0.113	0.149	0.105	0.078	0.063	0.015	-4.208	0.000	0.005	0.030	0.017	0.042	0.019	0.063
KE JRight Fruit	0.056	0.093	0.054	0.070	0.069	0.009	0.035	-0.252	0.004	0.015	0.013	0.019	0.047	0.064
KE Nutri Grain	0.093	0.131	0.086	0.081	0.068	0.013	0.065	0.000	-4.748	0.024	0.016	0.033	0.038	0.072
RA Corn Chex	0.108	0.145	0.100	0.078	0.065	0.014	0.080	0.000	0.005	-3.992	0.017	0.039	0.021	0.065
RA Wheat Chex	0.093	0.133	0.087	0.079	0.069	0.013	0.066	0.000	0.005	0.025	-3.399	0.033	0.029	0.068
RA Rice Chex	0.111	0.147	0.102	0.078	0.064	0.015	0.083	0.000	0.005	0.029	0.017	-4.123	0.020	0.064
QO 100 Natural	0.058	0.093	0.054	0.075	0.067	0.009	0.035	0.001	0.005	0.015	0.014	0.019	-3.968	0.077
QO Capn Crunch	0.086	0.125	0.080	0.079	0.069	0.012	0.059	0.001	0.005	0.023	0.016	0.030	0.038	-3.761

Industry competition	Parameter Value $\Theta = 0$	Parameter Value $\Theta = 0.25$	Parameter Value $\Theta = 0.5$	Parameter Value $\Theta = 1$
Wheat	0.123**	0.112***	0.091**	0.046**
Rice	0.057**	0.058***	0.061***	0.066***
Oat	0.019***	0.017***	0.014***	0.007***
Corn	0.092***	0.085***	0.077***	0.057***
Sugar	0.001***	0.001***	0.000***	0.000***
Vitamin	0.028***	0.027***	0.024***	0.02***
Firm Dummies	Yes	Yes	Yes	Yes
Median mc^{pre}	0.1036	0.0961	0.0864	0.0606
Mean mc^{pre}	0.1005	0.0930	0.0837	0.0586
Std. mc^{pre}	0.0669	0.0616	0.0564	0.0481
Median mc^{post}	0.1138	0.1062	0.0971	0.0721
Mean mc^{post}	0.1161	0.1087	0.0994	0.074
Std. mc^{post}	0.0807	0.0709	0.0608	0.0489

Table 2.8: Cost function estimates γ^S

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Num Obs: 7840. Columns 2-5 show estimates for different forms of symmetric industry conduct, raging from full competition between firms to full collusion. Wheat, Oat, Corn, Sugar, Rice, Vitamins reflect coefficients for interaction between input prices and relative ingredient content in product.

Industry competition	$\Theta = 0$ (Nash) Estimated $\tilde{\theta}$	$\Theta = .25$ Estimated $\tilde{\theta}$	$\Theta = .5$ Estimated $\tilde{\theta}$
1-2 quarters	.30 (1.03)	.11 (.69)	.02 (.51)
3-4 quarters	.36 (1.05)	.19 (.71)	.02 (.49)
5-6 quarters	.43 (1.41)	.22 (.87)	.07 (.94)
7-8 quarters	.02 (.92)	.02 (.79)	.02 (.65)
9-10 quarters	.98 (80.2)	.98 (103.7)	.98 (95.5)

Table 2.9: Joint profit maximization estimates $\tilde{\theta}$

Note: Standard errors in parentheses. Num Obs: 9800. Degree of merging firms' joint profit maximization over time, $\tilde{\theta}$, for different degrees of symmetric industry competition, Θ. Numerical minimization is restricted to values of $\tilde{\theta} \in [.018, .982]$.

	Conduct Parameter	Std. Dev
Inter-Firm Conduct	.708***	.023
Type of competition	Median PCM	St. Dev PCM
Estimated Conduct	.548	.216
Multi-brand Nash	.408	.285
Full collusion	.631	.217

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Num Obs: 7840. Recovered parameters from multi-product Nash pricing reflects the case of $\theta = 0$ for all cross-firm conduct parameters. Recovered parameters from full collusion reflects the case of of $\theta = 1$ for all cross-firm conduct parameters.

Table 2.10: Estimation of single conduct parameter

	Conduct Parameter	Std. Dev.
General Mills	.981***	.318
Ralston	.070	.206
Kellogg	.402***	.013
Post	.102	.106
Nabisco	.314***	.025
Quaker Oats	.195***	.070
Type of competition	Med. PCM	Std.PCM
Estimated Conduct	.494	.244
Multi-brand Nash	.408	.285
Full collusion	.631	.217

Table 2.11: Conduct estimates under symmetry to all firms

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Num Obs: 11760. Recovered parameters from multi-product Nash pricing reflects the case of $\theta = 0$ for all cross-firm conduct parameters. Recovered parameters from full collusion reflects the case of of $\theta = 1$ for all cross-firm conduct parameters.

	GM	RA	KE	PO	NA	QO
General Mills	1	.886	.046	.555	.799	.982
Ralston		1	.932	.856	.070	.982
Kellogg			1	.841	.086	.022
Post				1	.456	.977
Nabisco					1	.018
Quaker Oats						1
Type of competition	Med. PCM	Std.PCM				
Estimated Conduct	.476	.193				
Multi-brand Nash	.408	.285				
Full collusion	.631	.217				

Table 2.12: Conduct estimates under bilateral firm symmetry

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Num Obs: 11760. Recovered parameters from multi-product Nash pricing reflects the case of $\theta = 0$ for all cross-firm conduct parameters. Recovered parameters from full collusion reflects the case of of $\theta = 1$ for all cross-firm conduct parameters.

Cross-firm Conduct	$\Theta_{ij} = 0.25$	$\Theta_{ij} = 0.5$	$\Theta_{ij} = .75$	$\Theta_{ij} = 1$
$\Theta_{ij} = 0$	14.10	12.64	11.17	9.75
$\Theta_{ij} = .25$		11.35	9.84	8.35
$\Theta_{ij} = .5$			8.26	6.67
$\Theta_{ij} = .75$				4.90

Table 2.13: Selection method results using Rivers and Vuong test

Note: Test statistic follows a standard normal distribution. Num obs: 11760.

Figure 2.1: Pre-merger market shares

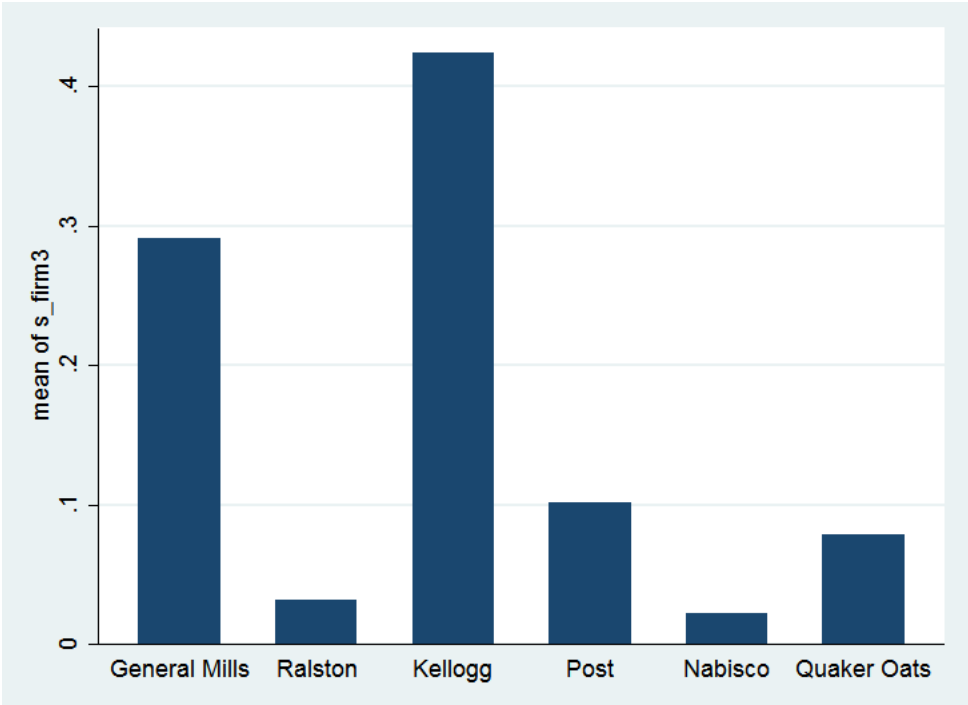


Figure 2.2: Geographical location of stores in dataset

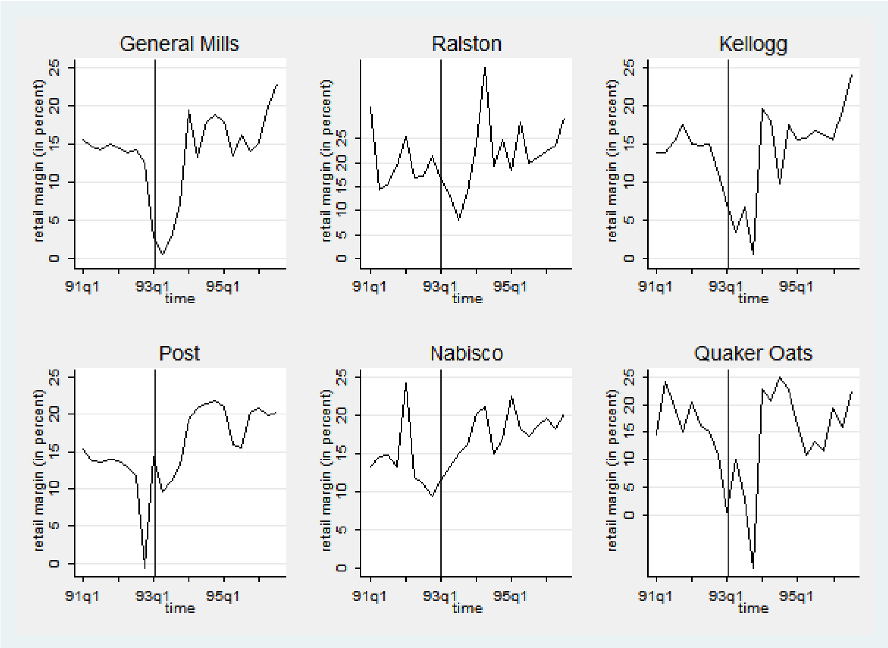
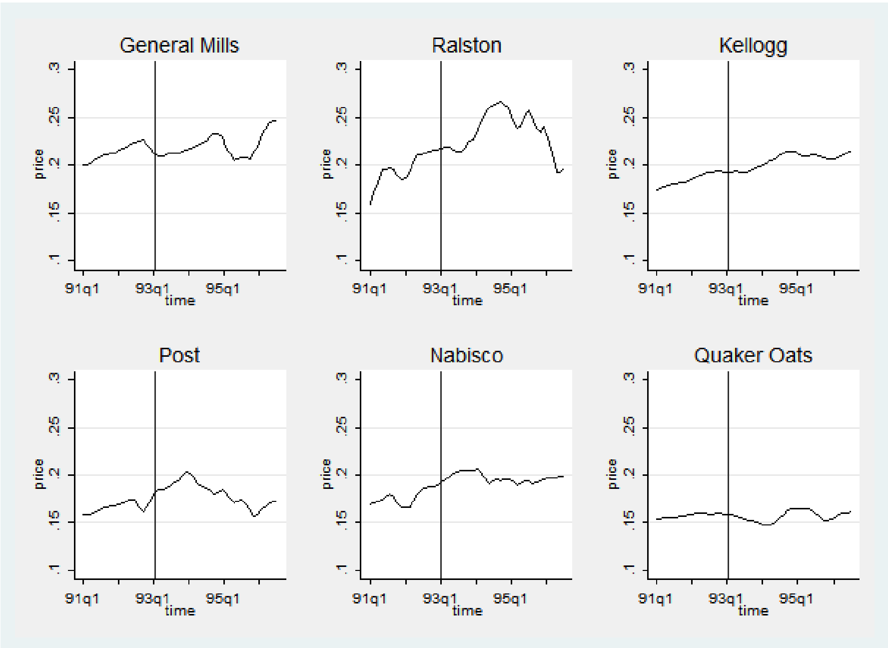


Figure 2.3: Retail margin development per firm
Note: Retail margin proxy computed as $1 - AAC$, where AAC are the average acquisition costs observed in the data for a firm per quarter. AAC reflects the percentage of the retail price paid to a producer.



Note: Firm prices computed as deflated quantity-weighted average prices over all products sold of same firm per quarter.

Figure 2.4: Average price development per firm

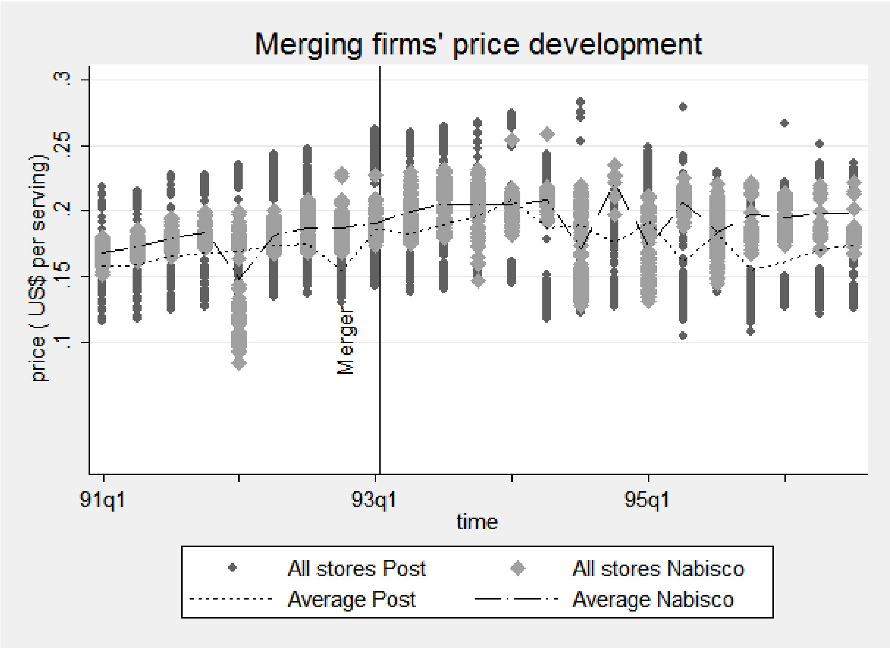


Figure 2.5: Price development of merging firms across stores
Note: Squares indicate quantity-weighted average deflated firm prices per store per quarter. Lines indicate average firm prices across all stores.

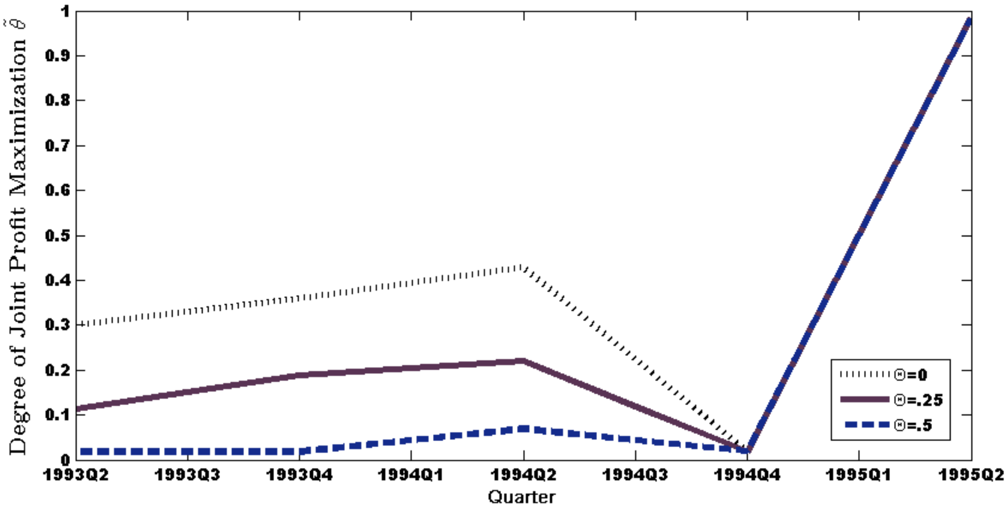


Figure 2.6: Degree of joint profit maximization $\tilde{\theta}$ over time
Note: Lines reflect degree of merging firms' joint profit maximization over time, $\tilde{\theta}$, for different degrees of symmetric industry competition, Θ. Degree of joint profit maximization is estimated separately for two-month intervals, and then harmonized between neighboring intervals.

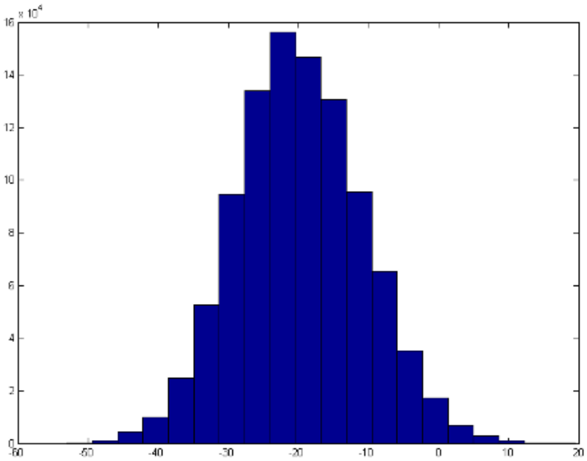


Figure 2.7: Distribution of random price coefficient $\tilde{\alpha}$

Chapter 3

Contractual Structures and Consumer Misperceptions – Warranties as an Exploitation Device

This chapter analyzes how contractual non-price features can be used to increase firm profits in the presence of biased consumers. We develop a model in which some consumers underestimate their costs of claiming a warranty payment. When a monopolist can signal his product quality by offering warranties, this gives rise to two different forms of consumer exploitation. First, he can offer contracts such that naive consumers overestimate the product quality and pay more than their willingness to pay for a good. Second, he can offer extended warranties for an excessive price, which also leads to overpaying consumers. Our results show that industry competition among firms is not able to prevent naive consumers from making non-optimal decisions. We introduce consumer protection policies in terms of minimum warranty standards. Such policies can sometimes mitigate consumer misperceptions, but also have the risk of distorting the market negatively if consumer utilities are sufficiently differentiated. The results match several empirical facts that are hard to explain with fully-rational consumers alone, namely the existence of consumer quality misperceptions, the high profitability of extended warranties, and the profitability

of mail-in rebates.[1][2]

[1]I would like to thank Erik Eyster, Michael Grubb, Volker Nocke, Matthew Rabin, Andre Stenzel, Elu von Thadden, as well as participants at the CDSE Seminar Mannheim, EARIE 2011 Stockholm, and at the ENTER Jamboree 2011 Tilburg for helpful comments. A previous version circulated under the title "When Consumer Naiveté affects Product Qualities – Warranties as an Exploitation Device."

[2]This chapter extends my own work of Michel (2009). Several sections, especially the main monopoly section, simultaneous oligopoly section, and the introduction largely include similar or identical parts compared to my previous work. There are, however, several key contributions that are new to this thesis. First, the previous work neither focused on nor explicitly mentioned extended warranties, which are a key policy component of the chapter. Second, the sequential oligopoly section covering Propositions 3.3 and 3.4 is not included either, as have all graphs in this chapter. Third, the analysis of consumer protection policies and all other extensions in section 3.5 are new.

3.1 Introduction

Recent work in economics and psychology has focused on how psychological biases affect contract design, with a particular focus on the pricing of goods. In many instances contractual complexity and other non-price features can also be reasons for consumer confusion. A warranty applicable in case of a product breakdown is such a feature that is included in many contracts. Next to the insurance effect, warranties often serve as a signal for product quality.[3]

Warranty practices have recently come under scrutiny from policy makers. The British Office of Fair Trading (OFT) concluded in its 2002 market investigation that "The similar behaviour of electrical retailers limits consumers' ability to make accurate assessments of the value of buying extended warranties" and that "Consumer protection in this market is inadequate" (OFT 2002). Subsequently, British legislature added protection policies such as mandatory information to consumers that extended warranties are optional, as well as a 45 day cancellation period. In 2011, the OFT concluded that these measures were not reducing consumer confusion, and that that common practices where "unfair and uncompetitive." Further evidence suggests that extended warranties are among the main profit drivers for retailers in the consumer electronics industry.[4] The market for redeemable rebates shares some patterns with the warranty market. There is evidence that firms using redeemable rebates as a promotional tool also because many consumers will not exercise them and end up paying the full price, see for example Jolson et al. (1987).[5]

This chapter formalizes a model in which some consumers overestimate their probability of claiming a warranty payment in case of a product failure. The model can match the previously discussed stylized facts, i.e. the ex-ante excessive purchase of extended

[3] Erdem et al. (2010) and Choi and Ishii (2010) empirically analyze the main reasons consumers purchase warranties. Both find a dominating effect of consumer quality signaling over other reasons, i.e. risk aversion and price-prediction, respectively.

[4] See also "The warranty windfall" in Business Week, December 20, 2004. An extended warranty can also be seen as an add-on good. Since the price of the warranty is observable, this is however not be consistent with a "shrouded attribute", as for example in Gabaix and Laibson (2006).

[5] Dhar and Hoch (1996) conduct field experiments in order to compare how consumers react to both redeemable coupons and off-the-shelf price discounts. They find that coupons lead to both a higher increase in the number of sales and higher profits compared to the price discounts, for only an average of 55% of the consumers redeem the coupons.

warranties, cases of consumer quality misperceptions, and the profitability of mail-in-rebates for firms by introducing a single consumer bias. Moreover, it introduces a novel form of consumer exploitation, namely an overestimation of product quality through false inference via the warranty. We focus on optimal contract setting and consumer choices both under monopoly and oligopoly, and analyze the effectiveness of consumer protection policies in terms of minimum warranty standards.

The model's intuition is the following: Many goods have a certain probability of having a defect. This probability can be reduced by putting more quality into the production process, which is costly for the producing firm. Firms can issue warranty rights to consumers in terms of reimbursements or repair should a defect occur. When consumers consider buying a good whose quality they cannot observe, they make inferences about the quality that has been put into the good based on the predicted firm's profits for each possible product quality. Thus, they take into account a firm's trade-off between lower effort costs and higher costs of warranty payments. If all consumers are fully rational, firms can use high warranties to signal high product quality to consumers and therefore overcome a moral hazard problem.

However, we argue that some people underestimate the costs of returning the good in order to request a warranty payment. Such costs can be for example additional unforeseen shipment costs and opportunity costs of time spent sending the good back to the firm.[6] Because the firm's reimbursements costs in case of a low product quality is lower than expected by some consumers, a warranty grant may mislead such consumers into thinking that a good has a higher quality than is actually the case. A wrong inference about the quality of such a good through a warranty leads to a higher willingness to pay for the good ex-ante, which can be exploited by a monopolist. A second way for a monopolist to extract excessive rents from naive consumers is to bundle a high quality product together with an extended warranty. Since naive consumers make correct predictions about the quality of a product in such a case, the monopolist benefits solely from an overprediction of the extended warranty valuation. If the share of naive consumers is sufficiently high, both cases then lead to a market

[6] Further examples are the foregone utility of consuming features of a good that are still partly working, or costs due to losing the warranty certificate of a good which is necessary to claim such a warranty.

distortion that also persists in oligopoly.

The consumer bias is modeled as a systematic underestimation of the return costs consumers incur when claiming a warranty. There is empirical, experimental, and anecdotal evidence that such a bias exists in many markets. Evidence from both psychology and marketing suggesting that a large fraction of consumers does not complain to producers after having experienced a product failure. The TARP (1996) study for example shows that conditional on having a service failure, more than 70% of the customers do not report it. Chebat et al. (2005) argue that passive emotions such as resignation and avoidance are a strong factor for this kind of behavior. Huppertz (2007) finds a positive relationship between the leniency of the refund system and the number of consumer complaints. This suggest that firms can influence consumer response behavior via the complexity of their refund system.

Besides proposing a new concept of exploitation in this model, we are also interested in the resulting implications on the introduction of consumer policies. One problem consumer protection policies face is how to deal with supposedly naive or uninformed consumers, and whether market competition alone can save naive consumers from being exploited by firms. We introduce a policy in terms of a retailer's minimum warranty standards. The results show that these policies can overcome the problem of consumer quality misperceptions. In case of extended warranties, however, these policies have no effect on consumer welfare. When allowing for differentiated preferences with respect to product qualities, such protection policies can even hurt consumer welfare. From a consumer policy perspective, the literature reviews of Armstrong (2008) and especially Vickers (2004) give a useful summary of the current literature. Vickers also explores whether reputation can overcome a firm's commitment problem of providing a high quality to consumers. As a necessary condition, he finds that firms in such an equilibrium have to make positive profits, however full efficiency in terms of the optimal price-quality bundle can still not be achieved.

Our model relates to two different branches in economics, the literature on warranties in industrial organization and contract theory, as well as the exploitative contracting literature in the field of behavioral economics. Spence (1977) first formalizes a signaling role of warranties. In his model, the firm side is perfectly competitive, and consumers vary in their degree of risk-aversion. The marginal costs of consumption are increasing and convex in the probability of the good working without a defect, which he shows to be sufficient for warranties to be a potential signal for product quality in terms of a good's reliability. Emons (1988) focuses on the double moral hazard problem that arises from warranties, i.e. the trade-off between a firm's moral hazard of producing a low quality good when only offering a low warranty and a consumer moral hazard that arises because consumers do not use a good carefully when having a high warranty. Mann and Wissink (1990) assess the effectiveness of money-back and replacement warranties both when the product quality is observable and when it is not. They find that a money-back warranty is always better except for an intermediate range of replacement costs of the firm. Lutz and Padmanabhan (1998) develop a model in which independent non-manufacturing firms can enter the market to sell extended warranties in the presence of a manufacturing monopoly. They find an ambiguous effect on the profits of the manufacturer.

There is a growing literature of Industrial Organization models with a Behavioral Economics foundation, see Ellison (2006) and DellaVigna (2009) for broad literature reviews. DellaVigna and Malmendier (2004) focus on how firms can design contracts in order to maximize their revenue when facing consumers with hyperbolic preferences. Gabaix and Laibson (2006) give a bounded rationality explanation for why firms shroud prices of add-on costs in equilibrium. They develop a competitive model in which the purchase of a base good implicates add-on costs later on. Sophisticated consumers foresee the add-on costs even without advertising and substitute away from them if they exceed the costs of a private substitution; in contrast, myopes are unaware of these costs if information on add-on prices is shrouded. Armstrong and Chen (2009) develop a model in which some consumers do not pay attention to the quality component of products when making their purchasing decision under firm competition. In a symmetric mixed strategy equilibrium, the existence of inattentative consumers is responsible for postive firm profits. Drago and Kadar (2006) explain empirical evidence for people not sending back mail-in rebates offered in combination with the purchase of certain goods. They provide a model in which consumers have both hyperbolic preferences and suffer from a so-called "sunk-cost" effect. They show that a relatively short rebate period increases the number of people who complete the rebate, while a longer duration of the rebate period increases the number of consumers who purchase a good but decreases the number of consumers who complete the rebate. Inderst and Ottaviani (2013) explore the provision of consumer cancellation rights in case sellers act as advisors with respect to the suitability of a product for a consumer. They distinguish between a case in which all consumer's anticipate the seller's profit-maximizing intentions and a case in which all consumers always believe the seller's claims.

The remainder of this chapter is structured as follows. Section 3.2 presents the setup of our model. The analysis of the monopoly and oligopoly settings is provided in Section 3.3 and 3.4, respectively. Section 3.5 develops extensions of the baseline model such as rebates and consumer protection policies. Section 3.6 concludes with a discussion of the results.

3.2 Monopoly Setting

3.2.1 Baseline model

A risk-neutral monopolist offers a good to a unit mass of potential consumers. The monopolist can choose whether to produce a product of low quality q_L or high quality q_H, with $q_H > q_L$; $q_H, q_L \in [0, 1]$. The quality of the good reflects its probability of working properly, so that this probability is increasing in the quality of the good. A good's quality cannot be observed by the consumers prior to its purchase. Marginal cost c_i, $i \in \{L, H\}$ in production is constant for both quality types, with $c_H > c_L$, and there exists no fixed cost of production. The monopolist can issue warranty rights in terms of a payment to the consumers for the event of a breakdown of the good. In order to receive the warranty payment, the consumer has to send the good back to the firm, which incurs a cost r to the consumer. A cost draw r is unknown for each consumer prior to the purchase of the good. It is distributed according to the cumulative distribution function $F(r)$, with the according probability density function $f(r)$. We assume that consumers cannot affect the probability of a product failure, i.e. abstract from consumer moral hazard. In the following we will assume that there are two kinds of consumers: A fraction of $1 - \theta$ consumers is "sophisticated" in the sense that they correctly predict their distribution of return costs. The other θ consumers are "naive" in a sense that they erroneously underestimate their costs of returning the good, i.e

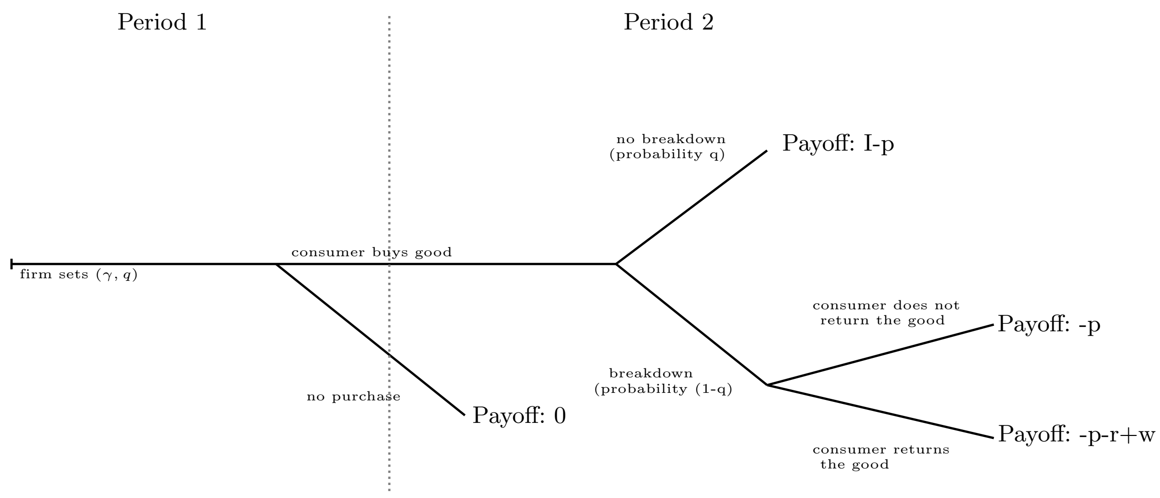


Figure 3.1: Timeline of the game

their anticipated distribution of return costs. Despite having the same return cost cdf $F(r)$, as the sophisticated consumers, naive consumers think that their distribution of return costs r can be described by a density function $\tilde{f}(r)$ and the corresponding cumulative distribution function $\tilde{F}(r)$. F first-order stochastically dominates $\tilde{F}(r)$, i.e. $\tilde{F}(r) \geq F(r) \forall r \in (0, \infty)$, with $\tilde{F}(r) > F(r)$ for the full support of F. This implies that naive consumers underestimate their return costs for all cost levels.
Furthermore, it is necessary to define the consumers' beliefs about the other players' warranty return costs. Naive consumers believe that all consumers have the return cost distribution $\tilde{F}(r)$ and that the monopolist has the same beliefs. Sophisticated consumers correctly believe that all consumers have the return cdf $F(r)$, but also correctly foresee that naive consumers underestimate the return cost distribution of all consumers, and that the monopolist shares the sophisticated consumers' beliefs. Thus, only sophisticated consumers and the monopolist are aware of a disagreement with respect to the return cost distribution.
Our equilibrium results are also consistent with different belief structures, which will be discussed later.

Structure of the game Figure 3.1 shows the timing of the game. In period 1, the monopolist proposes a contract $\gamma = (p, w)$ consisting of a price p and a warranty payment w to the consumer in case of a return of a defective product, and sets a quality $q \in \{q_L, q_H\}$. Each consumer values a functioning good with utility $I > 0$, whereas the utility from a non-functioning good is normalized to 0. Therefore, in case that a consumer buys the good from the monopolist, she derives a net utility of $I-p$ if the product works properly. In case of a defect of the good after the purchase, she has two options: If she sends the good back to the producer, she receives a compensatory warranty payment and hence experiences a utility of $-p+w-r$. In the case that she does not return the good to the store she experiences a utility $-p$. Thus, she will only return the good if the warranty payment exceeds the return costs, $w \geq r$. Therefore, the expected ex-ante utility in case of a defect is $-p+\int_0^w (w-r)f(r)dr$ for a sophisticated consumer, and $-p+\int_0^w (w-r)\tilde{f}(r)dr$ for a naive consumer. The outside option of not buying the good is set to 0.

In period 2, firms have to pay warranty payments to those consumers that claim the warranty after their good has had a defect.

The discount rate between periods is equal to 1. Given a quality q, the expected utility of consuming a good for the sophisticated consumer can be written as

$$V(q,\gamma) = qI - p + (1-q)\int_0^w [w-r]f(r)dr.$$

With $\frac{1}{\tilde{F}(w)}\int_0^w [w-r]f(r)dr \equiv E[w-r|w>r]$, this can be rewritten as

$$V(q,\gamma) = qI - p + (1-q)F(w)E[w-r|w>r]. \tag{3.1}$$

The expected utility of a naive consumer can be written as

$$\tilde{V}(q,\gamma) = qI - p + (1-q)\int_0^w [w-r]\tilde{f}(r)dr.$$

Define $\frac{1}{\tilde{F}(w)}\int_0^w [w-r]\tilde{f}(r)dr \equiv \tilde{E}[w-r|w>r]$. This then yields

$$\tilde{V}(q,\gamma) = qI - p + (1-q)\tilde{F}(w)\tilde{E}[w-r|w>r]. \tag{3.2}$$

We make the assumption that production and consumption of a high quality good leads to a higher welfare than the production of a low quality good and that the difference in terms of welfare is also high enough toovercome potential cost inefficiencies caused by signaling the high quality through a warranty.[7]

[7] Assumption 3.1 can be rewritten as $(q_H-q_L)I-c_H-c_L > \frac{c_H-c_L}{q_H-q_L}$. This gives a minimum required difference between the willingness to pay differences and cost differences of high and low quality goods, respectively.

Assumption 3.1. *(Consumer preference for high quality)*
$\frac{(q_H - q_L)^2 I}{1 + q_H - q_L} > c_H - c_L$.

We will now explain our equilibrium concept.

3.2.2 Equilibrium concept

To solve this game of imperfect information, we use a modified version of the Weak Perfect Bayesian Equilibrium concept with an additional belief refinement. Consumers have beliefs for each possible contract γ about the probability of the product being of high quality. Consumers then buy one unit of the good when it gives them a non-negative expected utility. Same as under the regular Weak Perfect Bayesian Equilibrium concept, we require that sophisticated consumers' beliefs match the true outcome in equilibrium. Unlike in Weak Perfect Bayesian Equilibrium, we require the beliefs of naive consumers to coincide with the hypothetical outcome in which the naive consumers' predicted distribution of return costs was the true distribution, i.e. in which the naive consumers have true expectations about their own distribution of costs of returning the good. This deviation from rationality is a crucial step in our model for establishing false beliefs about a product's quality. As a belief refinement, we impose the condition that all consumers believe that the producers will select the product quality that yields the highest profits given the contract γ.[8] As a convention, if producers are indifferent between providing a low or a high quality product for a given contract γ, we assume that consumers believe that the firm will produce a good of high quality. Given the structure of our model, it turns out that sophisticated consumers always have correct beliefs about the whole game after observing the monopolist's contract γ. Therefore, all non-rational behavior in the market is caused by naive consumers. We will now formalize the belief structure of the game.

3.2.3 Belief structure

Denote by $\mu(\gamma)$ a sophisticated consumer's belief about the probability of the monopolist's product being of high quality after having observed the contract γ offered by the monopolist. Analogously, denote by $\tilde{\mu}(\gamma)$ a naive consumer's belief about the probability of the monopolist's product being of high quality after having observed γ.

[8] This refinement is similar to that in Emons (1988).

Denote a sophisticated consumer's predicted consumption utility after having observed the contract γ by $U(\mu(\gamma),\gamma)$. This can then be written as

$$U(\mu(\gamma),\gamma) = \mu(\gamma)V(q_H,\gamma) + (1-\mu(\gamma))V(q_L,\gamma). \tag{3.3}$$

The analogous predicted consumption utility for a naive consumer after having observed γ, $\tilde{U}(\tilde{\mu}(\gamma))$, can be written as

$$\tilde{U}(\tilde{\mu}(\gamma),\gamma) = \tilde{\mu}(\gamma)\tilde{V}(q_H,\gamma) + (1-\tilde{\mu}(\gamma))\tilde{V}(q_L,\gamma). \tag{3.4}$$

A consumer will only buy a good if her predicted consumption utility is non-negative. Thus, the sophisticated consumers' demand for the good, $D(\gamma)$, can be written as

$$D(\mu(\gamma),\gamma) = \begin{cases} 1-\theta, & \text{if } U(\mu(\gamma),\gamma) \geq 0 \\ 0, & \text{if } U(\mu(\gamma),\gamma) < 0. \end{cases}$$

The naive consumers' demand can be written as

$$\tilde{D}(\tilde{\mu}(\gamma),\gamma) = \begin{cases} \theta, & \text{if } \tilde{U}(\tilde{\mu}(\gamma),\gamma) \geq 0 \\ 0, & \text{if } \tilde{U}(\tilde{\mu}(\gamma),\gamma) < 0. \end{cases}$$

The monopolist's profit maximization problem can then be written as

$$\max_{q\in\{q_L,q_H\},\gamma}[D(\mu(\gamma),\gamma) + \tilde{D}(\tilde{\mu}(\gamma),\gamma)](p-(1-q)F(w)w - c(q)), \tag{3.5}$$

where $c(q_L) = c_L; c(q_H) = c_H$.

We now characterize the threshold warranty level at which the monopolist is indifferent between producing a good with low or with high quality. If the monopolist wants to attract naive and sophisticated consumers, it is profitable for him to produce a high quality good given a warranty w if $p-(1-q_H)F(w)w - c_H \geq p-(1-q_L)F(w)w - c_L$ This is equivalent to

$$F(w)w \geq \frac{c_H - c_L}{q_H - q_L}. \tag{3.6}$$

Denote the minimum level of w that satisfies the weak inequality by w^S. Recall that sophisticated consumers believe both that naive consumers have the return cost cdf $F(s)$ and that the monopolist shares the same beliefs about naive consumers. Therefore, sophisticated consumers will only believe that a good is of high quality if

$w \geq w^S$. We assume that sophisticated consumers' beliefs are such that they reflect the dominant actions of the monopolist given the contract γ. Thus, the beliefs $\mu(\gamma)$ of the sophisticated consumers have the following form:

$$\mu(\gamma) = \begin{cases} 1, & \text{if } w \geq w^S \\ 0, & \text{if } w < w^S. \end{cases} \tag{3.7}$$

This implies our belief convention that if the monopolist is indifferent between producing a high or a low quality good, he will produce the high quality good. Because naive consumers believe that their costs are distributed according to the cdf $\tilde{F}(s)$, the constraint for a naive consumer if she believes that only naive consumers buy the good given a contract γ can turn to

$$\tilde{F}(w)w \geq \frac{c_H - c_L}{q_H - q_L}. \tag{3.8}$$

Denote the minimum level of w that satisfies this weak inequality by w^N. Therefore, given a contract γ from the monopolist, a naive consumer's belief $\tilde{\mu}$ will be of the following form:

$$\tilde{\mu}(\gamma) = \begin{cases} 1, & \text{if } w \geq w^N \\ 0, & \text{if } w < w^N. \end{cases} \tag{3.9}$$

Therefore, we again use the convention that if a monopolist is indifferent between producing with high or with low quality, the naive consumers believe that the monopolist is producing with high quality.

3.3 Monopoly Analysis

3.3.1 Derivation of the optimal contracts

Given the system of beliefs as addressed above, there are three potential candidates for an equilibrium that can be broadly characterized by the quality level and by whether sophisticated consumers or naive consumers become fully exploited by the monopolist:

1. *No exploitation*: All consumers buy a high quality good, and sophisticated consumers have a 0 consumption utility: $q = q_H$, $\mu(\gamma) = \tilde{\mu}(\gamma) = 1$, $U(\mu(\gamma), \gamma) = 0$, $\tilde{U}(\tilde{\mu}(\gamma), \gamma) > 0$, $D(\mu(\gamma), \gamma) = 1 - \theta$, $\tilde{D}(\tilde{\mu}(\gamma), \gamma) = \theta$

2. *Return cost exploitation*: Only naive consumers buy a good which is of high quality, and predict a 0 consumption utility: $q = q_H$, $\mu(\gamma) = \tilde{\mu}(\gamma) = 1$, $U(\mu(\gamma), \gamma) < 0$; $\tilde{U}(\tilde{\mu}(\gamma), \gamma) = 0$, $D(\mu(\gamma), \gamma) = 0$, $\tilde{D}(\tilde{\mu}(\gamma), \gamma) = \theta$

3. *Quality misperception exploitation*: Only naive consumers buy a good which is of low quality, and predict a 0 consumption utility: $q = q_L$, $\mu(\gamma) = 0$, $\tilde{\mu}(\gamma) = 1$, $U(\mu(\gamma), \gamma) < 0$, $\tilde{U}(\tilde{\mu}(\gamma), \gamma) = 0$, $D(\mu(\gamma), \gamma) = 0$, $\tilde{D}(\tilde{\mu}(\gamma), \gamma) = \theta$

Because of Assumption 3.1 it is never optimal for a monopolist to set a low quality while also attracting sophisticated consumers. When the monopolist wants to attract both consumer types, whenever $U(\mu(\gamma), \gamma) > 0$, the monopolist can raise prices without losing consumer demand, which is thus profitable. When the monopolist is selling to naive consumers only, analogously, whenever $\tilde{U}(\tilde{\mu}(\gamma), \gamma) > 0$, the monopolist can also increase the price without losing consumer demand, which therefore increases profits. We will now explore the three different cases in more detail.

Case 1: Both sophisticated and naive consumers buy a high quality good. We start this section with the following claim.

Claim 3.1. *No contract with $w > w^S$ in which $U(\mu(\gamma), \gamma) \geq 0$ can yield a higher profit to the monopolist than the contract specified by $w = w^s$ and $U(\mu(\gamma), \gamma) = 0$.*

Proof. Under the system of beliefs we have specified above, for every contract γ we have $\tilde{U}(\tilde{\mu}(\gamma), \gamma) \geq U(\mu(\gamma), \gamma)$. If the monopolist wants to attract all consumers to buy the good, the highest price the monopolist can set given any warranty level is a price such that the sophisticated consumers are indifferent between buying and not buying the good. Assumption 3.1 implies that it is more profitable for the monopolist in such a case to sell a high quality product to the consumers than a low quality product. In order to credibly signal a high quality level to sophisticated consumers, the warranty level has to be at least w^S. We now show that no warranty level can yield a higher profit to the monopolist than the level w^S. Denote the monopolist's profit when attracting all consumers by π^S. Thus, the firm's maximization problem can be written as $\max_{\gamma,q}\pi(\gamma, q) = p - (1-q)F(w)w - c(q)$ s.t. $Iq + (1-q)F(w)E[w-r|w > r] - p \geq 0$. The Lagrangean of the problem can be written as

$$L^S = p - (1-q)F(w)w - c(q) + \lambda[Iq + (1-q)F(w)E[w - r|w > r] - p].$$

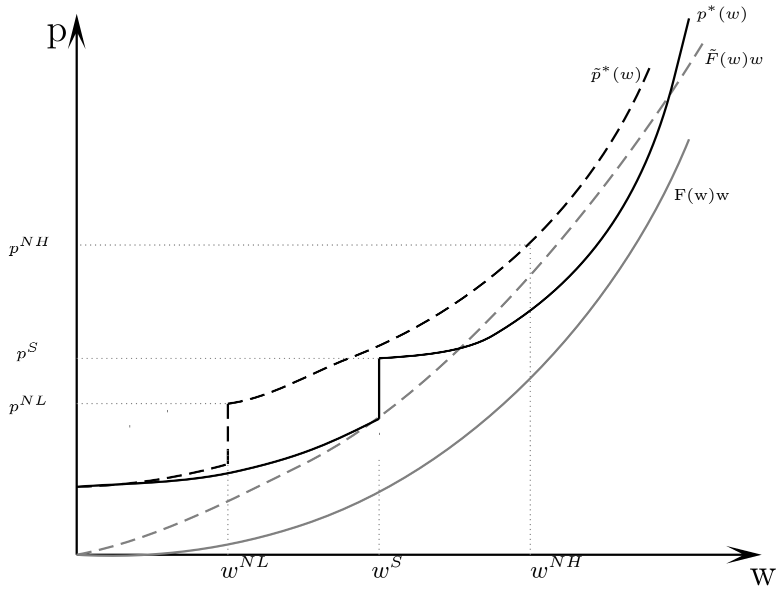


Figure 3.2: Consumer willingness to pay given warranty level

Taking the derivative with respect to p yields $\lambda = 1$. Intuitively, this implies that the monopolist will maximize the difference between the willingness to pay of sophisticated consumers and the production costs. With $\int_0^w rf(r)dr = [rF(r)]_0^w - \int_0^w F(r)dr$, this can then be written as $L = Iq - (1-q)[\int_0^w F(r)dr] - c(q)$. The first-order derivative with respect to w then yields

$$\frac{\partial L^S}{\partial w} = -(1-q)[F(w)] \leq 0 \forall w \geq 0. \tag{3.10}$$

Therefore, given $w \geq w^S$, the sum of consumer and producer surplus is never increasing in w. From Assumption 3.1, it follows that it is always profitable to offer a high quality q_H to sophisticated consumers, such that $w \geq w^S$; because only then $\mu(\gamma) = 1$ is fulfilled. Since the monopolist can at most obtain the whole rent $\pi^S + V(q_H, \gamma)$ by selling the good to the consumers, which happens when $V(q_H, \gamma) = 0$, it follows that he cannot earn a higher profit than when setting $w = w^S$ while setting a price such that $U(\mu(\gamma), \gamma) = 0$. This completes the proof. □

Claim 3.2. *Let $F(w)$ be continuous and strictly increasing around w^S. Then w^S is the unique profit-maximizing warranty level when selling to both consumer types.*

Proof. This follows immediately from Assumption 3.1 and equation 3.10. □

Case 2: Only naive consumers buy a high quality good. Again, from Assumption 3.1 it follows that it is always profitable for a firm to make the naive consumers believe that the good is of high quality. In this segment, it is thus best for the monopolist to set a warranty level equal or above w^S in combination with producing a high quality q_H. Denote the monopolist's profit in this case by π^S. The Lagrangean can be written as

$$L^{NH} = \theta \left[p - (1 - q_H)F(w)w - c_H + \lambda[Iq_H + (1 - q_H)\tilde{F}(w)\tilde{E}[w - r|w > r] - p]\right].$$

The first-order derivative with respect to p yields $\lambda = \theta$. Plugging this in the above equation and deriving with respect to w yields

$$\frac{\partial L^{NH}}{\partial w} = \theta(1 - q_H)\left[\tilde{F}(w) - F(w) - wf(w)\right]. \tag{3.11}$$

The second-order derivative yields

$$\frac{\partial^2 L^{NH}}{\partial w^2} = \theta(1 - q_H)\left[\tilde{f}(w) - 2f(w) - wf'(w)\right]. \tag{3.12}$$

Equation (3.11) shows that without further restrictions, there is not necessarily a unique maximum of L^{NH} with respect to w. A sufficient condition would be $\frac{w}{2} > \frac{f(w) - \tilde{f}(w)}{f'(w)}$ for all $w > w^S$. An optimal contract for this segment consists of a high quality q_H, a warranty level w that maximizes π^{NH}, and a price such that $\tilde{U}(1, \gamma) = 0$.
We introduce another assumption in order to better characterize the monopolist's optimal strategy.

Assumption 3.2. *(Bounds on consumer misperceptions) For all $w \geq w^N$, $\frac{\tilde{F}(w)}{F(w)} < \frac{1-q_L}{1-q_H}$.*

Case 3: Only naive consumers buy a low quality good. In this case, it is best for the monopolist to set a warranty level strictly below w^S but equal or above w^N in combination with producing the quality q_L. Denote the monopolist's profit in this segment by π^{NL}.

$$L^{NL} = \theta \left[p - (1 - q_H)F(w)w - c_H + \lambda[Iq_H + (1 - q_L)\tilde{F}(w)\tilde{E}[w - r|w > r] - p]\right].$$

The first-order derivative with respect to p yields $\lambda = \theta$. Deriving with respect to w yields The first-order derivative with respect to w yields

$$\frac{\partial L^{NL}}{\partial w} = \theta \left[(1 - q_H)\tilde{F}(w) - (1 - q_L)F(w) - (1 - q_L)f(w)w \right]. \tag{3.13}$$

Under Assumption 3.2 and equation (3.13), it follows immediately that w^N is the unique optimal warranty level in this case. Figure 3.3.1 shows the maximum price each consumer type is willing to pay given a warranty level w, and the resulting optimal contracts for the monopolist in each case.

3.3.2 Monopolist's optimal choice of a contract

Without further restrictions on the relationship between $F(r)$ and $\tilde{F}(r)$, it is not possible to characterize a unique warranty level by some first-order conditions. When the monopolist is targeting naive consumers only in each case, it is clear that an optimal warranty level exists for both the interval $[w^N, w^S]$ and the interval $[w^S, \infty)$. For notational ease, we define the difference between a sophisticated consumer's expected net warranty benefit of a high quality good and the monopolist's expected warranty costs per product when offering a product of quality q as $R(w, q) \equiv (1 - q_H)F(w)E[w - r|w > r] - (1 - q)F(w)w$. Analogously, define $\tilde{R}(w, q) \equiv (1 - q_H)\tilde{F}(w)\tilde{E}[w - r|w > r] - (1 - q)F(w)w$ as the difference between a naive consumer's expected net warranty benefit of a high quality good and the monopolist's expected warranty costs per product when offering a quality q. As shown above, the value w^S is already an optimal warranty value if the monopolist targets both naive and sophisticated consumers.[9] Denote a monopolist's optimal warranty level in the "return cost exploitation" case 2, i.e. when only naive consumers buy a high quality good, as $w^{NH} \in [w^S, \infty)$. Analogously, denote an optimal warranty level in the "quality misperception" case 3, i.e. when selling a high quality product only to naive consumers, by $w^{NL} \in [w^N, w^S]$. Furthermore, denote the maximum profits the firm makes in the three different cases as: π^{S*} when the monopolist also attracts sophisticated consumers; π^{NH*} when the monopolist sells a high quality good to naive consumers only; and π^{NL*} when the monopolist sells a low quality good to naive consumers only. We are now able to characterize the size of the measure of

[9]Here, uniqueness for a maximum in this segment is neither required nor assumed.

naive consumers, θ, for which the monopolist is indifferent between selling a good only to naive consumers and selling the good to all consumers.
From Case 1 we can see that the monopolist's maximum profit when attracting both consumer types becomes

$$\pi^{S*} = q_H I + R(w^S, q_H) - c_H. \tag{3.14}$$

The maximum profit in Case 2 can be written as

$$\pi^{NH*} = \theta \left[q_H I + \tilde{R}(w^{NH}, q_H) - c_H \right]. \tag{3.15}$$

Finally, the maximum profit in Case 3 is

$$\pi^{NL*} = \theta \left[q_H I + \tilde{R}(w^{NL}, q_L) - c_L \right]. \tag{3.16}$$

Therefore, the difference between the maximum profit in Case 3 and Case 2 can be written as

$$\pi^{NL*} - \pi^{NH*} = \theta \left[(\tilde{R}(w^{NL}, q_L) - \tilde{R}(w^{NH}, q_H)) - (c_L - c_H) \right]. \tag{3.17}$$

As long as this expression is positive, it is more profitable for the monopolist to sell a low quality product to the naive consumers than a high quality product. Selling the good to both sophisticated and naive consumers is at least as profitable as selling to naive consumers only if $\pi^{S*} \geq max\{\pi^{NL*}, \pi^{NS*}\}$. From equations (3.14)-(3.16), this can be written as

$$q_H I + R(w^S, q_H) - c_H \geq \theta[q_H I + \max\{\tilde{R}(w^{NL}, q_L) - c_L, \tilde{R}(w^{NH}, q_H) - c_H\}.$$

Solving this for θ, we obtain our first proposition.

Proposition 3.1. *The monopolist prefers to sell to both consumer types rather than to only naive consumers if and only if* $\theta \leq \frac{q_H I + R(w^S, q_H) - c_H}{q_H I + max\{\tilde{R}(w^{NL}, q_L) - c_L, \tilde{R}(w^{NH}, q_H) - c_H\}}$.

Proof. In text. □

Intuitively, the monopolist will only choose to sell to naive consumers if the increase in the mark-up from selling only to naive consumers offsets the drop in demand due to not selling to sophisticated consumers. Figure 3.3.2 shows how the choice of the

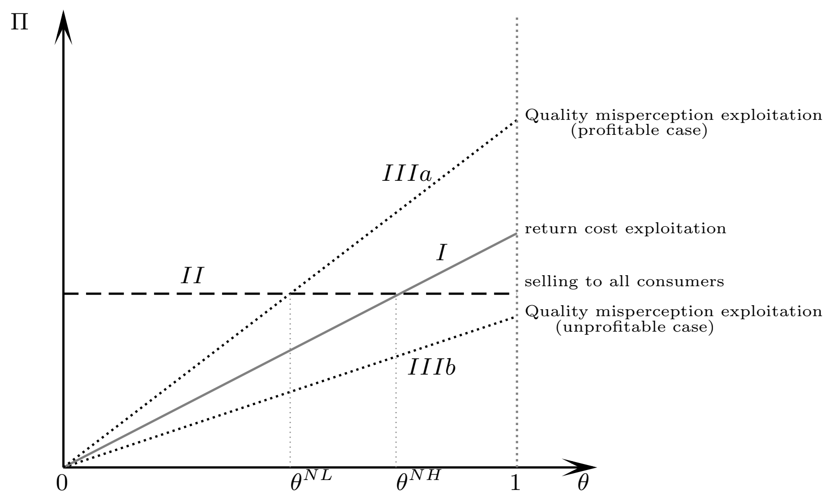


Figure 3.3: Profitability of contract choices by fraction of naive consumers

optimal contract is determined by the fraction of naive consumers θ. As explained above, under some circumstances it will never be optimal for a monopolist to exploit naive consumers' quality misperceptions. This can be seen from the lower grey dotted line $IIIb$. However, for sufficiently high θ, a monopolist will always have an incentive to exploit the consumers' return cost misperceptions. This is because the mark-up per consumer is always higher for the case of return cost exploitation (dashed line II) compared to the case when all consumers buy the good. The minimum threshold level of θ for which it is profitable to target naive consumers only results at the point at which the highest curve when targeting naive consumers only intersects with the flat curve I when targeting also sophisticated consumers.
We now give a numerical example.

3.3.3 Example

Example 3.1. *Suppose sophisticated consumers believe that the return costs s are uniformly distributed over* 0 *and* 9, $r \sim U[0,9]$, *while naive consumers believe that the return costs* $\hat{r}$ *are distributed uniformly over 0 and 4,* $\hat{r} \sim U[0,4]$. *Therefore* $F(r) = \frac{r}{9}, r \in [0,9]$, *and* $F(s) = 1 \forall r \geq 9$. *Analogously,* $\tilde{F}(r) = \frac{r}{4}, r \in [0,4]$, *and* $\tilde{F} = 1\ \forall r \geq 4$. *Let* $q_H = 0.75, q_L = 0.25$*;* $c_H = 3$, $c_L = 1$, $I = 8$. *From equation (3.6), it follows that* $\mu(\gamma) = 1$ *whenever* $F(w)w = \frac{w^2}{9} \geq \frac{c_H - c_L}{q_H - q_L} = 4$. *It follows that the weak inequality binds with strict equality if* $w^S = 6$. *From equations (3.8) and (3.9), it follows similarly for naive consumers that* $\tilde{\mu}(\gamma) = 1$ *whenever* $\tilde{F}(w)w = \frac{w^2}{4} \geq 4$. *This weak inequality* *binds with strict equality if* $w^N = 4$. *From Claim* 3.1, *we know that no contract can be more profitable when selling to both naive and sophisticated consumers than the one specified by* $\gamma = (p, w) = (Iq_H + \int_0^{w^S} F(r)dr, w^S) = (6.5, 6)$, *in which case it is profitable for the monopolist to set* $q = q_H$. *This yields an expected profit of* $\pi^{S*} = \theta(p - c_H - (1 - q_H)F(w^S)w^S = 2.5$. *When selling to naive consumers only, one can see that the two first-order conditions (3.11) and (3.13) are weakly decreasing in the warranty level* w *for the relevant ranges, i.e. for* $w \in [4, 6]$ *in equation (3.13), and for* $w \in [6, \infty)$ *in equation (3.11). Therefore, the optimal contracts when only attracting naive consumers are given by* $\gamma = (p, w) = (Iq_H + \int_0^{w^{NL}} \tilde{F}(r)dr, w^{NL}) = (6.5, 4)$ *when providing a low quality, and* $\gamma = (p, w) = (Iq_H + \int_0^{w^{NH}} \tilde{F}(r)dr, w^{NH}) = (7, 6)$ *when providing a high quality. The profit in the former case case is* $\pi^{NL*} = 4\frac{1}{6}\theta$, *and is thus higher than in the latter case with with* $\pi^{NH*} = 3\theta$. *So it follows that the monopolist rather sells to both consumer types whenever* $\pi^{S*} \geq \pi^{NL*}$, *which is equivalent to* $\theta \geq \frac{3}{5}$.

3.3.4 Discussion of the monopolist's choice

Exploitation device As shown before, the monopolist will optimally pick the quality-contract combination that yields the highest utility among all combinations located in the three segments above. Only the first case yields a non-negative consumer surplus. In this case, the sophisticated consumers' correct evaluation of the good saves the naive consumers from becoming exploited because of their false beliefs about both the good's quality and their return cost distributions. This is optimal for the monopolist if the increase in demand due to also attracting the sophisticated consumers compensates for the lower maximum profit per consumer when also targeting sophisticates. When the monopolist chooses to attract only naive consumers, he is faced with two choices: The first choice is to set a high quality and a warranty level above w^S in order to exploit naive consumers' false beliefs about their return costs. This is similar to many overconfidence or self-control models in the literature because consumers only mispredict their costs but not the good's quality.[10] The second possibility for the monopolist is to make naive consumers believe that the good is of high quality and thus charge a high price while producing a good of low quality. In this case, the main form of exploitation works through successfully selling a low quality product for the price of a high quality

[10] See for example DellaVigna and Malmendier (2004), or Eliaz and Spiegler (2006).

product, while still exploiting false believes about consumer replacement costs.
As most Behavioral Industrial Organization models rely on exploitation through pricing schemes, this usually works best through multi-period contracts, as for example in gym contracts, or through hidden add-on prices not anticipated by some consumers before consuming the good. Our model also has a multi-period characteristic in terms of a possible future breakdown after purchasing the good which then leads to a replacement stage of the game. However, the form of exploitation in the second case does not arise from the pricing scheme, but from the warranty setting through false signaling. In the two cases in which only the naive consumers buy the good, their willingness to pay is above their "true" expected utility that takes into account the correct return cost cdf $F(w)$. Therefore, the consumer surplus in these cases is negative, and the monopolist will benefit from it. In each case, sophisticated consumers always earn 0 consumer surplus. Thus, naive consumers do not exert a negative externality on sophisticated consumers.

Minimum assumptions on naiveté In order to establish that naive consumers misperceive product qualities in the model, the baseline setting also requires that these consumers do not condition their beliefs on the overall profitability of the firm. This means that these consumers do not ask themselves whether the offered contract to consumers is the profit-maximizing contract for a firm, but rather whether given the consumers in the market, it makes sense for the firm to produce a high quality good as compared to a low quality good.

Prices as signal for demand In our model, unless there is a strict money-return policy, i.e. $w = p$, prices cannot signal quality. This is because the revenue from prices a firm receives is irrespective of the probability of a product breakdown. This in turn implies that the compensatory warranty payment in case of a breakdown is the sole channel for signaling the product quality.

3.4 Oligopoly

We present two different oligopoly settings. In the first one, firms set their contracts and quality levels simultaneously. In the second one, firms set their contracts and qualities sequentially. Similar to the monopoly setting, all firms' qualities are unobservable to all consumers. After the firms have set their contracts and quality levels, consumers observe all contracts and make inferences about each firm's quality level. We now focus on the consumers' decision making process in more detail.

3.4.1 Consumers' contract choices

After all firms have set their contracts, the consumers observe all contracts $\Gamma = (\gamma_1, .., \gamma_M)$ and can choose whether to buy one unit of one of the M products. For each contract γ_i, consumers assign beliefs about the firm i producing a product of high quality. Denote $\mu = (\mu_1, .., \mu_M)$ the vector of beliefs of the sophisticated consumers about the probability of each firm being of high quality. Analogously, denote $\tilde{\mu} = (\tilde{\mu}_1, .., \tilde{\mu}_M)$ the vector of beliefs of naive consumers about the probability of each firm being of high quality. Given the beliefs about the quality of each product conditional on the contracts observed, a consumer chooses the contract that maximizes her expected utility if this utility is non-negative. In case of $K \leq M$ products give her the highest expected non-negative utility, she chooses all firms with probability $\frac{1}{K}$. In the case that a consumer has bought a good that breaks down, she can decide to take the good back to the producer in order to claim her warranty payment. She will do so if the warranty payment is at least as high as her return costs, i.e. if $w_i \geq r$. The utility of a sophisticated consumer when consuming firm i's good of quality $q_i \in \{q_H, q_L\}$ can be written as

$$V_i(q_i, \gamma_i) = q_i I - p_i + (1 - q_i) \int_0^{w_i} [w_i - r] f(r) dr. \tag{3.18}$$

The predicted utility for a naive consumer of consuming one unit of firm i's product becomes

$$\tilde{V}_i(q_i, \gamma_i) = q_i I - p_i + (1 - q_i) \int_0^{w_i} [w_i - r] \tilde{f}(r) dr. \tag{3.19}$$

Therefore, after having observed all contracts $\Gamma = (\gamma_1, .., \gamma_M)$, a consumer's willingness to pay for a good $i \in \{1, .., M\}$, $U_i(\mu(\Gamma), \gamma_i)$, yields

$$U_i(\mu(\Gamma), \gamma_i) = \mu_i(\Gamma) V_i(q_H, \gamma_i) + (1 - \mu_i(\Gamma)) V_i(q_L, \gamma_i). \tag{3.20}$$

Similarly, a naive consumer's willingness to pay for good i can be written as

$$\tilde{U}_i(\tilde{\mu}(\Gamma), \gamma_i) = \tilde{\mu}_i(\Gamma)\tilde{V}_i(q_H, \gamma_i) + (1 - \tilde{\mu}_i(\Gamma))\tilde{V}_i(q_L, \gamma_i). \tag{3.21}$$

We will now focus on the firms' demand and profit functions.

3.4.2 Firms' maximization problem

We define the set of M firms $\{1, .., M\} \equiv \mathtt{M}$. For a firm $i \in \mathtt{M}$, denote the set of its competitors $\mathtt{M}_{-i} \equiv \mathtt{M}\backslash\{i\}$. The demand of sophisticated consumers for firm i's product given the set of contracts Γ and beliefs $\mu(\Gamma)$, $D_i(\Gamma, \mu(\Gamma))$, then becomes

$$D_i(\Gamma, \mu(\Gamma)) = \begin{cases} 1-\theta, & \text{if } U_i(\mu(\Gamma), \gamma_i)) > U_j(\mu(\Gamma), \gamma_j) \forall j \in \mathtt{M}_{-i} \text{ and } U_i(\mu(\Gamma), \gamma_i) \geq 0 \\ \frac{1-\theta}{K}, & \text{if } U_i(\mu(\Gamma), \gamma_i)) \geq U_j(\mu(\Gamma), \gamma_j) \forall j \in \mathtt{M}_{-i}, U_i(\mu(\Gamma), \gamma_i)) = U_j(\mu(\Gamma), \gamma_j) \\ & \text{ for } K-1 \text{ different } j \in \mathtt{M}_{-i} \text{ and } U_i(\mu(\Gamma), \gamma_i) \geq 0 \\ 0, & \text{if } U_i(\mu(\Gamma), \gamma_i)) < U_j(\mu(\Gamma), \gamma_j) \text{ for at least one } j \in \mathtt{M}_{-i} \\ & \text{ or if } U_i(\mu(\Gamma), \gamma_i) < 0. \end{cases}$$

The demand of naive consumers for firm i's product given the contracts Γ and beliefs $\tilde{\mu}(\Gamma) = (\tilde{\mu}_i(\Gamma), \tilde{\mu}_{M_{-i}}(\Gamma))$, $\tilde{D}_i(\Gamma, \tilde{\mu}(\Gamma))$, turns to

$$\tilde{D}_i(\Gamma, \tilde{\mu}(\Gamma)) = \begin{cases} \theta, & \text{if } \tilde{U}_i(\tilde{\mu}(\Gamma), \gamma_i)) > \tilde{U}_j(\tilde{\mu}(\Gamma), \gamma_j) \forall j \in \mathtt{M}_{-i} \text{ and } \tilde{U}_i(\tilde{\mu}(\Gamma), \gamma_i) \geq 0 \\ \frac{\theta}{K}, & \text{if } U_i(\tilde{\mu}(\Gamma), \gamma_i)) \geq U_j(\tilde{\mu}(\Gamma), \gamma_j) \forall j \in \mathtt{M}_{-i}, U_i(\tilde{\mu}(\Gamma), \gamma_i)) = U_j(\tilde{\mu}(\Gamma), \gamma_j) \\ & \text{ for } K-1 \text{ different } j \in \mathtt{M}_{-i} \text{ and } U_i(\tilde{\mu}(\Gamma), \gamma_i) \geq 0 \\ 0, & \text{if } U_i(\tilde{\mu}(\Gamma), \gamma_i)) < U_j(\tilde{\mu}(\Gamma), \gamma_j) \text{ for at least one } j \in \mathtt{M}_{-i} \\ & \text{ or if } U_i(\tilde{\mu}(\Gamma), \gamma_i) < 0. \end{cases}$$

Therefore, given these demand functions, firm i's profit function becomes

$$\pi_i(\Gamma, \mu, \tilde{\mu}, q_i) = [D_i(\Gamma, \mu(\Gamma)) + \tilde{D}_i(\Gamma, \tilde{\mu}(\Gamma))](p_i - c(q_i) - (1 - q_i)F(w_i)w_i). \tag{3.22}$$

Firm i's profit maximization problem can thus be written as

$$\max_{\gamma_i, q_i} \pi_i(\Gamma, \mu, \tilde{\mu}, q_i) \text{ s.t. } (p_j, \gamma_j) \in \arg\max_{\gamma_j, q_j} \ \pi_j(\Gamma, \mu, \tilde{\mu}, q_j) \ ; \forall j \in \mathtt{M}_{-i}. \tag{3.23}$$

3.4.3 Belief structure

Similar to the monopoly case, we impose an equilibrium refinement in which consumers' beliefs follow the producers' dominant actions, with the convention that when producers are indifferent between producing a high or a low quality product, consumers believe that the firm produces a high quality product. Furthermore, we require the beliefs of the sophisticated consumers to be true in equilibrium, and those of naive consumers to be true under the hypothetical situation that their expected distribution of return costs is their true distribution of return costs. Thus, sophisticated consumers' beliefs about firm $i \in \{0, 1\}$ delivering a high quality product, $\mu_i(\Gamma)$, can be written as

$$\mu_i(\Gamma) = \begin{cases} 1, & \text{if } w_i \geq w^S \\ 0, & \text{if } w_i < w^S. \end{cases} \tag{3.24}$$

Given the set of contracts, Γ, from all firms in the market, including contract γ_i for firm i, a naive consumer's belief about firm $i \in \mathtt{M}_{-i}$ being of high quality, $\tilde{\mu}_i$, will be of the following form:

$$\tilde{\mu}_i(\Gamma) = \begin{cases} 1, & \text{if } w_i \geq w^N \\ 0, & \text{if } w_i < w^N. \end{cases} \tag{3.25}$$

3.4.4 Simultaneous contract setting

We first make some useful definitions. Recall the optimal warranty levels for the three different cases from the monopoly section, w^{NL}, w^S, and, w^{NH}. Regarding multiplicity of maxima in each segment, w^{NL} and w^{NH} will now always refer to the smallest of these warranty levels, respectively. We now define the sum of marginal costs of production and expected warranty payments of a firm as its quasi-marginal costs: $\hat{c}_{NL} \equiv c_L + (1 - q_L)F(w^{NL})w^{NL}$; $\hat{c}_S \equiv c_H + (1 - q_H)F(w^S)w^S$; $\hat{c}_{NH} \equiv c_H + (1 - q_H)F(w^{NH})w^{NH}$. Furthermore, we define Λ as the difference in a naive consumer's predicted utility of consuming a product with warranty level w^{NH} priced at quasi-marginal cost and of consuming a product with warranty level w^{NL} priced at quasi-marginal costs:

$$\begin{aligned} \Lambda &\equiv \tilde{U}(\hat{c}_{NH}, w^{NH}) - \tilde{U}(\hat{c}_{NL}, w^{NL}) \\ &= c_H - c_L - (1 - q_H) \int_{w^{NL}}^{w^{NH}} \tilde{F}(r)dr - (1 - q_L)F(w^{NL})w^{NL} + (1 - q_H)F(w^{NH})w^{NH}. \end{aligned}$$

When there are at least four firms in the market, we can show that there always exists an equilibrium in which all firms make zero profits. This is because given our belief refinement, in such a case Bertrand competition in both the segment for naive and the one for sophisticated consumers drives prices down.

Proposition 3.2. *Let $M \geq 4$. Then there always exists a pure strategy equilibrium in which at least two firms set $\gamma_S = (p_S, w^S) = (\hat{c}_S, w^S)$ and $q = q_H$, and*
i) if $\Lambda < 0$, at least two firms set $\gamma_{NL} = (p_{NL}, w^{NL}) = (\hat{c}_{NL}, w^{NL})$ and $q = q_L$
ii) if $\Lambda \geq 0$, at least two firms set $\gamma_{NH} = (p_{NH}, w^{NL}) = (\hat{c}_{NH}, w^{NH})$ and $q = q_H$.
All sophisticated consumers will buy a product that maximizes their utility, while all naive consumers will not.

Proof: See Appendix.
Proposition 3.2 implies that no matter how big the fraction of naive consumers in the market, if there are at least four firms in the market, competition will lower the price for all goods down to zero profits for all firms, while also providing an inefficient product choice for naive consumers. This follows again from the wrong inferences consumers make from the warranties. Intuitively, Bertrand competition drives prices down to the quasi-marginal costs.
If less than four firms are present in the market, there is the problem that for many parameter constellations there does not necessarily exist an equilibrium in which all firms set contracts simultaneously. Here, due to the presence of naive consumers, competition does not immediately move prices down to zero profits for the firms. Given one firm sets a high warranty level and prices such that its profits are close to zero, the other firm has an incentive to set a warranty level such that only naive consumers believe the good is of high quality with a lower price than the good of the other firm, and earn a positive profit. In order to overcome these problems, we introduce at a variation of the game, in which firms set their contracts and qualities not simultaneously.

3.4.5 Sequential contract setting

If firms set their contracts sequentially, each firm has to take into account its followers' strategies when setting its own contract and product quality. When there are only two firms in the market, they both always make positive profits. This is because of the existence of two different consumer types, namely sophisticated and naive consumers. This ensures that the first firm can always set a price above its quasi-marginal costs, for which it makes sense for the second firm not to undercut, but rather to fully serve only one consumer type. This leads to our next proposition.

Proposition 3.3. *If $M = 2$, then in each equilibrium of the sequential contract setting game, both firms make positive profits, and the following firm's profits are at least as high as the leading firm's profits.*

Proof: See Appendix.

If there are more than two firms in the market, then there is the problem of a potential multiplicity of equilibria. We restrict our analysis to cases in which all firms in the market have a positive market share. Under this restriction, there always exist two types of equilibria. In the first equilibrium type, all firms set prices equal to their quasi-marginal costs, and naive consumers always choose a non-utility maximizing contract. In the second equilibrium type, the first $M-1$ firms all set the same contract, in which they set prices equal to quasi-marginal cost, which is the optimal contract for one consumer type. Firm M then sets a price above marginal cost and attracts the other consumer type. Only if firm M serves sophisticated consumers in such an equilibrium, all consumers make utility-maximizing choices. We sum this up in the next proposition.

Proposition 3.4. *If $M \geq 3$, then in all equilibria of the sequential contract setting game in which all firms have positive market shares, at most firm M makes positive profits. In case firm M makes no positive profit, this is sufficient for all naive consumers to make non-utility maximizing choices.*

Proof: See Appendix.

All in all, this section has shown that competition lowers prices, but often does not prevent all consumers from making non-optimal decisions. In the sequential setting, if there are only two firms in the market, it might still be possible that naive consumers pay more for the good than their true willingness to pay.

3.5 Extensions

In this section, we present several extensions to the baseline model.

Money return policies A common form of warranties are money return policies, in which case $w = p$. This puts some restrictions on the warranty-price relationship. Interestingly, depending on the magnitude of Iq_H, it is never be optimal for the monopolist to exploit naive consumers' quality misperceptions. This is because offering a high quality product is optimal for the monopolist as soon as $Iq_H + \tilde{F}(w)\tilde{E}[p-r|p > r] \geq F(p)p \geq \frac{c_H - c_L}{q_H - q_L}$. Under Assumption 3.1, this always holds. Without the $w = p$ restriction, it is always possible for the monopolist to set a warranty level just high enough to (falsely) signal a high quality product to naive consumers, while extracting the willingness to pay by setting a high price. When also aiming the good at sophisticated consumers, the $w = p$ restriction decreases the overall welfare. This is because once $p > w^S$, which is always the case, there is a welfare loss due to the increase in return cost frictions. When exploiting consumers' return cost misperceptions only, the monopolist's profits can fall compared to the case when $w = p$ does not hold. This is because the monopolist has to steer the extraction of the consumer rent and to set a price as to maximally distort consumers' beliefs through the same channel.

Mail-in rebates Once there is no uncertainty about product qualities, our model can be used to explain the existence of money-return mail-in rebate policies. Because of the quality certainty, a consumer is always willing to pay a total amount I for a good. When facing sophisticated consumers only, the contract $\gamma = (p, w) = (I, 0)$ is always optimal. However, when facing naive consumers only, the optimal contract can be written as $\gamma = (p, w) = (I + F(w)\tilde{E}[w - r|w > r], w)$, where $w \in \arg\max_w I + \tilde{F}(w)\tilde{E}[w - r|w > r] - F(w)w$. The optimal warranty level w therefore maximizes the term $\int_0^w \tilde{F}(r)dr - F(w)w$. This arises because by offering a coupon worth w, the monopolist can increase the base price of the good, while exploiting the naive consumers' false predictions about the propensity to return a good. Furthermore, if the monopolist is able to offer a variety of contracts to consumers, he can set both of the above contracts in order to maximize his profits. This is a consequence of an absence of coupon payment in a contract targeted towards sophisticated consumers. Therefore, both naive and sophisticated consumers have the same valuation for the good, and only disagree in their valuation for the "coupon-contract". This is not an optimal strategy for the earlier sections in which the monopolist has to signal a high quality to sophisticated consumers by offering a warranty payment.

Other potential reasons for consumer misperceptions There are several channels through which consumer misperceptions can be modeled. Inderst and Ottaviani (2013) model a relationship between a financial advisor and a private investor in which granting a cancellation right to consumers can increase a seller's credibility in a cheap talk game. In contrast to our model, naive consumers in their model always believe every advice the seller gives them, leading the seller always claiming that a product is the most suitable for a consumer. Thus, a consumer's key concern in their model is the suitability of a product to the individual taste, which amounts to horizontal product differentiation rather than vertical product differentiation as in our model. From a policy perspective, their recommendations are essentially opposite to ours: Cancellation rights will offset some bad advice from sellers.

Another way of viewing our model would be that the naive consumers do not see some hidden warranty return costs, such as a consumer's shipping costs if she wants to return the product. These shrouded costs would lead to an overestimation of a warranty, which would then cause a misperception of product qualities. This kind of consumer myopia would be similar to Gabaix and Laibson (2006). There are however several aspects that differ from this model. First, we believe that the psychological costs of dealing with a warranty reimbursement is very difficult to be unshrouded in a Gabaix and Laibson (2006) fashion. Such costs are very hard to be quantified, and rival warranty suppliers will hardly be able to unshroud such costs.[11] Second, in our model, a warranty cannot be simply seen as a substitutable add-on. This is because of the signaling role of a warranty with respect to product quality.

Self-control costs due to hyperbolic discounting could be seen as one of the explanations for return costs. However, a purely hyperbolic model would have to imply naivete about self-control costs in each period in order to be able to predict a behavior of never returning a product. In any other case, there would be a last period in which returning the product to claim a warranty would be optimal. Drago and Kadar (2006) overcome this problem in a mail-in-rebate setting by furthermore introducing a sunk-costs regret effect.

Consumer policy intervention: minimum warranty level In the monopoly case, it should be obvious that a policy intervention in terms of a minimum warranty level $w^{min} \geq w^S$ can only be useful if the monopolist is exclusively selling to naive

[11] According to OFT (2012), one of the major problems in the extended warranty business is that retailers have a near-monopoly status at the point of sale. In such a case, firms have even less incentive to unshrouding warranty costs.

Without intervention	$w \geq w^S$	Change due to intervention
$\Pi^{NL} > \Pi^S > \Pi^{NH}$	Π^S	no more exploitation
$\Pi^{NL} > \Pi^{NH} > \Pi^S$	Π^{NH}	now return cost exploitation
$\Pi^{NH} > \Pi^{NL} > \Pi^S$	Π^{NH}	no change
$\Pi^{NH} > \Pi^S > \Pi^{NL}$	Π^{NH}	no change

Table 3.1: Effects of introduction of a minimum warranty level

consumers, and produces a low quality. However, even in such a case a minimum warranty level might not prevent naive consumers from becoming exploited. This is because of the existence of two different exploitative equilibria. If it is most profitable to sell a low quality to naive consumers in equilibrium, but it is more profitable to sell high quality goods to naive consumers when compared to selling high quality goods to all consumers, then such a policy intervention is not sufficient to prevent exploitation of naive consumers. Table 3.1 sums up the changes from such a consumer intervention when the monopolist had initially been targeting naive consumers only.

Differentiated tastes for quality in oligopoly So far we have assumed a homogeneous preference for product quality. When preferences with respect to qualities are differentiated and for some part of the consumers, the added utility from higher quality does not offset the increase in marginal costs, these consumers prefer low quality products to be offered.

Assume that a share of μ consumers in the population has preferences $u(q)I$ for consuming a product of quality q. We refer to these consumers as low responsiveness types, and to the others as high responsiveness types. In both groups, the share of naive and sophisticated consumers remains distributed at fractions θ and $1-\theta$, respectively. Assume further that $(q_H - q_L)I > \hat{c}_S - c_L > [u(q_H) - u(q_L)]I$. This implies that the difference in the marginal costs of production including warranty expenses between a low and high product quality exceeds the difference in the consumption utilities of low responsiveness types. Assume for example that $u(q_H) = u(q_L) = Iq_H$. This is an extreme case, in which the low responsiveness consumers do not value any increase in product quality. In such a case, one can show instances in which a policy intervention in terms of a minimum warranty level is harmful to the low responsiveness consumers. In monopoly, a shift from an exploitative low quality contract to a high quality contract increases the utility of high responsiveness consumers. For low responsiveness consumers, however, when an increase in quality causes an increase in the price, such an intervention hurts their consumption utility. In the oligopoly case, if the number of firms is sufficiently high, then there always exists an equilibrium in the simultaneous contract setting game in which sophisticated low types buy a low quality good without any warranty offered. Imposing a minimum warranty level $w^{min} = w^S$ then has two detrimental effects: Since all firms then offer high product qualities in the market, naive high types are better off because of the intervention. However, low preference consumers lose from such a policy intervention. Therefore, if the fraction of low preference consumers in the population is sufficiently high, such a policy intervention reduces overall consumer surplus.

3.6 Conclusion

This chapter formalizes an economic model that captures several traits with respect to consumers' return behaviors, choices of extended warranties, and quality misperceptions. The presented evidence suggests that return costs play a significant role in the consumers' complaint behavior. We show that wrong inferences about a firm's quality, which increase the consumer's willingness to pay in the buying period, can lead to consumer exploitation by firms once the share of naive consumers is sufficiently high. Consumer policy interventions in terms of a minimum warranty payment can prevent firms from selling low quality products. This does not always stop consumers from becoming exploited by overpaying for an extended warranty contract. Competition always decreases prices for the goods in the market. However, this does not crowd low quality goods out of the market, and naive consumers often not buy their most suitable product. Note that the implications of our model are in line with some specific formalized psychological biases. One example is the concept of projection bias, which was introduced by Loewenstein et al. (2003). Projection bias describes situations in which people fail to fully project their future taste changes relative to their current taste. In our case one could think of different states as instances in which consumers are either "active" with respect to both the purchase or return of a good, or instances in which they are "inactive", and thus have higher return costs. Therefore, if consumers always buy in an active state, they on average put too much weight on low return cost events, and are thus also susceptible to exploitative contracts.

We show that in a differentiated oligopoly, consumer policy interventions can hurt overall consumer welfare because of restricting consumers with a low preference for quality to purchase high quality goods. This illustrates one problem decision-makers face when using behavioral models such as ours for policy purposes. Even though consumer protection policies may prevent some consumers from making mistakes, it can have adverse effects on other people because of restricting their choices. Therefore, any such policy should be implemented with caution, and only after weighting its advantages and disadvantages.

There are some important open questions that we have not addressed in this chapter. First, it is of importance to understand to what extent consumers learn from previous bad experiences. Agarwal et al. (2009) study how consumers' behavior changes after having to pay overlimit or cash advance fees in a certain month. They find that shortly after such an incident, consumers avoid paying fees by changing their behavior, but that in later months, they fall back into old manners and eventually pay higher fees again. Second, we do not focus much on interactions between different consumers. We believe that interactions between consumers, such as leadership of some consumers, or the presence of network effects, can give further insights into the determinants of contractual structures in the presence of consumer misperceptions.

3.A Proofs

Proof of Proposition 3.2

Proof. Define a strategy s_i of a firm i as the triple (p_i, w_i, q_i). We first consider the case $\Lambda \geq 0$. We show first that when $M = 4$, there exists an equilibrium in which two firms play the strategy $s^S \equiv (\hat{c}_S, w^S, q_H)$, all sophisticated consumers buy from these firms, two firms play the strategy $s^{NH} \equiv (\hat{c}_{NH}, w^{NH}, q_H)$, and all naive consumers buy from these firms. From our belief refinements it follows that sophisticated consumers assign belief $\mu_i = 1$ about firm i's product quality as long as $w_i \geq w^S$. From Claim 3.1 and Assumption 3.1 it follows that the utility maximizing contract for a sophisticated consumer includes a warranty level w^S, which signals a high quality level q_H. Because $\hat{c}_S$ is already the quasi-marginal costs for a firm that produces a high quality product with warranty level w^S, it follows that no firm can profitably deviate to attract sophisticated consumers by playing a strategy different from s^S, which yields 0 profits. Similarly, per definition, if $\Lambda \geq 1$, a contract that maximizes a utility of a naive consumer includes the warranty level w^{NH}, which also implies the production of a high quality level for a firm. Because $\hat{c}_{NH}$ is also the quasi-marginal cost of a firm offering such a contract, it follows that no firm can profitably deviate in order to attract naive consumers by setting a contract that is different from the contract s^{NH}, which also yields 0 profits. Because both s^S and s^{NH} yield 0 profits, it thus follows that no firm can profitably deviate. Assume now that $M > 4$. Then, once at least 2 firms play the strategy s^S, and two firms play the strategy s^{NH}, then it follows again that no firm can profitably attract any consumer type by setting a different contract.

Now consider the case $\Lambda \leq 0$. We show next that when M=4, there exists an equilibrium in which two firms play the strategy $s^S \equiv (\hat{c}_S, w^S, q_H)$, all sophisticated consumers buy from these firms, two firms play the strategy $s^{NL} \equiv (\hat{c}_{NL}, w^{NL}, q_L)$, and all naive consumers buy from two of these firms. The proof goes similar to the proof for $\Lambda \geq 0$. Again, in such a situation, it follows from Claim 3.1 and Assumption 3.1 that no firm can profitably attract sophisticated consumers by setting a strategy different to s^S, which yields exactly 0 profit. Per definition, if $\Lambda \geq 1$, a contract that maximizes a naive consumer's predicted consumption utility includes the warranty level w^{NL}, which also implies the production of a low quality level q_L for a firm. Because $\hat{c}_{NL}$ is also the quasi-marginal cost of a firm offering such a contract, it follows that no firm can profitably deviate in order to attract naive consumers by setting a contract that is different from the contract s^{NL}, which yields exactly 0 profits. Because both s^S and s^{NL} yield 0 profits, it thus follows that no firm can profitably deviate. Assume now that $M > 4$. Then, it follows again that once at least 2 firms play the strategy s^S, and two firms play the strategy s^{NL}, then no firm can profitably attract any consumer type by setting a different strategy. This completes the proof. $\square$

Proof of Proposition 3.3

Proof. We denote the firm that sets its contract first as firm 1, and the second firm as firm 2. Recall first, that per definition,

1.) $U(\hat{c}_S, w^S) > U(\hat{c}_{NH}, w^{NH})$ and 2.) $\tilde{U}(\hat{c}_S, w^S) < \tilde{U}(\hat{c}_{NH}, w^{NH})$. From this, it follows that $\forall\theta \in (0,1) \exists\ p_1, > \hat{c}_S$ s.t.

$\exists p_2$ s.t.

1. $U(p_1, w^S) > U(p_2, w^{NH})$
2. and $\tilde{U}(p_2, w^{NH}) > \tilde{U}(p_1, w^S)$
3. and $\theta(p_2 - \hat{c}_{NH}) \geq (p_1 - \hat{c}_S)$.

Therefore, there exists a strategy (p_1, w^S, q_H) with a price p_1 for which firm 1 can ensure strictly positive profits. But since the game is sequential and only has two stages, it thus follows that firm 1 can always ensure positive profits. Next, we show, that firm 2's profits in equilibrium must be at least as high as firm 1's profits in equilibrium. Suppose not. Then firm 2 can always just slightly undercut firm 1's price while setting same warranty level and product quality and therefore having a higher profit than firm 1, which is a contradiction. This completes the proof.

□

Proof of Proposition 3.4

Proof. In the following, we will denote the i^{th} firm to set its contract and product quality as firm i. Consider $M = 3$, and $\Lambda > 0$. We first show that in this case, player 1 will never make positive profits. Assume the contrary. Then it follows that one of the followers either will make 0 profit or set the identical contract. Furthermore, it follows that player 1's strategy $s_1 \notin \{(\hat{c}_S, w^S, q_H), (\hat{c}_{NH}, w^{NH}, q_H)\} \equiv S_1^*$. Focus now on firm 3. Given strategy s_1 of firm 1 and strategy $s_2(s_1)$ of firm 2 such that it makes non-negative profits, since $s_1 \notin S_1^*$, this already implies that firm 3 can make positive profits, because at least one consumer segment, i.e. naive or sophisticated consumers, does not face a price equal to quasi-marginal costs. Therefore, by undercutting one of the two predecessors in terms of price while setting the same warranty level, it can ensure a positive profit. Focus now on firm 2. As long as $s_1 \notin S_1^*$, from the proof of Proposition 3.3 we know that firm 2 can now always set a contract $s_2 \notin S_1^*$, such that firm 3 will set a different warranty level and firm 2 thus makes positive profits. However, in this case, it can never be optimal for firm 2 to set $s_2(s_1) = s_1$. Suppose $s_1 = (p_1, w_1, q_1)$ is such that $s_2(s_1) = s_1$ would ensure firm 2 and firm 1 positive profits. In such a case, firm 2 can strictly increase its profits by playing the strategy $(p_1 - \epsilon, w_1, q_1)$, where ϵ is infinitesimally small. Therefore, firm 1 can never make a positive profit. Thus, it follows that the only possibility for firm 1 of having a positive market share is whenever $s_1 \in S_1^*$. Denote such a strategy as s_1^*. But from a backward induction argument it then follows that for firm 2 to have a positive market share, $s_2^*(s_1^*) \in S_1^*$. Otherwise, firm 3 could always profitably set a contract such that firm 2 would have 0 demand. But then it follows that there are only 2 different types of equilibria. Whenever $s_1^* \neq s_2^*(s_1^*)$, it follows that firm 3 will only have a positive market share without making negative profits whenever $s_3(s_2^*(s_1^*)) \in S_1^*$. In such a case, all sophisticated consumers will buy a contract with warranty level w^S at the price of the quasi-marginal costs $\hat{c}_S$, and all naive consumers will buy a contract with warranty level w^{NH} at the price of the quasi-marginal costs $\hat{c}_{NH}$, which is clearly not utility maximizing. On the contrary, if $s_2^*(s_1^*) = s_1^*$, then it follows again from the proof of Proposition 3.2 that firm 3 can play a strategy $s_3(s_2^*(s_1^*)) \notin S_1^*$ that ensures a strictly positive profit. Now suppose that $s_1^* = s_2^*(s_1^*) = (\hat{c}_{NH}, w^{NH}, q_H)$ Then it follows that firm 3 will maximize its profits by setting $w = w^S$, and p_3 s.t. $U(p_3, w^S, q_H) = U(\hat{c}_{NH}, w^{NH}, q_H)$. In such a case it follows that naive consumers will make profit maximizing choices.

Now suppose that $M > 3$. From a simple induction argument it follows again that firm 1 will never be able to make positive profits in equilibrium. Conditional on firm 1 playing strategy $s_i \in S_1^*$, it then follows that all firms $i \in \{2, .., M-1\}$ then always have to play strategies $s_i(s_1^*, ...s_{i-1}^*) \in S_1^*$ in order to have a positive market share without having negative profits.

The proof for $\Lambda \leq 0$ goes analogously, now with $S_1^* \equiv \{(\hat{c}_S, w^S, q_H), (\hat{c}_{NL}, w^{NL}, q_L)\}$. This completes the proof.

□

3.B Different beliefs for naive consumers

Naive consumers only underestimate their own return costs It is furthermore necessary to define the consumers' beliefs about the other players' warranty return costs. Sophisticated consumers correctly believe that all consumers have the return cdf $F(r)$. Naive consumers are aware of the beliefs of the sophisticated consumers, and also believe that sophisticated consumers have the return cost distribution $F(r)$. However, they disagree with the sophisticated consumers in that they believe that their own return cost distribution is $\tilde{F}(r)$. Furthermore, both consumer types think that the monopolist's belief about the cost distribution is the same as their own. This therefore results in two "agreements to disagree", one between both consumer types, and one between the naive consumers and the monopolist. Such agreements to disagree are present in several exploitative contracting models, see for example Eliaz and Spiegler (2006).

Note that if a naive consumer believes that also sophisticated consumers will buy the good for a w with $w^N \leq w < w^S$, a naive consumer will only believe that the good is of high quality if the fraction of naive consumers is sufficiently high compared to the fraction of sophisticated consumers, because the naive ones know that at such a warranty level the sophisticates believe that the good is of low quality. Recall that the naive consumers (falsely) think that the monopolist shares the same belief about the true naive consumers return cost cdf, $\tilde{F}$. Under the assumption that these are the true return costs, it would be unprofitable for a firm to set a low quality level given a warranty level bigger or equal than w^N when only selling to naive consumers.

Therefore, given a contract γ from the monopolist, a naive consumer's belief $\tilde{\mu}$ will be of the following form:

$$\tilde{\mu}(\gamma) = \begin{cases} 1, & \text{if } w \geq w^S \\ 1, & \text{if } w^N \leq w < w^S \text{ and } U(\mu(\gamma), \gamma) > 0 \text{ and Condition 1 is satisfied} \\ 0, & \text{if } w^N \leq w < w^S \text{ and } U(\mu(\gamma), \gamma) > 0 \text{ and Condition 1 is not satisfied} \\ 0, & \text{if } w < w^N, \end{cases} \tag{3.26}$$

where Condition 1 says that given demand $D + \tilde{D}$,

$$\begin{gathered}(D + \tilde{D})[p - (1 - q_H)(DF(w)w + \tilde{D}\tilde{F}(w)w) - c_H] \geq \\ (D + \tilde{D})[p - (1 - q_L)(DF(w)w + \tilde{D}\tilde{F}(w)w) - c_L].\end{gathered}$$

Therefore, we again use the convention that if a monopolist is indifferent between producing with high or with low quality, the naive consumers believe that the monopolist is producing with high quality. The first line of equation (3.26) refers to the case for which the warranty level is so high that also sophisticated consumers believe that it is more profitable for the monopolist to produce a high quality good. The second line refers to the case in which sophisticated consumers infer a low quality from the contract offered, but still have a non-negative consumption utility. Since naive consumers have the same and correct belief about the sophisticated consumers' return cost distribution, they also infer that the sophisticated consumers think the good is of low quality. However, given naive consumers' beliefs about their own return cost distribution, when Condition 1 holds they believe that it is still more profitable for the monopolist to produce a high quality good than a low quality good. The third line present the opposite case, while in the fourth line the warranty level is so low that also the naive consumers believe that it is always a dominant strategy for the monopolist to produce a low-quality good.

As in the monopoly case, naive consumers' beliefs do not depend solely on the warranty for the good but also on the perceived demand from the sophisticated consumers. In a case in which a firm sells a product to both sophisticated and naive consumers, naive consumers will only believe that this firm will produce with high quality if this is optimal given the return cost functions for both naive and sophisticated consumers. This is because also naive consumers believe that sophisticated consumers will always believe that the good is of low quality as long as $w < w^S$. Therefore, given the set of contracts Γ from all firms in the market, including contract γ_i for firm i, a naive consumer's belief about firm i being of high quality, $\tilde{\mu}_i$, will be of the following form:

$$\tilde{\mu}_i(\Gamma) = \begin{cases} 1, & \text{if } w_i \geq w^S \\ 1, & \text{if } w^N \leq w_i < w^S \text{ and Condition 2 is satisfied} \\ 0, & \text{if } w_i < w^N \\ 0, & \text{if } w^N \leq w_i < w^S \text{ and Condition 2 is not satisfied,} \end{cases} \tag{3.27}$$

where Condition 2 implies that given demand $D_i + \tilde{D}_i$,

$$\begin{aligned} &(D_i + \tilde{D}_i)[p_i - (1 - q_H)(D_i F(w_i) w_i + \tilde{D}_i \tilde{F}(w_i) w_i) - c_H] \geq \\ &\quad (D_i + \tilde{D}_i)[p_i - (1 - q_L)(D_i F(w_i) w_i + \tilde{D}_i \tilde{F}(w_i) w_i) - c_L]. \end{aligned}$$

Analogous to Condition 1 in the monopoly case, Condition 2 states that given firm i's demand from both sophisticated and naive consumers, from a naive consumer's point of view it is more profitable for firm i to produce a high quality good than to produce a low quality good.

Chapter 4

Persuasive Advertising and Cooling-Off Laws under Non-Standard Beliefs

This chapter analyzes pricing and welfare in a model in which advertising is able to influence a state-dependent psychological bias that consumers experience when making purchasing decisions. Consumers attribute too much weight to their current psychological state when making a consumption decision for the future, which amounts to a "projection bias". We focus on how the bias distribution within the population affects the market outcome both under monopoly and competition. Unlike most of the persuasive advertising literature, the model allows for an unambiguous welfare evaluation. We assess consumer protection policies in terms of mandatory cooling-off periods. Our results show that if the share of initially motivated consumers is sufficiently low, such protection policies will decrease both overall and consumer welfare. For higher fractions of motivated consumers, the two welfare standards can give opposing assessments.[1] [2]

[1]I would like to thank Jana Friedrichsen, Paul Heidhues, Volker Nocke, Matthew Rabin, Xiaojian Zhao, as well as seminar participants at Mannheim and at the Maastricht Behavioral and Experimental Economics Symposium 2011 for helpful comments.

[2]This chapter extends my own work of Michel (2009) and Michel (2007). Several sections of this chapter, especially the introduction, the main model, the monopoly section, the duopoly section, and the welfare section largely include similar or identical parts compared to my previous work. However, there are also crucial differences compared to my previous work. First, previous work does not focus on one of the two key questions in this chapter, namely the consequences of mandatory cooling-off policies on consumer and overall welfare. Second, section 4.6, including Propositions 4.5 - 4.7 are not included either. Furthermore, the chapter contains several new graphs, tables, and additional motivation.

4.1 Introduction

It is common wisdom that one of the basic means of advertising is to persuade potential customers to buy goods. Although many advertisements inform customers about product features, a large fraction mainly tries to create emotional links between the good and the costumer. Research from cognitive psychology has shown that people who are in a specific emotional state tend to remember memories experienced in a such state better than those experienced in a different state. A related concept, projection bias, was introduced into the behavioral economics literature by Loewenstein et al. (2003). When an agent experiences different states over time, for example being hungry or satiated, motivated or unmotivated, this bias makes her attribute too much weight on her current state when predicting her future consumption utility. Thus, she overpredicts the likelihood of the current state in the future.

In this chapter, we model the effects of persuasive advertising as a change in the bias consumers experience when predicting their future consumption utility of a good. In our model, consumers are susceptible towards a projection bias. Advertising affects their decision-making process by influencing the magnitude of the bias. This has two effects: First, consumers who are unmotivated to buy a good and therefore underpredict their future expected consumption utility have an increased probability of recalling favorable memories. This in turn decreases their projection bias. Second, motivated consumers attribute an even higher probability than before on being motivated in the future. Advertising therefore further distorts their predicted future consumption utilities and enhances their bias.

We want to answer two main questions. The first question is how equilibrium prices, advertising intensities, and welfare are affected by the state and bias distributions in the population. There are many instances in which consumers' perceptions with respect to a good temporarily change due to a specific incident. Examples are the unexpected success of a domestic athlete in a sport, increased consumer insecurity because of new phenomena such as mad cow disease for beef consumption, or a rapid change in weather conditions.[3]. In this case, our interest lies in how such incidents affect a firm's strategy The second question is how consumer protection policies in the form of mandatory cooling-off periods affect overall welfare. Such policies are often advised to prevent consumers from making impulsive buying decisions. A convenient aspect of the model is that, unlike in most other persuasive advertising models, a consumer's true consumption utility is does not change due to the advertising intensity. This leads to an unambiguous welfare evaluation, independent of using an advertising or non-advertising welfare standard.

Our results indicate that the effects of advertising on pricing and consumer welfare under monopoly crucially depend on whether advertising induces the monopolist to target new consumer types. If a monopolist sells only to motivated consumers without advertising, but also includes unmotivated consumers under advertising, this decreases market prices and increases consumer welfare. In any other case advertising causes an increase in prices and a decrease in consumer welfare. From an overall welfare standpoint, advertising can increase welfare also in cases in which only motivated consumers buy the good both with and without advertising.
The effects of cooling-off periods depend on the fraction of initially motivated consumers in the market. We show that whenever the fraction of motivated consumers in the population is sufficiently low, such policies decrease both overall and consumer welfare. A mandatory cooling-off law then causes a welfare loss due to the foregone utility of not consuming the good in the cooling-off period. If the fraction of motivated consumers is sufficiently high, this is sufficient for the policy to increase consumer welfare, because it makes a monopolist change his optimal strategy from targeting only motivated consumers to targeting all consumers. For intermediate fractions of motivated consumers in the industry, overall welfare may still decrease if the foregone consumption utility of motivated consumers exceeds the additional consumption utility of unmotivated consumers.

As an intuition for our model, consider the following example. Imagine a TV network that owns the license for showing live games of a professional sports league on pay-TV

[3]Conlin et al. (2007) find evidence for a projection bias in catalogue order after abrupt weather changes.

or over the Internet. It offers contracts to consumers; in this example only a season ticket contract is offered. In addition, the network also runs advertisements on different channels to promote subscriptions. Let us further assume that, at the beginning of the season, people are either motivated about buying the good or skeptical about it. In a motivated state, consumers always like to watch sports and think that it would be a good choice to buy the good even before it is advertised. Skeptical consumers, on the other hand, are afraid of paying too much for a season ticket and have a lower valuation for a season ticket than motivated consumers. It seems natural to us that the firm uses exciting scenes from earlier seasons in order to motivate potential consumers. When confronted with such scenes, viewers think that it is more likely that they will appreciate the good. This increases their predicted consumption utility when buying the season ticket. There are also real life examples for such behavior: Sport networks such as NBA TV often show some teaser highlights when advertising their season tickets.

We believe that motivated consumers further underestimate the possibilities of boring games or of a satiation effect after having watched several games. Consequently, they are willing to buy a good under these circumstances for a price for which they would not buy the good if they were fully rational.Skeptical consumers are reminded of previous sport events they liked as well, which decreases their degree of pessimism with respect to the purchase of the good. A monopolist has to decide whether a price should be set that only attracts the motivated consumers or whether the skeptical consumers should also be attracted by setting a lower price. The ratio of motivated to unmotivated consumers is crucial for this decision.

Loewenstein et al. (2003) were the first to explicitly define the notion of projection bias and also give the formal basis for our model. Their concept focuses on a bias that consumers experience when predicting their utility of consuming a good in a different state in the future. This projection bias occurs because people are said to rely partly on their current state of mind when making a prediction about a future state. Since their predicted utility differs from the expected utility of a rational consumer, this can lead to non-optimal behavior. As a formal illustration of the projection bias, Loewenstein et al. (2003) present a model in which consumers decide whether or not to buy a durable good. Consumers exhibit day-to-day fluctuations with respect to the valuation of a good they can consume in several periods. When in a state with a relatively high valuation, a consumer overpredicts his possible consumption utility. Consequently, the consumer would buy the good even for a price so high as to have a negative expected utility without a bias. At the same time, when being in a state with a low valuation, he would not want to buy the good for a price that would give him a positive expected consumption utility when making rational predictions about the future. Conlin et al. (2007) use a structural model in order to estimate the magnitude of a potential projection bias with respect to weather changes when consumers make catalogue orders for clothing. They find a significant projection bias when consumers are predicting their future tastes in this case.

The work on emotion and cognition in the branch of the cognitive psychology literature can give further intuitions for the existence of a projection bias. In our context, the concepts of mood-state dependent retrieval and mood congruity are of particular interest. Mood congruity "describes the case when people in a good mood remember emotionally positive material better than those in a bad mood, whereas the opposite is true for emotionally negative material"(Eysenck and Keane 1993, p. 443.). If a person finds herself in a happy mood, for example, under mood congruity she is thus said to remember emotionally positive characteristics better than negative ones. "Memory is said to be mood-state dependent in case the memories that subjects store when they are in one emotional state are more retrievable later if they re-enter that same emotional state; and their recall is worse if they attempt recall in a different emotional state from original learning"(Bower 1992, p.22). Braun (1999) analyzes the effects of post-experience advertising on consumer memory. She finds that post-experience advertising can make memories about a product experience more appealing. Also negative experiences are perceived more favorably.

There is a large and diverse literature on the different effects of advertising, see Bagwell (2007) for a broad overview of the advertising literature. As Bagwell points out, empirical evidence suggests that no theoretical approach seems to work in all cases. In Bloch and Manceau (1999), consumers' tastes are non-uniformly differentiated on a line and firms are located at the ends of the line. Bloch and Manceau explore cases where either both firms are owned by a multi-product monopolist or by competing firms. At most one firm is allowed to advertise in their model. If a firm advertises, this causes a shift in the distribution of consumers' tastes; these then move closer to the advertising firm. For the class of log-concave distributions, Bloch and Manceau show that a multi-product monopolist has an incentive to advertise the more favored product to generate an even more biased distribution. Bernheim and Rangel (2004) provide a model in which agents face stochastically varying environmental impulses over time that influence the propensity to consume an addictive substance. In a "cold" mode, agents choose according to their true preferences, while in a "hot" mode, their brain processes suffer from a distorted forecast mechanism. In a dynamic programming framework, such a distortion can lead to a consumption of the addictive substance which in turn increases the probability of being in a hot mode in the future.

This chapter also relates to a small literature of marketing models with a behavioral economics foundation. Ho et al. (2006) discuss how several behavioral economics concepts, such as reference dependence and hyperbolic discounting, can be implemented in a marketing context.

Section 4.2 presents the baseline model. The monopoly and oligopoly outcomes are analyzed in section 4.3 and 4.4, respectively. In section 4.5, we present welfare evaluations, and analyze a mandatory cooling-off periods in section 4.6. Section 4.7 concludes.

4.2 The Model

A monopolist is located at one point on a line of infinite length and offers a single good to potential consumers, whose tastes are uniformly distributed on the line. Consumers first make the decision of whether or not to buy a good and consume it at a later point. This applies to nearly all forms of both internet and catalogue shopping of commodity goods. We assume that there is no depreciation of the good between these two periods, i.e. we have a unit discount rate. Furthermore, each consumer either no or exactly one good. Each consumer has two potential states. We refer to these as being motivated about a good, $\overline{o}$, and being unmotivated about a good, $\underline{o}$. Depending on the states, the associated utilities of consuming an ideal product c, i.e. one that perfectly matches a consumer's tastes, is $u(c, \overline{o})$ for a motivated consumer, and $u(c, \underline{o})$ for an unmotivated consumer. A consumer's consumption utility is higher in a motivated state than in an unmotivated state, $u(c, \overline{o}) > u(c, \underline{o}) > 0$. The utility from not consuming the good is set to 0 for both states of the consumer. This reflects the consumption of another good as an outside option for which the consumer does not experience a projection bias in predicting her consumption utility. Consumers are aware of the fluctuations concerning their valuations of a good. Hence, they form predictions about how likely it is that they will be motivated when consuming the good in a later period. A rational consumer assigns a probability $q \in (0,1)$ to the event of being motivated when consuming the good, and a probability $1-q$ to the event of being unmotivated when consuming it. The expected utility of consuming an ideal product is thus the weighted sum of the utilities in the two states, namely being motivated and being unmotivated. This can be written as

$$u(c) = qu(c, \overline{o}) + (1-q)u(c, \underline{o}), \quad q \in (0,1). \tag{4.1}$$

Note that when a consumer has no bias, the expected future consumption utility is independent of the state in which she makes her buying decision. A bias occurs in our model when consumers make their buying decisions. In such a situation a consumer finds herself either in a motivated or in an unmotivated state. Motivated consumers overpredict the utility of buying the good whereas unmotivated consumers underpredict it. Loewenstein et al. (2003) refer to such time-inconsistencies as a projection bias. Let $\tilde{u}(c|o)$ denote the predicted utility of consuming a good in the consumption period while being in the state o in the buying period. This results in the following definition of a projection bias:

Definition (Projection Bias) 4.1. *Predicted utility exhibits a projection bias if there exists $\alpha \in [0,1]$ such that for all c,o: $\tilde{u}(c|o) = (1-\alpha)u(c) + \alpha u(c,o)$.*

Note that this definition extends the one of Loewenstein et al. (2003).[4] The difference is that in their model, Loewenstein et al. deal with the predicted utility of a single state, whereas we incorporate agents who predict their expected utility of a combination of states. We believe that our specification better describes the decision-making processes of many markets, because it reflects a possible uncertainty in the consumption utility. If a consumer buys a new computer game, for example, she does not know in advance how much she will enjoy playing it.

In period 1, consumers find themselves in the motivated state with probability $\theta \in$

[4] In particular, Loewenstein et al. (2003) define a projection bias in the following way: "Predicted utility exhibits simple projection bias if there exists $\alpha \in [0,1]$ such that for all $c,\ s,\ and\ s',\ \tilde{u}(c,s|s') = (1-\alpha)u(c,s) + \alpha u(c,s')$."

$(0,1)$, and in the unmotivated state with probability $1-\theta$. The probability θ in our model is independent of the probability q of having a high valuation for the good in the second period. This enables us to explore cases in which certain outside events might cause systematic differences between the probabilities of being in certain states in different periods. For example, during Olympic games people might be more motivated to buy sports equipment than before or after the event. In our model, $\theta > q$ reflects such a case.

We now introduce our advertising concept. We assume that advertising increases the positive projection bias for motivated consumers and decreases the negative bias for unmotivated consumers. One intuition for these assumptions is as follows: If advertising increases the emotional attachment to a good, a key purpose of advertising, people are supposed to value the good more highly and are consequently also willing to pay more for its consumption. We believe that work from cognitive psychology give an explanation for a projection bias: Due to mood-state dependent recall, people build their expectations based on state-correlated probabilities.

For simplicity reasons, we restrict our analysis to only two advertising intensities, $\psi \in \{0,1\}$. We distinguish between the bias of a motivated consumer, $\overline{\alpha}(\psi)$, and the bias of an unmotivated consumer, $\underline{\alpha}(\psi)$. Advertising tries to emotionally attach the consumers to their goods by creating a higher emotional captivity associated with positive or sometimes also with scary moods. Since advertising only gives positive associations about the good, this increases the memory selectivity in the case of a motivated consumer and reduces the selectivity towards negative thoughts of an unmotivated consumer. Thus, $\overline{\alpha}(1) > \overline{\alpha}(0)$, and $\underline{\alpha}(1) < \underline{\alpha}(0)$.

The price p of the monopolist's good, as well as the distance of the good's product designs, i, from the consumer's original preferences , x, affect her utility negatively in a linear way. In the following, we normalize the monopolist's location to the point $i = 0$ on the line. The consumption utility of a consumer located at $x \in (-\infty, \infty)$ who is in state o for the good, $\tilde{U}_x(c|o)$, thus becomes

$$\tilde{U}_x(c|o) = \tilde{u}(c|o) - p_i - t|x|; \quad t > 0. \tag{4.2}$$

$\tilde{u}(c|o)$ is the perceived expected utility for consuming her ideal variety of the good produced by the monopolist when in state o, and t is a differentiation parameter. Note

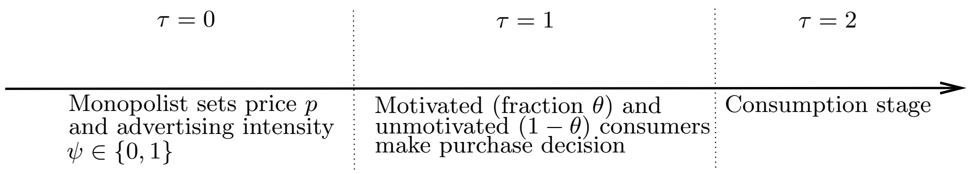


Figure 4.1: Structure of the game

that the notion "ideal product variety" refers to the horizontal differentiation in our market. In that way, if a product differs from a consumer's ideal variety, she ends up having a lower utility than for consuming a product that matches her ideal preferences and has the same price. This implies that the effect of advertising on predicted utility is independent of the consumer's horizontal preferences, i.e. her location on the line. The reservation utility $\overline{U}$ of not consuming a good is 0: $\overline{U}(0|o) = 0 \; \forall o$.
Given a consumer's upward bias function $\overline{\alpha}(\psi)$, the definition for a consumer's predicted consumption utility in state $\overline{o}$ can then be rewritten as

$$\tilde{u}(c|\overline{o}) = (1 - \overline{\alpha}(\psi))u(c) + \overline{\alpha}(\psi)u(c, \overline{o}) = (1-q)\overline{\alpha}(\psi)(u(c, \overline{o}) - u(c, \underline{o})) + u(c). \quad (4.3)$$

The first term, $\overline{\alpha}(\psi)(1-q)(u(c,\overline{o}) - u(c,\underline{o}))$, can be interpreted as the additional predicted utility attributable to the upward bias. This markup consists of two components; the upward-bias $\overline{\alpha}(\psi)$ itself, which varies with the advertising intensity ψ; and a fixed upward utility component for a given probability q, $(1-q)(u(c,\overline{o}) - u(c,\underline{o}))$. We define this component as $\overline{V}(q) \equiv (1-q)(u(c,\overline{o}) - u(c,\underline{o}))$. Therefore, for a motivated consumer, we can rewrite the predicted utility as

$$\tilde{u}(c|\overline{o}) = \overline{\alpha}(\psi)\overline{V}(q) + u(c). \quad (4.4)$$

Similarly, for the unmotivated consumers, the definition of the utility function under projection bias can be rewritten as $\tilde{u}(c|\underline{o}) = u(c) + \underline{\alpha}(\psi)q[u(c,\underline{o}) - u(c,\overline{o})]$, where $q[u(c,\underline{o}) - u(c,\overline{o})]$ denotes the fixed downward utility component, and $\underline{\alpha}(\psi)$ the variable downward bias. Define the fixed downward utility as $\underline{V}(q) \equiv q(u(c,\underline{o}) - u(c,\overline{o}))$. This gives us another expression for the predicted utility for unmotivated consumers:

$$\tilde{u}(c|\underline{o}) = \underline{\alpha}(\psi)\underline{V}(q) + u(c). \quad (4.5)$$

We can rewrite the definitions for the predicted utility in each state to emphasize consumers mispredicting the actual probabilities of being in a certain state in the future. Define $\tilde{q} \equiv (1-\alpha)q + \alpha > q$, and $\hat{q} \equiv (1-\alpha)q < q$. Then we obtain the following expression for the predicted utilities of motivated and unmotivated consumers, respectively:

$$\tilde{u}(c|\overline{o}) = \tilde{q}u(c,\overline{o}) + (1-\tilde{q})u(c,\underline{o}); \quad \tilde{u}(c|\underline{o}) = \hat{q}u(c,\overline{o}) + (1-\hat{q})u(c,\underline{o}). \tag{4.6}$$

We assume that production of the good is free, but positive advertising, $\psi = 1$, incurs a cost $s(\psi)$: Therefore, $s(0) = 0$ and $s(1) > 0$.
We assume that consumers are unaware of their projetion bias. Therefore, they take their predicted future consumption utility for their unbiased expected future consumption utility. This then leads to consumers maximizing their their predicted utility in the buying period. The monopolist is aware of consumers suffering from projection bias, and maximizes his profits while anticipating consumers' behavior correctly. We will now focus on the monopolist's optimal decision-making.

4.3 Monopoly Analysis

The monopolist's decision problem consists of setting the optimal price p and advertising intensity ψ as to maximize his profits.
First, we derive the monopolist's demand function. The monopolist's total demand, $D(p,\psi)$, is the sum of the demand of motivated consumers, $\overline{D}(p,\psi)$, and unmotivated consumers, $\underline{D}(p,\psi)$: $D(p,\psi) = \overline{D}(p,\psi) + \underline{D}(p,\psi)$. The demand of motivated consumers is

$$\overline{D}(p,\psi) = 2\theta \max\left\{\frac{\overline{\alpha}(\psi)\overline{V} + u(c) - p}{t}, 0\right\}.$$

Similarly, the demand of unmotivated consumers is

$$\underline{D}(p,\psi) = 2(1-\theta) \max\left\{\frac{\underline{\alpha}(\psi)\underline{V} + u(c) - p}{t}, 0\right\}.$$

Let $\overline{x}(p,\psi) \equiv max\{0, \frac{\overline{\alpha}(\psi)\overline{V}+u(c)-p}{t}\}$ be the distance from the monopolist to the most distant motivated consumers who purchase a good, and let $\underline{x}(p,\psi) \equiv max\{0, \frac{\underline{\alpha}(\psi)\underline{V}+u(c)-p}{t}\}$ be the distance from the monopolist to the most distant unmotivated consumers who purchase a good, respectively. This yields $\overline{D}(p,\psi) =$

$2\theta\overline{x}(p,\psi)$, and $\underline{D}(p,\psi) = 2(1-\theta)\underline{x}(p,\psi)$ as expressions for demand. The firms' maximization problem can be expressed as

$$\max_{p,\psi} \Pi(p,\psi) = D(p,\psi)p - s(\psi).$$

The monopolist's demand function has a kink because of the different consumer states in the buying period. Therefore, there exist two equilibrium candidates. Based on the fraction of motivated consumers, θ, the monopolist has to decide whether to set a price so high as to only attract motivated consumers or whether to set a price that also attracts some unmotivated consumers.

Monopolist sells to both motivated and unmotivated consumers A candidate equilibrium price in this segment has to lie within the price range $p \in [0, \underline{p})$, where $\underline{p} \equiv \underline{\alpha}(\psi)\underline{V} + u(c)$. Under this condition, differentiating the monopolist's profit function with respect to p yields the first-order condition

$$\frac{\partial \Pi}{\partial p} = \frac{2}{t}[\theta\overline{\alpha}\overline{V} + (1-\theta)\underline{\alpha}\underline{V} + u(c) - 2p] = 0.$$

From the above equation, it follows that the price when selling to both consumer types is

$$p = \frac{1}{2}[u(c) + \theta(\overline{\alpha}(\psi^{*All})\overline{V} + (1-\theta)\underline{\alpha}(\psi^{*All})\underline{V})]. \tag{4.7}$$

Necessary, but not sufficient, for this price to be optimal is that, given the price p and advertising intensity ψ^{*All}, the demand of unmotivated consumers, $\underline{D}(p,\psi^{*All})$, is positive. This will only be the case if $p < \underline{\alpha}(\psi^{*All})\underline{V}$, which can be rewritten as $\theta < \frac{u(c)+\underline{\alpha}(\psi^{*All})\underline{V}}{\overline{\alpha}(\psi^{*All})\overline{V}-\underline{\alpha}(\psi^{*All})\underline{V}}$. When targeting both consumer types, the monopolist will optimally set the advertising intensity according to the following rule:

$$\psi^{*All} = \begin{cases} 0, & \text{if } [\theta\overline{\alpha}(1)\overline{V} + (1-\theta)\underline{\alpha}(1)\underline{V} + u(c)]^2 - [\theta\overline{\alpha}(0)\overline{V} + (1-\theta)\underline{\alpha}(0)\underline{V} + u(c)]^2 \le 2ts(1) \\ 1, & \text{if } [\theta\overline{\alpha}(1)\overline{V} + (1-\theta)\underline{\alpha}(1)\underline{V} + u(c)]^2 - [\theta\overline{\alpha}(0)\overline{V} + (1-\theta)\underline{\alpha}(0)\underline{V} + u(c)]^2 \ge 2ts(1) \end{cases}$$

In such a case, the monopolist's profit depends only on the advertising intensity ψ^{*All}, and can be expressed as

$$\Pi^{All}(\psi^{*All}) = \frac{1}{2t}[\theta\overline{\alpha}(\psi^{*All})\overline{V} + (1-\theta)\underline{\alpha}(\psi^{*All})\underline{V} + u(c)]^2 - s(\psi^{*All}). \tag{4.8}$$

Monopolist only sells to motivated consumers Intuitively, such a case can only be optimal if, for a given price p and advertising intensity ψ, no unmotivated consumer but at least some motivated consumers demand the good, i.e. if $p \in \mathtt{P} \equiv [u(c) + \underline{\alpha}(\psi)\underline{V}, u(c) + \overline{\alpha}\overline{V})$. Clearly, each price equal or above $\overline{p} \equiv u(c) + \overline{\alpha}(\psi)\overline{V}$ cannot be optimal, for it always yields the monopolist a non-positive profit, which cannot be optimal, given that a lower price leads to a positive profit. Therefore, for a price in this interval, differentiating the monopolist's profit function with respect to p yields

$$\frac{\partial \Pi}{\partial p} = \frac{2\theta}{t}[u(c) + \overline{\alpha}\overline{V} - 2p] = 0.$$

We denote the monopolist's optimal advertising intensity in this segment by ψ^{*Mot}. The optimal advertising level depends on the relationship between advertising costs and increase in consumer willingness to pay due to advertising. It is defined by

$$\psi^{*Mot} = \begin{cases} 0, & \text{if } \theta[(\overline{\alpha}(1) + \overline{\alpha}(0))\overline{V} + u(c)][(\overline{\alpha}(1) + \overline{\alpha}(0))\overline{V}] \leq 2ts(1) \\ 1, & \text{if } \theta[(\overline{\alpha}(1) + \overline{\alpha}(0))\overline{V} + u(c)][(\overline{\alpha}(1) + \overline{\alpha}(0))\overline{V}] \geq 2ts(1). \end{cases}$$

The monopolist's optimal price in this case is given by

$$p = \frac{1}{2}[u(c) + \overline{\alpha}(\psi^{*Mot})\overline{V}], \tag{4.9}$$

The monopolist's profit in this segment, Π^{Mot}, as a function of ψ^{*Mot} while solving for the optimal price, can be expressed as

$$\Pi^{Mot}(\psi^{*Mot}) = \frac{\theta}{2t}[\overline{\alpha}(\psi^{*Mot})\overline{V} + u(c)]^2 - s(\psi^{*Mot}). \tag{4.10}$$

Given the specific bias functions, whether or not to sell a good only to motivated consumers depends on the fraction of motivated consumers in the market in the first period. For our welfare analysis, it will be helpful to focus on the conditions under which it is best for the monopolist to sell also to unmotivated consumers both under advertising and when advertising is not feasible. We can show as a benchmark that if advertising is not feasible, there exists a unique threshold level θ for which a monopolist is indifferent between attracting both consumer types and attracting motivated consumers alone. When advertising is feasible, there is no clear threshold level. This is because the optimal advertising intensity for the monopolist differs when attracting only motivated consumers or attracting both motivated and unmotivated consumers. From equations (4.8) and (4.10) we obtain the expressions for a firm's optimal profit, given the optimal advertising intensities ψ^{*Mot} and ψ^{*All}. The monopolist will only attract both consumer types whenever $\frac{1}{2t}[\theta\overline{\alpha}(\psi^{*All})\overline{V} + (1 -$

$\theta)\underline{\alpha}(\psi^{*All})\underline{V} + u(c)]^2 - s(\psi^{*All}) \geq \frac{\theta}{2t}[\overline{\alpha}(\psi^{*Mot})\overline{V} + u(c)]^2 - s(\psi^{*Mot})$. Lemma 4.1 sums up these results.

Lemma 4.1. *When advertising is feasible, a monopolist will set a price to attract both consumer types whenever* $\Pi^{All}(\psi^{*All}) \geq \Pi^{Mot}(\psi^{*Mot})$.
When advertising is not feasible, a monopolist will set a price as to attract both consumer types, if and only if $\theta \leq \left(\frac{\underline{\alpha}(0)\underline{V}+u(c)}{\overline{\alpha}(0)\overline{V}-\underline{\alpha}(0)\underline{V}}\right)^2$ *and* $\underline{\alpha}(0)\underline{V} + u(c) > 0$.

Proof: See Appendix.
Intuitively, when advertising increases unmotivated consumers' demand sufficiently more than motivated consumers' demand, this can lead to circumstances for which the monopolist sells to both consumer types under advertising and only to motivated consumers when advertising is not feasible. Advertising can also cause a fall in the monopoly price in this case. Empirical evidence on the relationship between price and advertising is ambiguous, but in some studies this type of inverse relationship is documented, see for example Bagwell (2007), pp. 1743-1746. Our model allows both for a positive and a negative correlation between monopoly price and advertising, which differs from some models in the literature, see for example Dixit and Norman (1978), where the price is assumed to be weakly increasing in the advertising level. The next section introduces and analyzes a duopoly variant of the model.

4.4 Duopoly analysis

We present a competitive model in which two competitors, firms A and B, are located at the two ends of a line of length one, on which consumers tastes are uniformly distributed. In particular, we assume that firm A is located at $x_A = 0$ and firm B is located at $x_B = 1$.
Many advertising campaigns begin long before a new product is sold. Most of the time, persuasive advertising does not focus on prices, but rather conveys "soft" information (Nelson 1974). TV ads often do not contain information about prices. When dealing with competing firms in the market, we assume that firms set their advertising intensities first and set prices later. The advertising intensity of one firm is perfectly observable for the other firm before the prices are set. This results in a two-stage game. Firms first set advertising intensities independently of each other and in the

second stage, after observing the competitor's advertising intensities, set prices non-cooperatively. We assume that the advertising perception of one brand does not affect the advertising perception of another one, so that there are no advertising externalities Our setting also allows us to analyze cases in which both firms have different bias functions. This reflects for example situations in which one brand has either a higher initial reputation or a better marketing firm.[5] Throughout our analysis, we however restrict our focus to cases in which both firms have the same advertising functions. In particular, we assume that $\overline{\alpha}_A(\psi) = \overline{\alpha}_B(\psi)$; $\underline{\alpha}_A(\psi) = \underline{\alpha}_B(\psi),\ \psi \in \{0,1\}$.
For our equilibrium concept, we first require, as in the monopoly case, that consumers maximize their predicted utility in the buying period. Second, firms anticipate both consumers' and their competitor's behavior correctly. Given these behaviors, firms maximize their expected profits in every subgame.[6] The associated profit function for firm i is

$$\Pi_i(p_i, p_j, \psi_i, \psi_j) = D_i(p_i, p_j, \psi_i, \psi_j)p_i - s_i(\psi_i). \tag{4.11}$$

To find the equilibria of the game, we will solve the game via backward induction, starting with finding firms' optimal prices for any possible advertising history (ψ_A, ψ_B). The next assumption excludes cases in which advertising is unprofitable for both firms and gives an upper bound for the firms' advertising costs.

Assumption 4.1. $\frac{t}{2} > s(1);\quad min\{(\underline{\alpha}_i(1) - \underline{\alpha}_i(0))\underline{V}, (\overline{\alpha}_i(1) - \overline{\alpha}_i(0))\overline{V}\} > \frac{t}{2} + s(1).$

The first part of the assumption gives an upper bound on the degree of horizontal differentiation of tastes among consumers. As we will see later, this implies that firms can coexist in an equilibrium in which they both advertise under the requirement that the unmotivated consumers' willingness to pay is sufficiently high. Under the same requirement, the second part of the assumption assures that at least one firm will invest in advertising. Since firms are located at the two ends of the line, the firms' profit functions in the pricing subgame are always be continuous, assures the existence of an equilibrium.

[5]In their functional magnetic resonance imaging (fmri) experiment, McClure et al. (2004) found out that the "brand-cued" delivery of Coca-Cola significantly increased participants' brain activity compared to the anonymous delivery. This was not the case for Pepsi-Cola or no-name cola. In our model, this would imply that some firms can better influence consumers than others through advertising. If consumers have such a subconscious pre-ordering of firms, it would be likely that advertising of highly appreciated brands also increases the consumption utility of a good.

[6]Our equilibrium concept is close to the notion of subgame perfect equilibrium. However, this notion does not apply in our case, since consumers have a projection bias.

In the next subsection we explore under which circumstances a symmetric equilibrium exists.

4.4.1 Symmetric equilibria

Fully served market First, we present the case when $\underline{\alpha}_i(1)\underline{V}+u(c) \geq \frac{3t}{2}$. This implies that the differentiation in the market is relatively low compared to the unmotivated consumers' willingness to pay for the two goods without advertising. When setting the same advertising intensities in the first stage, both firms will target all consumer types and set the same prices. Whether they set the same advertising intensities depends on the equilibrium profits in the subgame after they have set different advertising intensities. This is our next proposition.

Proposition 4.1. *Suppose $\underline{\alpha}_i(1)\underline{V} + u(c) \geq \frac{3t}{2}$. If in the pricing subgame following history $(\psi_A, \psi_B) = (1, 0)$ firm A's equilibrium profits are less than $\frac{t}{2} - s(1)$, then there exists a symmetric equilibrium in which $p_A = p_B = t$ and $\psi_A = \psi_B = 1$.*

Proof: See Appendix.
This is a relatively standard result. A firm's reaction function is increasing in the other firm's price, and the only time both firms' reaction functions intersect is at $p_A = p_B = t$. We now show that there are circumstances in which only motivated consumers are served also in the duopoly case when advertising is not allowed.

Equilibria in which only motivated consumers are served when advertising is not allowed When considering welfare effects of advertising, as we will do in detail in the next section, it is important to analyze the equilibrium outcomes when advertising is not allowed. In such a case, since two different consumer types are present in our model. It is especially interesting for us to discuss what happens if unmotivated consumers only have a small willingness to pay for a good relative to the differentiation of tastes in the market. If $\underline{\alpha}(0)\underline{V} + u(c)$ is sufficiently smaller than the differentiation parameter t, then there exist equilibria in which both firms set prices so that no unmotivated consumer will buy a good. We have shown in Proposition 4.2 that if the market is fully served, a symmetric equilibrium will yield prices $p_A = p_B = t$. If $\overline{\alpha}(0)\overline{V} + u(c) \geq \frac{3t}{2}$. if advertising costs are sufficiently low, then there are circumstances in which both firms still set $p_A = p_B = t$ and thus attract motivated consumers only. If the willingness to pay for the unmotivated consumers is sufficiently low, a firm would have to cut its price too much to also attract unmotivated consumers for this to be a profitable deviation. Recall that in the symmetric equilibrium where the market is fully served, a firm's reaction function is increasing in the other firm's price. In case only motivated consumers are served after the symmetric advertising history $(\psi_A, \psi_B) = (0, 0)$, the reaction functions still cut at $p_A = p_B = t$. The proof of the existence of such equilibria is given in the Appendix.

The next subsection deals with the possibility of equilibria with asymmetric advertising strategies.

4.4.2 Asymmetric equilibria

The last subsection has shown that even if the willingness to pay for motivated consumers is extremely high, in a symmetric advertising equilibrium prices will never be above t. However, under some circumstances there exists an equilibrium in which firms set different advertising intensities. In our two-stage game, advertising can also be considered as having a commitment value. Since the firms can see each other's advertising intensities in the pricing stage, their advertising decisions thereby signal their further optimal actions in the market. We show in the Appendix that at least a mixed strategy equilibrium exists in the subgame after firms set asymmetric prices. Because of the complementary effects in prices, advertising asymmetries can lead to higher prices for both consumer types. In some cases a full market segmentation is even possible, in which motivated consumers are served exclusively by the advertising firm, and unmotivated consumers are served by the other firm. As an intuition for this, one can think of a situation in which advertising increases the motivated consumers' willingness to pay so dramatically that a firm that does not advertise can barely attract these consumers even when setting a very low price. It can be optimal for this firm to act as a monopolist for the unmotivated consumers, given that it is optimal for the advertising firm to only focus on motivated consumers.

The next proposition gives us the conditions for an asymmetric equilibrium:

Proposition 4.2. *Suppose $\underline{\alpha}_i(1)\underline{V} + u(c) \geq \frac{3t}{2}$. If in the pricing subgame following history $(\psi_A, \psi_B) = (1, 0)$, firm 1 will make profits of at least $\frac{t}{2} - s(1)$, then there exists an equilibrium in which the two firms set different advertising intensities.*

Proof: See Appendix.

In the next section, we will focus on welfare properties of our model.

4.5 Welfare Analysis

Common problems regarding welfare evaluations in persuasive advertising models

When assessing welfare effects in models including persuasive advertising, these models usually have the problem that advertising induces a change in consumers' tastes. Dixit and Norman (1978) argue that in models with such a taste change, one has to use the same standard, i.e. either pre-advertising tastes or post-advertising tastes, in order to compare the market outcome with and without advertising. They provide a welfare analysis for both pre-advertising and post-advertising standards. In their model, if a marginal increase in advertising is welfare beneficial under post-advertising tastes, this is sufficient for advertising to always be beneficial under the pre-advertising tastes. There are circumstances under which an additional amount of advertising can be considered to be excessive under the former tastes and beneficial under the latter. This implies that one has to restrict to a fix standard in order to be able to make unambiguous welfare evaluations.

The complementarity view of advertising is one way to overcome these problems, see for example Becker and Murphy (1993). Becker and Murphy treat advertising as part of given meta-tastes, such that advertising increases consumption utility when it is a good and decreases utility when it is a bad. Because of the stable preferences, welfare evaluations are unambiguous. Unlike Becker and Murphy's model, in our model consumers have no specific preference for advertising. As an example, this implies that a consumer's utility from watching a sport event on TV does not depend on how much the TV network advertises its broadcasting. The unambiguous welfare evaluation in our model comes at a cost. In many behavioral models like ours, consumers make decisions that on average do not maximize their utility. In such cases, consumers' choices do not necessarily reflect their optimal preferences. The question then arises of whether or not such models can still be used for a normative analysis. There is an ongoing debate relating to this question, for which the different viewpoints in the edition of Caplin and Schotter (2008) can serve as a good guidance. Despite the non-applicability of revealed preferences, we still believe that a welfare analysis in our model will give useful insights.

While consumers' choices do not necessarily reveal their true preferences, they are still optimal given the perceived preferences.

4.5.1 Monopoly welfare

From a normative viewpoint, we are interested in two main questions. The first question is under which circumstances monopoly advertising leads to a higher overall welfare in the market as compared to the case in which advertising is not allowed. The second question is whether or not monopoly advertising is always excessive from a consumer surplus perspective. We define advertising to be excessive if the increase in profits due to advertising exceeds the overall increase in welfare.
Let W denote the overall welfare when advertising is feasible, and let W^N denote the welfare when advertising is not feasible. Advertising is socially desirable if and only if the change in welfare due to advertising is positive: $\Delta W = W - W^N > 0$. Recall $\overline{x}(p,\psi) \equiv \max\left\{0, \frac{\overline{\alpha}(\psi)\overline{V}+u(c)-p}{t}\right\}$ and $\underline{x}(p,\psi) \equiv \max\left\{0, \frac{\underline{\alpha}(\psi)\underline{V}+u(c)-p}{t}\right\}$.
Apart from the expected utility of consuming her ideal good $u(c)$ and the good's price p, the consumer's utility depends on the distance between her own taste x on the line and the monopolist's good located at point 0 in the taste space. Adding those terms for all consumers who purchase a good, we obtain an expression for the consumer surplus CS:

$$CS(p,\psi) = 2\theta \int_0^{\overline{x}(p,\psi)} [u(c) - p - t\tau]d\tau + 2(1-\theta)\int_0^{\underline{x}(p,\psi)} [u(c) - p - t\tau]d\tau. \qquad (4.12)$$

It is straightforward to see that, given an optimal advertising intensity ψ^* and according monopoly price p^* under advertising, and optimal monopoly price p^N without advertising, we can easily check whether advertising increases overall consumer surplus in the market compared to the case when there is no advertising. This is the case whenever $CS(p^*,\psi^*) - CS(p^0, 0) > 0$. The monopolist's profit Π can be written as the sum of motivated and unmotivated consumers who demand the good in the buying period times the selling price p minus the advertising costs $s(\psi)$:

$$\Pi(p,\psi) = 2\left(\theta \int_0^{\overline{x}(p,\psi)} p\ d\tau + (1-\theta)\int_0^{\underline{x}(p,\psi)} p\ d\tau\right) - s(\psi).$$

Adding consumer surplus and the monopolist's profit, we obtain the overall welfare W in the market:

$$W(p,\psi) = CS(p,\psi) + \Pi(p,\psi)$$

$$= 2\left(\theta\int_0^{\overline{x}(p,\psi)}[u(c)-t\tau]d\tau + (1-\theta)\int_0^{\underline{x}(p,\psi)}[u(c)-t\tau]d\tau\right) - s(\psi). \tag{4.13}$$

Inserting the expressions (4.7) and (4.9) for the optimal price for each case, respectively, into $\overline{x}(p,\psi)$ and $\underline{x}(p,\psi)$, we obtain the locations of the most distant consumers in a monopoly as a function of the advertising intensity ψ only. Consequently, we can write overall welfare as a function of the advertising intensity ψ only. However, we still have to distinguish between cases in which the monopolist attracts both consumer types and cases for which he sets a price as to sell to motivated consumers only.

If the monopolist only sells to motivated consumers, this yields the following expression for overall welfare W^{Mot} in the market:

$$W^{Mot}(\psi) = 2\theta\int_0^{\frac{\overline{\alpha}(\psi)\overline{V}+u(c)}{2t}}[u(c)-t\tau]d\tau - s(\psi). \tag{4.14}$$

If the monopolist decides to attract both consumer types, we obtain the following expression for the overall welfare W^{All}

$$W^{All}(\psi) = 2\theta\int_0^{\overline{x}(\psi)}[u(c)-t\tau]d\tau + 2(1-\theta)\int_0^{\underline{x}(\psi)}[u(c)-t\underline{x}]d\tau, \tag{4.15}$$

where

$$\overline{x}(\psi) = \frac{1}{2t}[u(c)+(2-\theta)\overline{\alpha}(\psi)\overline{V}-(1-\theta)\underline{\alpha}(\psi)\underline{V}];\ \underline{x}(\psi) = \frac{1}{2t}[u(c)+(1+\theta)\underline{\alpha}(\psi)\underline{V}-\theta\underline{\alpha}(\psi)\underline{V}].$$

When looking at equations 4.12, 4.14 and 4.15, one can see a big potential difference between a welfare evaluation of advertising and a consumer surplus evaluation of advertising. From an overall welfare perspective, it is desirable that a consumer buys a product, as long as the consumption utility $u(c)$ is bigger or equal than her differentiation costs tx, $u(c) \geq tx$. When investigating how advertising affects consumer surplus, however, we also have to consider how the monopoly price changes due to advertising, which is often more restrictive. On the other hand, it is easy to see that if advertising increases consumer surplus, this will be sufficient for advertising to increase welfare. This is because when advertising is feasible, the monopolist still has the choice not to advertise at all, such that his profit under advertising must be at least as high as his profit without advertising.

We define the distance between the most distant motivated consumer under monopoly advertising and the most distant motivated consumer in a monopoly when advertising is not feasible as $\Delta\overline{x} \equiv \overline{x}(,p^*,\psi^*) - \overline{x}(p^N,0)$. Define the analogous distance for unmotivated consumers as $\Delta\underline{x} \equiv \underline{x}(,p^*,\psi^*) - \underline{x}(p^N,0)$. We can derive conditions for which advertising increases overall welfare, and conditions for which it increases consumer surplus. This results in the following proposition.

Proposition 4.3. *Advertising increases monopoly welfare if and only if*
$\frac{2u(c)}{t}[\theta\Delta\overline{x}+(1-\theta)\Delta\underline{x}] > \theta[\overline{x}(,p^*,\psi^*)^2-\overline{x}(p^N,0)^2]+(1-\theta)[\underline{x}(p^*,\psi^*)^2-\underline{x}(p^N,0)^2]+\frac{s(\psi^*)}{t}$.
Advertising increases consumer surplus if and only if
$2\theta[(u(c)-p^*)\Delta\overline{x}+(p^N-p^*)\overline{x}(p^N,0)]+2(1-\theta)[(u(c)-p^*)\Delta\underline{x}+(p^N-p^*)\underline{x}(p^N,0)] >$
$\theta t[\overline{x}(p^*,\psi^*)^2-\overline{x}(p^N,0)^2]+(1-\theta)t[\underline{x}(p^*,\psi^*)^2-\underline{x}(p^N,0)^2]$.

Proof: See Appendix.

Intuitively, advertising increases welfare if the rent coming from consumers that become persuaded to buy a good because of advertising exceeds the additional costs of differentiation and the overall costs of advertising. When looking at equation (4.9), we can see that if the monopolist only sells to motivated consumers with and without advertising, advertising always increases the price. This leads to the following corollary.

Corollary 4.1. *If only motivated consumers buy a good both under advertising and when advertising is not feasible, then advertising always decreases consumer surplus.*

Proof: See Appendix.

In a case in which the monopolist sells to only motivated consumers both with and without advertising, if the bias is such that the most distant motivated consumer who buys a good without advertising obtains a non-positive consumption utility, then advertising always lowers welfare. This leads to our second corollary.

Corollary 4.2. *If only motivated consumers buy a good both under advertising and when advertising is not feasible, and $\overline{\alpha}(0)\overline{V} \geq u(c)$, this is sufficient for advertising to decrease welfare.*

Proof: See Appendix.

4.2 graphically shows some implications of the two Corollaries. We present two cases. In the first case, a motivated consumer's upward bias is sufficiently low for advertising to be potentially welfare beneficial, i.e. $\overline{\alpha}_I(0)\overline{V} < u(c)$. In the second case, the opposite holds: $\overline{\alpha}_{II}(0)\overline{V} > u(c)$. The vertical axis in the figure represents a motivated consumer's

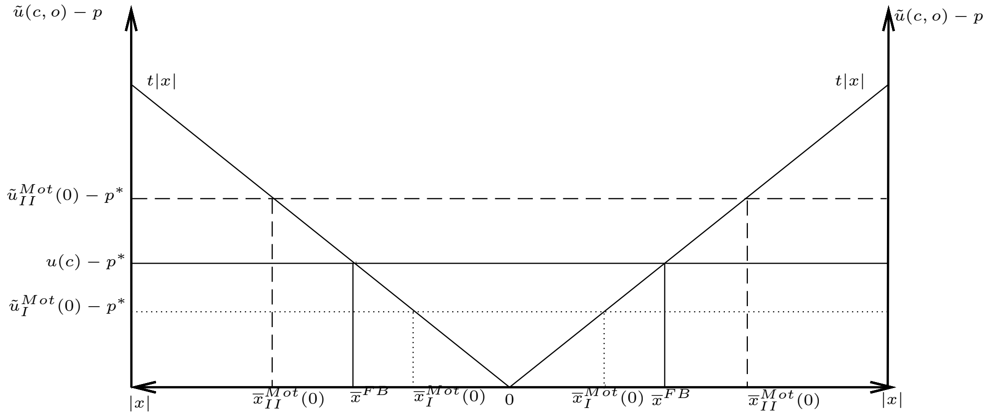


Figure 4.2: Monopoly effects if only motivated consumers are targeted

consumption utility of an ideal product variety given price p, $\tilde{u}(c|\overline{o}) - p$. As shown in section 4.4, when only motivated consumers are targeted, this difference between ideal product variety and monopoly price is equal to the monopoly price. Thus, a consumer's surplus can be seen as the difference between the unbiased consumption utility $u(c)$ and the monopoly price, minus her location costs $t|x|$. The horizontal axis represents the absolute distance between a producer located at point x and the monopolist located at 0, $|x|$, in the product space. From a welfare standpoint, it is desirable that a consumer buys a good, as long as $u(c) \geq t|x|$, i.e. as long as a consumer's consumption utility is at least as high as her differentiation costs.

In our first case, we furthermore assume that a motivated consumer's utility of consuming her ideal variety under monopoly pricing and advertising, $\frac{\overline{\alpha}_I(\psi^*)\overline{V}+u(c)}{2}$, is also below the consumption utility $u(c)$ that she contributes to overall welfare. Consequently, the distance between the most distant motivated consumer and the monopolist both with and without advertising, $\overline{x}_I(1)$ and $\overline{x}_I(0)$, respectively, is always below the welfare maximising distance, $\overline{x}^{FB}$. Thus, if advertising costs are sufficiently low, advertising can potentially lead to an increase in overall welfare in this case. However, because advertising causes an upward shift in the monopoly price from $\frac{\overline{\alpha}_I(0)\overline{V}+u(c)}{2}$ to $\frac{\overline{\alpha}_I(1)\overline{V}+u(c)}{2}$, one can see that this shift hurts all consumers who buy the good also without advertising, because the difference between the horizontal $u(c)$-line and the monopoly price under advertising, $\frac{\overline{\alpha}_I(1)\overline{V}+u(c)}{2}$, is smaller than the difference between $u(c)$ and the monopoly price without advertising. For motivated consumers who only buy under advertising, the differentiation costs are too big to have a positive consumer surplus from consuming the good. This demonstrates the potential trade-off between overall welfare and consumer surplus in this case.

In the second case, a motivated consumer's upward bias is always bigger than her consumption utility. As one can see from the figure, this implies that in a monopoly without advertising, consumption of the most distant consumers already decreases overall welfare, since their location costs $t|\overline{x}_{II}(0)|$ exceed their expected consumption utility $u(c)$. Thus, any further upward shift in the predicted consumption utility will always hurt welfare. Advertising can only increase consumer surplus if unmotivated consumers buy the good under advertising. Intuitively, such a case might arise where advertising leads to a decrease in the monopoly price.

Thus, we can conclude that depending on the different bias terms $\overline{\alpha}\overline{V}$, $\underline{\alpha}\underline{V}$, and on the fraction of motivated consumers in the buying period, there exist cases in which advertising can be both welfare beneficial and increase consumer surplus. However, as we have shown, after certain outcomes advertising neither increases welfare nor consumer surplus. Our results differ from the normative analysis in Dixit and Norman (1978), where advertising is always excessive. This difference occurs because in our model, fluctuating states of motivation result in the existence of two different consumer types in the buying period, i.e. motivated and unmotivated consumers, who have to make their consumption decisions.

4.5.2 Duopoly welfare

Analogously to the monopoly case, we firstly give expressions for consumer surplus and overall profits in the market. Denote $\overline{x}_0$ the distance between firm 0 and the most distant motivated consumer who purchases a good from this firm, $\overline{x}_1$ is the analogue distance from firm 1 to its most distant purchasing motivated consumer, and the distances from the two firms to their most distant purchasing unmotivated consumers are denoted $\underline{x}_0$ and $\underline{x}_1$ respectively. In case a consumer type is not served by a firm, this distance will be set equal to 0. Let D denote the overall demand for both goods in the market. In addition, we split up this overall demand into the demand for good A, D_A, and the demand for good B, D_B. These can be written as $D_A = \theta \int_0^{\overline{x}_A} d\tau + (1-\theta) \int_0^{\underline{x}_A} d\tau$, and $D_B = \theta \int_0^{\overline{x}_B} d\tau + (1-\theta) \int_0^{\underline{x}_B} d\tau$. Consequently, the overall demand for both goods can be written as

$$D := D_A + D_B = \theta\left(\int_0^{\overline{x}_A} d\tau + \int_0^{\overline{x}_B} d\tau\right) + (1-\theta)\left(\int_0^{\underline{x}_A} d\tau + \int_0^{\underline{x}_B} d\tau\right). \tag{4.16}$$

The total consumer loss in form of cumulated costs due to the variation in consumers' tastes, $T(x)$, can now be written as $T(x) = \theta t(\int_0^{\overline{x}_A} \tau d\tau + \int_0^{\overline{x}_B} \tau d\tau) + (1-\theta)t(\int_0^{\underline{x}_A} \tau d\tau + \int_0^{\underline{x}_B} \tau d\tau)$. This can be rewritten as

$$T(x) = \frac{t}{2}\left[\theta\overline{x}_A^2 + \theta\overline{x}_B^2 + (1-\theta)(\underline{x}_A)^2 + (1-\theta)(\underline{x}_B)^2\right]. \tag{4.17}$$

This gives us the expression for the change in welfare due to advertising in duopoly:

$$\Delta W = W - W^N = Du(c) - T(x) - s_A(\psi) - s_B(\psi) - D^N u(c) - T^N(x). \tag{4.18}$$

As in the monopoly case, we can now write a both necessary and sufficient condition for the social desirability of advertising in a duopoly, which is our second proposition.

Proposition 4.4. *Advertising is welfare beneficial in a duopoly if and only if* $D^A u(c) - T^A(x) - \sum_{i\in\{A,B\}} s_i(\psi) - D^N u(c) + T^N(x) > 0.$

The next corollary shows that advertising is never beneficial if the unmotivated consumers' willingness to pay is sufficiently high without advertising.

Corollary 4.3. *If* $\underline{\alpha}_A(0)\underline{V} + u(c) = \underline{\alpha}_B(0)\underline{V} + u(c) \geq \frac{3t}{2}$*, then advertising is never socially desirable in a duopoly.*

Proof: See Appendix.

The intuition for this corollary is as follows. If $\underline{\alpha}(0)\underline{V} + u(c) > \frac{3t}{2}$, then in the duopoly case the full market will be served. Due to a sufficiently high initial willingness to pay of the unmotivated consumers, the persuasive effects of advertising are not needed here. In fact, under these conditions, in any equilibrium without advertising, prices will always be $p_A = p_B = t$, and all consumers will be served. Advertising can thus never be optimal. The only positive effect of advertising, namely to persuade more people to buy the good, vanishes. Therefore, the competitive setting leads to a welfare loss due to the advertising expenditures.

Firms would also appreciate an advertising ban in the case of a symmetric equilibrium under advertising because of constant equilibrium prices and demands compared to increasing advertising costs. This is not necessarily true in an asymmetric equilibrium. If a firm in such an equilibrium gains a higher profit than $\frac{t}{2}$, then it benefits from advertising compared to the non-advertising case.

Since we have two different consumer types, a relatively low initial willingness to pay of unmotivated consumers can have several implications on the market outcomes. As mentioned before, one can think of situations in which the willingness to pay of motivated consumers is sufficiently high, i.e. $\overline{\alpha}(0)\overline{V} + u(c) \geq \frac{3t}{2}$, and at the same time, unmotivated consumers' willingness to pay is sufficiently smaller than t. When advertising is not allowed, we show in the Appendix that there exist equilibria in which firms only serve motivated consumers.

If both firms engage in advertising when allowed, and furthermore because $\underline{\alpha V}(1) + u(c) > t$ due to Assumption 4.1, at least in a symmetric equilibrium in which both firms advertise, this will also increase the number of purchased goods in a duopoly. This is because at least some unmotivated consumers will now buy the good. If the additional utility of the new consumers offsets the advertising costs in the market, then advertising is welfare beneficial.

Consider as an example consumers that plan to fly to a region with a high risk for malaria infections. Even though consumers know about the risks, they can misperceive them and hence become unwilling to pay the price for a prophylaxis. Other consumers who are afraid of an infection would like to buy the medicine even for very high prices at a given risk level. When being confronted with unpleasant pictures from the disease, consumers might then be willing to spend more money for the good, which reduces the risk of an infection.

4.6 Cooling-off periods

Mandatory cooling-off periods Loewenstein et al. (2003) mention the possibility that "Cooling-off periods that force consumers to reflect on their decisions for several days can decrease the likelihood that they end up owning products that they should not.[7]" However, their model does not include any analysis of such cooling-off laws. In this section, we analyze how a mandatory cooling-off period would affect overall and

[7] Loewenstein et al. (2003).

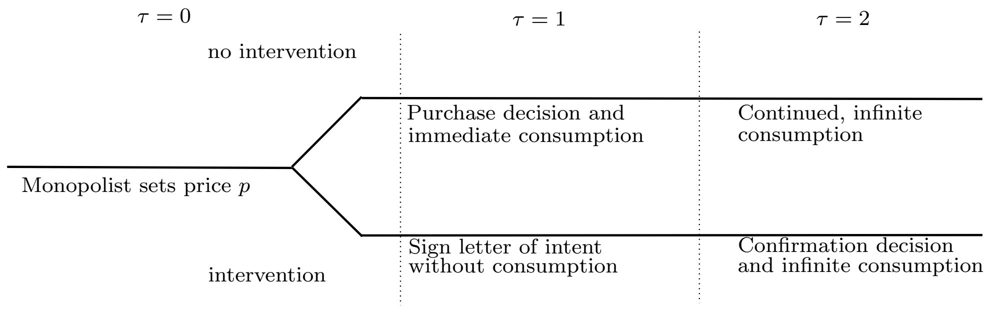


Figure 4.3: Game tree with and without policy intervention

consumer welfare in a variant of our model.

Suppose now that consumers can decide whether or not to purchase a durable good c that can be consumed today and for infinitely many periods in the future, with a constant discount rate δ. Denote $m_t \in \{\overline{o}, \underline{o}\}$ a consumer's state in period τ.

We abstract from the Coase (1972) conjecture by assuming that the monopolist can only offer the good once to consumers. For simplicity, we assume that $s(1) = 0$, such that the monopolist will always advertise costlessly, and that consumers' tastes are not horizontally differentiated. A consumer's predicted lifetime consumption utility when buying the good in period t and being in state o is

$$\tilde{U}_\tau(c|m_\tau = o) = \delta^\tau u(c, o) + \sum_{j=\tau}^{\infty} \delta^{j+1} \tilde{u}(c|o). \tag{4.19}$$

Define $\min\{a, b\} \equiv [a, b]^-$ and $\max\{a, 0\} \equiv [a]^+$. The monopolist's maximization problem can now be written as

$$\max_{p,\psi} \theta[[p, \tilde{U}(c|\overline{o})]^-]^+ + (1-\theta)[[p, \tilde{U}(c|\underline{o})]^-]^+.$$

It follows that the monopolist will target motivated consumers only if $\theta\tilde{U}(c|\overline{o}) > \tilde{U}(c|\underline{o})$, which is equivalent to $\theta > \frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}$

Assume following policy intervention. If a consumer wants to purchase a good, she has to sign a letter of intent first, and is only able to receive the good when agreeing to buy it again in the next period. We assume that the probability of being in a certain state is independent of the state in the previous period, i.e. $Prob(m_\tau = \overline{o}|m_{\tau-1} = \overline{o}) = Prob(m_\tau = \overline{o}|m_{\tau-1} = \underline{o}) = \theta$. Furthermore, we assume that a consumer cannot invest her money at a positive interest rate in between periods. Figure 4.3 shows the timing

θ consumer share	Net overall welfare effect	Net consumer welfare effect
$\theta \in \left(0, \frac{\tilde{U}(c\|\underline{o})}{\tilde{U}(c\|\overline{o})}\right)$	negative	negative
$\theta \in \left(\frac{\tilde{U}(c\|\underline{o})}{\tilde{U}(c\|\overline{o})}, \sqrt{\frac{\tilde{U}(c\|\underline{o})}{\tilde{U}(c\|\overline{o})}}\right)$	positive (negative) if $\theta < (>) [1 + \frac{1-\delta}{\delta} \frac{u(c,\overline{o})}{u(c)}]^{-1}$	positive
$\theta \in \left] \sqrt{\frac{\tilde{U}(c\|\underline{o})}{\tilde{U}(c\|\overline{o})}}, 1\right)$	negative	positive

Table 4.1: Net effect of mandatory policy intervention

of the game both with and without the policy intervention.

The monopolist's maximization problem can be written as

$$\max_{p,\psi} \theta^2 [[p, \tilde{U}(c|\overline{o})]^-]^+ + (1-\theta^2)[[p, \tilde{U}(c|\underline{o})]^-]^+. \tag{4.20}$$

One can see that the monopolist will sell to motivated consumers only whenever $\theta > \sqrt{\frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}}$, which is a higher threshold compared to the case when there is no cooling-off period. If it is best in both cases to sell either to both consumer types, or to motivated consumers only, this policy intervention will always hurt overall welfare. The reason for this is that the price will not change as a result of an intervention that does not change the consumer clientele, and the cooling-off period will lead to a loss due to the foregone immediate consumption. The only case in which such an intervention might increase overall welfare is if the monopolist start selling to both consumer types instead of selling to only motivated consumers. This leads to a drop in price which causes a demand increase. In such a case, the change in overall welfare is positive whenever the demand increase counterweighs the loss of immediate gratification of motivated consumers, i.e. if $\frac{\delta}{1-\delta}u(c) - \theta[\frac{\delta}{1-\delta}u(c) + u(c,\overline{o})] > 0$. This is equivalent to $(1-\theta)\frac{\delta}{1-\delta}u(c) > \theta u(c,\overline{o})$. Rewriting this, we obtain the following proposition.

Proposition 4.5. *A mandatory cooling-off period increases overall welfare if and only if the monopolist changes from targeting only motivated consumers to selling to both consumer types, and* $\theta < [1 + \frac{1-\delta}{\delta}\frac{u(c,\overline{o})}{u(c)}]^{-1}$.

Proof: See Appendix.

Things change when looking at consumer welfare only. Naturally, consumers will benefit from a price drop. If only motivated consumers buy the good with and without the intervention, this will furthermore decrease their consumer welfare losses. However, if a policy intervention is not necessary, i.e. if the monopolist always sells to both consumer types, then this will lead to a loss in consumer welfare due to the foregone immediate consumption in the first period. We sum this up in our next proposition.

Proposition 4.6. *A mandatory cooling-off period increases (decreases) consumer welfare if* $\theta > (<) \frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}$.

Proof: See Appendix.

Table 4.1 sums up the findings of the mandatory cooling-off intervention.

Voluntary return policies So far, we have only considered the case of a mandatory cooling-off period. Such policy instruments try to prevent consumers from making impulse purchases up front by forcing them to wait an extra period before acquiring a product. There are, however, also prominent voluntary return policies that correspond to money-return warranties. In chapter 3 of this thesis, we have analyzed the effects of frictions with respect to return behavior. When abstracting from such frictions, in the case of a one-period return period, the monopolist's maximization problem is identical to the mandatory case maximization problem in equation 4.20. This is because when the monopolist targets motivated consumers only, a fraction $1 - \theta$ of the consumers that buy a product in the first period will return the product in the second period. From a normative perspective, there is however an important change compared to the mandatory cooling-off period case. Since consumers can now buy a product in the first period, there is no foregone initial loss due to the policy intervention. Furthermore, this policy makes targeting all consumers from the monopolist's side more likely. Consequently, in the case without return costs, a voluntary return period will never hurt consumers. This leads to our next proposition.

Proposition 4.7. *In the case without return costs, a voluntary return period will never decrease consumer welfare.*

Proof. In text. □

In reality, due to frictions in return behavior and often-enforced rules of only being able to return unused products, voluntary cooling-off periods will lose some of their effectiveness. If consumers face return costs that decrease the probability of returning a product, they change the incentives for firms: Ceteris paribus, targeting only motivated consumers becomes more likely. From a normative standpoint, the question then becomes whether the foregone initial utility loss from a mandatory cooling-off period outweighs the utility loss due to the return frictions in the voluntary case.

Dynamic aspects So far, we only included the possibility for consumers to buy a product in the first period.[8] One can also think of models in which a consumer can buy a product at a later date. In the most natural case a firm cannot commit to a fix price in later periods. In such a case, forward looking consumers will anticipate a decrease in prices in later periods, and thus the whole problem becomes redundant. If consumers are myopic and do not take different potential types in the distribution into account, this allows a monopolist to extract even more consumer rents than in the baseline case, sometimes also by varying prices in the different periods.

In case the monopolist can commit to the same price in later periods, there are some similarities to the baseline mandatory cooling-off outcome. First, the only two equilibrium price outcomes are $\overline{p} \equiv \tilde{U}(c|\underline{o})$ and $\underline{p} \equiv \tilde{U}(c|\underline{o})$, i.e. the willingness to pay for a motivated and unmotivated consumer, respectively. Second, assuming that the firm has the same discount rate δ as the consumers, the higher δ, the more likely is an outcome in which the monopolist sells a good at a price $\overline{p}$. In such a case, each consumer eventually buys a product at a high price, because the probability of being in a motivated state twice in a row goes to one as $t \to \infty$. Denote by D_t the demand for the product in period t at price p. Overall, the monopolist offers a high price whenever $E[\sum_{\tau=0}^{\infty} \delta^{\tau} D_{\tau}(\overline{p})]\overline{p} > \underline{p}$.

4.7 Conclusion

In this chapter we have formalized a model in which a projection bias is responsible for the demand shifts caused by persuasive advertising. Advertising increases the predicted future consumption utility for a good, which in turn leads to an increase in the consumers' willingness to pay.

In our model, the theoretical computation of overall and consumer welfare is straightforward. However, knowledge is needed about the exact form of the "revealed

[8]In the mandatory cooling-off period case, this attributes to the signing of the letter of intent in the first period.

mistakes" at least some consumers make. If consumers indeed suffer from such a bias, the question remains whether this behavior can also be rationalized by some kind of optimal decisionmaking by the consumer that is not accounted for in our model. One such example would arise if questioning their biases and computing their true future predicted consumption utility was costly for the consumers. If these costs on average exceed the costs of "impulsively" making a wrong purchase decision, then this behavior is on average optimal for the consumer. Another rationalization would be consumers deriving an intrinsic utility from making impulse purchases that are so strong that they dominate the disutility from overpaying for the product.

An important empirical questions is how one can potentially identify a change in the bias due to advertising as the source for the demand shifts. One could test in an experiment how consumers' accuracy for predicting future probabilities depends on advertising intensity.

From a policy standpoint, we have shown that consumer protection policies can also harm consumers under some circumstances, which suggests that policy-makers should be cautious not to over-protect consumers by introducing protection policies ad hoc. In general, we believe that more work should be devoted to exploring how exactly consumer biases affect consumers' decisionmaking in an industry environment.

4.A Proofs

Proof of Lemma 4.1

Proof. The first part of the lemma follows right from the text. Now we focus on the case in which advertising is not feasible. Since the price p is the only choice variable in this case, and $\Pi_{pp} < 0 \forall p$ in both cases, it follows that there always exists a unique profit-maximizing price p for each segment, i.e. for the segment in which only motivated consumers are targeted, and the segment for which only unmotivated consumers are targeted. From equation 4.8, it follows that the monopolist's profit when selling to both consumer types can be written as $\Pi^{All}(0) = \frac{1}{2t}[\theta\overline{\alpha}(0)\overline{V} + (1-\theta)\underline{\alpha}(0)\underline{V} + u(c)]^2$. From equation (4.10) it follows analogously that the price when only targeting motivated consumers can be written as $\Pi^{Mot}(0) = \frac{\theta}{2t}[\overline{\alpha}(0)\overline{V}+u(c)]^2$. Therefore, $\Pi^{All}(0) \geq \Pi^{Mot}(0)$, if $\frac{1}{2t}[\theta\overline{\alpha}(0)\overline{V} + (1-\theta)\underline{\alpha}(0)\underline{V} + u(c)]^2 \geq \frac{\theta}{2t}[\overline{\alpha}(0)\overline{V} + u(c)]^2 \geq$

Since the firms' profits are always non-negative, it follows from the above equation that $\theta\overline{\alpha}(0)\overline{V} + (1-\theta)\underline{\alpha}(0)\underline{V} + u(c) \geq \sqrt{\theta}(\overline{\alpha}(0)\overline{V} + u(c))$. This can be written as $(1-\sqrt{\theta})u(c)+(1-\sqrt{\theta})(1+\sqrt{\theta})\overline{\alpha}\overline{V} \geq \sqrt{\theta}(1-\sqrt{\theta})\overline{\alpha}\overline{V}$. This is equivalent to $\sqrt{\theta}(\overline{\alpha}(0)\overline{V} - \underline{\alpha}(0)\underline{V}) - \underline{\alpha}(0)\underline{V} \leq u(c)$. Rearranging this equation yields $\sqrt{\theta} \leq \frac{u(c)+\underline{\alpha}(0)\underline{V}}{\overline{\alpha}(0)\overline{V}-\underline{\alpha}(0)\underline{V}}$. Naturally, it can only be optimal for the monopolist to attract unmotivated consumers when the willingness to pay for unmotivated consumers is positive, i.e. if $\underline{\alpha}(0)\underline{V} + u(c) > 0$. Then, the right hand side of the previous weak inequality will always be positive, thus squaring this inequality will not change the sign of the weak inequality. This then finally leads to $\theta \leq \left(\frac{u(c)+\underline{\alpha}(0)\underline{V}}{\overline{\alpha}(0)\overline{V}-\underline{\alpha}(0)\underline{V}}\right)^2$. This completes the proof. □

Proof of Proposition 4.1 Recall our definition $\max\{a, 0\} \equiv [a]^+$. We now furthermore define $\min\{a, b\} \equiv [a, b]^-$.

Proof. We consider the case in which the unmotivated consumers' willingness to pay is sufficiently high, i.e. if $\underline{\alpha}(1)\underline{V} + u(c) \geq \frac{3t}{2}$. Firm A's maximization problem when both firms attract and serve both consumer types becomes

$$\begin{aligned} \max_{p_A,\psi_A,} \Pi_A =& p_A\theta\left[\left[\frac{(\overline{\alpha}_A(\psi_A) - \overline{\alpha}_B(\psi_B))\overline{V} - p_A + p_B + t}{2t}, 1\right]^-\right]^+ + \\ & p_A(1-\theta)\left[\left[\frac{(\underline{\alpha}_A(\psi_A) - \underline{\alpha}_B(\psi_B))\underline{V} - p_A + p_B + t}{2t}, 1\right]^-\right]^+ - s_0(\psi_A); \\ & \text{s.t. } p_B \in BR_B(p_A); \end{aligned} \tag{4.21}$$

where $BR_B(.)$ is firm B's best response function for a given price of firm A. The analogue maximization problem for firm B becomes

$$\begin{aligned} \max_{p_B,\psi_B} \Pi_B =& p_B\theta\left[\left[\frac{(\overline{\alpha}_B(\psi_B) - \overline{\alpha}_A(\psi_A))\overline{V} - p_B + p_A + t}{2t}, 1\right]^-\right]^+ + \\ & p_B(1-\theta)\left[\left[\frac{(\underline{\alpha}_B(\psi_B) - \underline{\alpha}_A(\psi_A))\underline{V} - p_B + p_A + t}{2t}, 1\right]^-\right]^+ - s_1(\psi_B); \\ & \text{s.t. } p_A \in BR_A(p_B); \end{aligned} \tag{4.22}$$

where $BR_A(.)$ is firm A's best response function for a given price of firm B. After a history (ψ_A, ψ_B), we solve for the optimal profit functions using backward induction. In a subgame following histories $(\psi_A, \psi_B) = (0, 0)$ or $(\psi_A, \psi_B) = (1, 1)$, we obtain the following best-response functions by deriving the firms' profit functions with respect to p_A and p_B respectively:

$$\begin{aligned} p_A =& \theta \frac{(\overline{\alpha}_A(\psi_A) - \overline{\alpha}_B(\psi_A))\overline{V}}{2} \\ &+ (1-\theta)\frac{(\underline{\alpha}_A(\psi_A) - \underline{\alpha}_B(\psi_B))\underline{V}}{2} + \frac{t}{2} + \frac{p_B}{2} = \frac{t}{2} + \frac{p_B}{2}; \end{aligned} \tag{4.23}$$

$$\begin{aligned} p_B =& \theta \frac{(\overline{\alpha}_B(\psi_B) - \overline{\alpha}_A(\psi_A))\overline{V}}{2} + \\ &(1-\theta)\frac{(\underline{\alpha}_B(\psi_B) - \underline{\alpha}_A(\psi_A))\underline{V}}{2} + \frac{t}{2} + \frac{p_A}{2} = \frac{t}{2} + \frac{p_A}{2}. \end{aligned} \tag{4.24}$$

Because of our symmetry assumption with respect to the firms' bias functions, it therefore follows that in any subgame following a history in which both firms set the same advertising intensity, given that unmotivated consumers willingness to pay is sufficiently high, the resulting prices are $p_A = p_B = t$. In the pricing subgame following $(\psi_A, \psi_B) = (0,0)$, the profits are then $\Pi_A = \Pi_B = \frac{t}{2}$; after $(\psi_A, \psi_B) = (1,1)$, the resulting profits become $\Pi_A = \Pi_B = \frac{t}{2} - s(1)$.

In the subgame resulting after the history $(1,0)$, there is a multitude of possible equilibrium outcomes, because the concerned segments of the best-response correspondences depend on the specific bias functions $\underline{\alpha}(\psi)$ and $\overline{\alpha}(\psi)$, as well as on the upward and downward bias components $\overline{V}$ and $\underline{V}$. We know that each firm i setting prices above $p_i = \overline{\alpha}_i(\psi_i)\overline{V} + u(c)$ will have 0 demand, all prices above this mark are at least weakly dominated by prices $p_i \in [0, \overline{\alpha}_i(\psi_i)\overline{V} + u(c)]$. Therefore, the relevant action space is compact and non-empty. Moreover, given any pair of advertising intensities (ψ_A, ψ_B), the firms' profit functions are continuous. Although the relevant strategy space in this subgame is compact, the firms' profit functions are not necessarily quasi-concave, which would assure the existence of a pure strategy equilibrium. However, continuity of the profit functions as well as compact and non-empty action spaces of all firms suffice to prove the existence of a mixed-strategy Nash equilibrium in such a subgame, see Glicksberg (1952) for the proof.

We can deduce from Assumption 4.1 that, given that one firm does not advertise, the other firm always has an incentive to do so. For example, given that firm B does not advertise, firm A will make profits higher than $\frac{t}{2}$: In such a case, we know that in equilibrium, each firm will set prices higher than or equal to 0. If firm B sets $p_B = 0$, then if firm A sets $p_A = t$, which does not have to be firm A's best response to firm B's price, its profits will already be above $\frac{t}{2}$. One can see this by putting these prices into equation (4.21), this gives the profit $\theta \left[\frac{(\overline{\alpha}_A(1) - \overline{\alpha}_B(0))\overline{V}}{2}, t\right]^{-} + (1-\theta)\left[\frac{(\underline{\alpha}_A(1) - \underline{\alpha}_B(0)\underline{V}}{2}, t\right]^{-} - s_0(1)$, which is always bigger than $\frac{t}{2}$ because of Assumption 4.1. Because of the complementary effects in prices, firm A's profits will further increase in the prices of firm B, equations 4.23 and 4.24 shows this for the case in which the advertising firm serves both markets. Since it is under these circumstances always possible to serve both consumer types for this firm, it will only focus on motivated consumers solely if its profits are higher compared to the case where it serves both consumer types. Therefore, an incentive to advertise for at least one firm in the market is guaranteed. Whether both firms will invest in advertising thus depends on the obtained profits of firm B. If, given these advertising intensities, firm B will earn profits less than $\frac{t}{2} - s(1)$, then it is an optimal decision for this firm to invest in advertising in the advertising period. In such a case, no firm can profitably deviate by setting a price different from $p_A = p_B = t$ and advertising intensity other than $\psi_A = \psi_B = 1$. As mentioned before, this results in profits of $\Pi_A = \Pi_B = \frac{t}{2} - s(1)$ and completes the proof. □

Proof of Proposition 4.2

Proof. The proof is analogous to the proof of Proposition 4.1. Again, after any symmetric advertising history, i.e. either after $(\psi_A, \psi_B) = (0,0)$ or $(\psi_A, \psi_B) = (1,1)$, equilibrium prices will be $p_A = p_B = t$, if $\underline{\alpha}(0)\underline{V} + u(c) \geq \frac{3t}{2}$. Because of Assumption 4.1, the advertising history (0,0) is never an equilibrium advertising strategy pair, for firms have an incentive to deviate by setting a positive advertising intensity given that the other firm does not advertise. Furthermore, we know from Lemma 4.1 that equilibrium profits in a symmetric equilibrium will be $\frac{t}{2} - s(1)$ for both firms if unmotivated consumers' willingness to pay under advertising is equal or above $\frac{3t}{2}$. Therefore, if in the subgame following after $(1,0)$, firm B will make an equilibrium profit bigger than or equal to $\frac{t}{2} - s(1)$, then no firm can profitably deviate by setting another advertising strategy in the first period. This completes the proof. □

Existence of an equilibrium in which only motivated consumers are served without advertising

Proof. Our proof goes in two steps. Firstly, we will show that if only motivated consumers are in the market, $\theta = 1$, and if $\overline{\alpha}_i(0)\overline{V} + u(c) \geq \frac{3t}{2}\ \forall\ i\ \in \{A, B\}$, then if advertising is not allowed, firms will both set prices equal to t. Secondly, we will show that even if $\theta < 1$, i.e. if there are also unmotivated consumers in the market, and if the willingness to pay is sufficiently low for the unmotivated consumers, then under some circumstances, given that one firm sets its price equal to t, no firm can profitably deviate by setting a lower price and thereby also attract unmotivated consumers.

Suppose that the willingness to pay of the motivated consumers is sufficiently high, $\overline{\alpha}_i(0)\overline{V} + u(c) \geq \frac{3t}{2}\ \forall\ i\ \in \{0, 1\}$, and that there are only motivated consumers in the market, $\theta = 1$. The resulting maximization problem for firm A in the price setting stage when advertising is not allowed becomes

$$\begin{aligned}\max_{p_A}\Pi_A =& p_A \left[\left[\frac{(\overline{\alpha}_A(0) - \overline{\alpha}_B(0))\overline{V} - p_A + p_B + t}{2t}, 1\right]^-\right]^+ \\ & \text{s.t. } p_B \in BR_B(p_A).\end{aligned} \tag{4.25}$$

The analogous problem for firm B is

$$\begin{aligned}\max_{p_B}\Pi_B =& p_B \left[\left[\frac{(\overline{\alpha}_B(0) - \overline{\alpha}_A(0))\overline{V} - p_B + p_A + t}{2t}, 1\right]^-\right]^+ \\ & \text{s.t. } p_A \in BR_A(p_B).\end{aligned} \tag{4.26}$$

After the symmetric history $(\psi_A, \psi_B) = (0, 0)$, by taking the derivative of the profit functions with respect to p, one obtains the firms' best response functions:

$$\begin{aligned} p_A &= \frac{(\overline{\alpha}(0) - \overline{\alpha}(0))\overline{V}}{2t} + \frac{t}{2} + \frac{p_B}{2} = \frac{t}{2} + \frac{p_B}{2} \\ p_B &= \frac{(\overline{\alpha}(0) - \overline{\alpha}(0))\overline{V}}{2t} + \frac{t}{2} + \frac{p_A}{2} = \frac{t}{2} + \frac{p_A}{2}. \end{aligned}$$

This will result in the optimal pricing strategies

$$p_A = p_B = t.$$

Now suppose that $0 < \theta < 1$, so that both consumer types are in the market. Suppose furthermore that the willingness to pay for the unmotivated consumers without advertising is at most equal to d: $\underline{\alpha}(0)\underline{V} + u(c) \leq d$, where $d \in (0, \frac{t}{2})$.

Now in case both firms only attract motivated consumers, we know that after a

symmetric advertising history firms set $p_A = p_B = t$. If no firm is allowed to advertise, firms profits will be equal to $\Pi_A = \Pi_B = \frac{\theta t}{2}$. We will now show that under some circumstances, it is not profitable for a firm to set a price so low as to also attract unmotivated consumers. Given that the firm A sets the price $p_A = t$, the maximization problem of firm B after the advertising history $(\psi_A, \psi_B) = (0, 0)$ becomes $\max_{p_B} \Pi_B = p_B \theta \left[\left[\frac{-p_B+2t}{2t}, 1\right]^-\right]^+ + (1-\theta)\left[\left[\frac{\alpha_B(0)\underline{V}+u(c)-p_B}{t}; d\right]^-\right]^+$.
If $d < \frac{\theta t}{2}$, then this suffices to show that a firm cannot profitably deviate in the pricing stage by setting a price below t. As $d < \frac{\theta t}{2} < t$ by assumption, we know that when setting a non-negative price, a firm can never attract all unmotivated consumers in the market. Therefore, we know that if firm B undercuts price $p_A = t$ of firm A, so that also unmotivated consumers demand the good, its profit will always be below $d \cdot 1 = \frac{\theta t}{2}$. When setting a price equal to or above d, then we know that firms will only attract motivated consumers and that the resulting maximization problem in the subgames following advertising history $(\psi_A, \psi_B) = (0, 0)$ yields equilibrium prices of $p_A = p_B = t$. Since firms are not allowed to advertise, it follows that there is no advertising stage in which firms could possibly deviate. Therefore, it follows for symmetry reasons that under the given circumstances no firm can profitably deviate by setting a price different from t, given that the other firm sets its price equal to t, so that an equilibrium exists in which only motivated consumers are served with prices $p_A = p_B = t$. This completes the proof. □

Proof of Proposition 4.3

Proof. For the first part of the proposition, recall equation (4.13):
$W(p, \psi) = 2\left(\theta \int_0^{\overline{x}(p,\psi)} [u(c) - t\tau] d\tau + (1-\theta) \int_0^{\underline{x}(p,\psi)} [u(c) - t\tau] d\tau\right) - s(\psi)$.
Now, monopoly advertising is beneficial, if $W(p^*, \psi^*) > W(p^N, 0)$. Solving the integral for both sides, this then yields $2u(c)[\theta\overline{x}(p^*, \psi^*) + (1-\theta)\underline{x}(p^*, \psi^*)] - t\theta\overline{x}(p^*, \psi^*)^2 - (1-\theta)t\underline{x}(p^*, \psi^*)^2 - s(\psi) >$
$2u(c)[\theta\overline{x}(p^N, 0) + (1-\theta)\underline{x}(p^N, 0)] - t\overline{x}(p^N, 0)^2 - (1-\theta)t\underline{x}(p^N, 0)$.
Rearranging this inequality yields
$2u(c)[\theta(\overline{x}(, p^*, \psi^*) - \overline{x}(p^N, 0)) + (1-\theta)(\underline{x}(p^*, \psi^*) - \underline{x}(p^N, 0))] > t(\theta[\overline{x}(, p^*, \psi^*)^2 - \overline{x}(p^N, 0)^2] + (1-\theta)[\underline{x}(p^*, \psi^*)^2 - \underline{x}(p^N, 0)^2]) + s(\psi^*)$. Dividing the inequality by t and substituting $\Delta\overline{x}$ and $\Delta\underline{x}$ into this equation then yields the expression for the first part

of Proposition 4.4.

For the second part of the proposition, recall (4.12):

$CS(p,\psi) = 2\theta \int_0^{\overline{x}(p,\psi)}[u(c)-p-t\tau]d\tau + 2(1-\theta)\int_0^{\underline{x}(p,\psi)}[u(c)-p-t\tau]d\tau$. Thus, the condition $CS(p^*,\psi^*) > CS(p^N,0)$ can be rewritten as

$2\theta(u(c)-p^*)\overline{x}(p^*,\psi^*) - \theta t\overline{x}(p^*,\psi^*)^2 + 2(1-\theta)(u(c)-p^*)\underline{x}(p^*,\psi^*) - (1-\theta)t\underline{x}(p^*,\psi^*)^2 >$
$2\theta(u(c)-p^N)\overline{x}(p^N,0) - \theta t\overline{x}(p^N,0)^2 + 2(1-\theta)(u(c)-p^N)\underline{x}(p^N,0) - (1-\theta)t\underline{x}(p^N,0)^2$.

This can be rewritten as

$2\theta(u(c)-p^*)\overline{x}(p^*,\psi^*) - 2\theta(u(c)-p^N)\overline{x}(p^N,0) + 2(1-\theta)(u(c)-p^*)\underline{x}(p^*,\psi^*) - 2(1-\theta)(u(c)-p^N)\underline{x}(p^N,0) > \theta t(\overline{x}(p^*,\psi^*)^2 - \overline{x}(p^N,0)^2) + (1-\theta)t(\underline{x}(p^*,\psi^*)^2 - \underline{x}(p^N,0)^2)$

We further rearrange this to $2\theta u(c)[\overline{x}(p^*,\psi^*) - \overline{x}(p^N,0)] - 2\theta[[p^*[\overline{x}(p^*,\psi^*) - \overline{x}(p^N,0)] + (p^N-p^*)\overline{x}(p^N,0)]] + 2(1-\theta)[(u(c)-p^*)[[\underline{x}(p^*,\psi^*) - \underline{x}(p^N,0)] + (p^N-p^*)\underline{x}(p^N,0)]] > \theta t[\overline{x}(p^*,\psi^*)^2 - \overline{x}(p^N,0)^2] + (1-\theta)t[\underline{x}(p^*,\psi^*)^2 - \underline{x}(p^N,0)^2]$. Substituting $\Delta\overline{x}$ and $\Delta\underline{x}$ into this equation yields the second expression in the proposition. This completes the proof. □

Proof of Corollary 4.1

Proof. This proof goes in two steps. In the first step, we show that in the case when only motivated consumers buy a good both under and without advertising, advertising leads to a decrease in consumer surplus for every consumer that purchases a good even without advertising. In the second step, we show that those consumers who become persuaded to buy the good through advertising also all have a negative consumer surplus. For the first step, we can see that the overall utility of a consumer located at point $x \in \left[-\frac{\overline{\alpha}(0)\overline{V}+u(c)}{2t}, \frac{\overline{\alpha}(0)\overline{V}+u(c)}{2t}\right]$ who purchases a good also without advertising can be written as $u(c) - \frac{\overline{\alpha}(0)\overline{V}+u(c)}{2} - tx$. Advertising changes the utility of each consumer to $u(c) - \frac{\overline{\alpha}(\psi^*)\overline{V}+u(c)}{2} - tx$, which is smaller than the previous expression since $\overline{\alpha}(\psi^*)\overline{V} > \overline{\alpha}(0)\overline{V} \geq 0$. because of Assumption 4.3. This completes the first step of the proof. For the second step, recall that the consumer closest to the monopolist on the line that does buy a good under advertising but not without advertising is at least located at a distance $\overline{x}(p^N,0) = \overline{\alpha}(0)\overline{V} + u(c)$ away from the monopolist. Therefore, her consumption utility then becomes $u(c) - p(\psi^*) - t\overline{x}(p^N,0) = u(c) - \frac{\overline{\alpha}(\psi^*)\overline{V}+u(c)}{2} - \frac{\overline{\alpha}(0)\overline{V}+u(c)}{2} = -\frac{(\overline{\alpha}(0)+\overline{\alpha}(\psi^*))\overline{V}}{2} < 0$. Therefore, all other consumers located further away from the monopolist who have not bought a good without advertising will have a negative consumer surplus as well. From step 1 and step 2 it follows that all consumers who buy a good under advertising will be worse off under advertising compared to the situation when advertising is not feasible. This completes the proof. □

Proof of Corollary 4.2

Proof. We show that once $\overline{\alpha}(0)\overline{V} \geq u(c)$, and the monopolist only targets motivated consumers in all cases, then every additional motivated consumer buying a good under advertising will decrease welfare. When advertising is not feasible, it follows from (4.8) that $p = \frac{\alpha(0)\overline{V}+u(c)}{2t}$. Thus, the most distant consumers that buy a good lie at distance $\overline{x}(p^N, 0) = \frac{\overline{\alpha}(0)\overline{V}+u(c)}{2t}$ away from the monopolist. Therefore, the overall welfare gain from the most distant consumer can be written as $u(c) - t\overline{x}(p^N, 0) = \frac{u(c)-\overline{\alpha}\overline{V}}{2} < 0$. Furthermore, in this segment $\overline{x} = \frac{\alpha(\psi)\overline{V}+u(c)}{2t}$ is increasing in the advertising intensity ψ. Therefore, if $\overline{\alpha}(0)\overline{V} \geq u(c)$, any additional consumer buying a good will reduce overall welfare. Thus, any advertising leads to a welfare-harming increase in motivated consumers buying the good on top of welfare-reducing advertising costs $s(\psi)$. This completes the proof. □

Proof of Corollary 4.3

Proof. If $\underline{\alpha}_A(0)\underline{V} + u(c) = \underline{\alpha}_B(0)\underline{V} + u(c) \geq \frac{3t}{2}$, then one can see from the proof of Proposition 4.1 that there exists only one pure-strategy equilibrium when advertising is not allowed in which firms set $p_A = p_B = t$. Therefore, if all consumers are served in the market without advertising, advertising cannot increase the number of consumers in the market. Furthermore, we will now show that each firm serving exactly one half of the market minimizes overall transportation costs. Recall equation (4.17). We know that the differentiation costs are the aggregated distances for all consumers in the market times a parameter t. When the market is fully served, we can rewrite equation (4.17) as $T(x) = \theta t \left(\int_0^{\overline{x}_A} \tau d\tau + \int_0^{1-\overline{x}_A} \tau d\tau\right) + (1-\theta)t \left(\int_0^{\underline{x}_A} \tau d\tau + \int_0^{1-\underline{x}_A} \tau d\tau\right)$
$= \frac{t}{2}[\theta\overline{x}_A^2 + \theta(1-\overline{x}_A)^2 + (1-\theta)(\underline{x}_A)^2 + (1-\theta)(1-\underline{x}_A)^2]$.
Taking the derivatives of this expression with respect to $\overline{x}$ and $\underline{x}$ yields the first-order conditions: $\frac{\partial T}{\partial \overline{x}_A} = \frac{t\theta}{2}(4\overline{x}_A - 2) \stackrel{!}{=} 0 \quad \Leftrightarrow \overline{x}_A = \frac{1}{2}$.
$\frac{\partial T}{\partial \underline{x}_A} = \frac{t(1-\theta)}{2}(4\underline{x}_A - 2) \stackrel{!}{=} 0 \quad \Leftrightarrow \underline{x}_A = \frac{1}{2}$.
Taking the second-order derivatives with respect to $\overline{x}$ and $\underline{x}$, one then obtains $\frac{\partial^2 T}{\partial \overline{x}_A^2} = 2\theta t$;

$\frac{\partial^2 T}{\partial \underline{x}_A^2} = 2(1-\theta)t$, which are both positive for all $\theta \in (0,1)$. Since the second-order cross-partial derivatives are 0, this suffices for showing that the Hessian matrix of the second-order derivatives is positive, so that the differentiation costs function has a minimum at $\overline{x}_A = \frac{1}{2}$; $\underline{x}_A = \frac{1}{2}$ when all consumers are served. Thus, if the market is fully served, differentiation costs are minimized if both firms serve exactly half of the market for the motivated consumers and half of the market for the unmotivated consumers. As in the proof for Corollary 4.1, firms therefore cannot increase consumer demand because of advertising and also cannot lower differentiation costs in the market because of it. Since advertising is costly, this completes the proof. □

Proof of Proposition 4.5

Proof. The overall welfare consists of the monopolsit's profit, π, and the consumer surplus, $CS(p,\psi)$. Note that we assume $s(1) = 0$ in this section, such that the monopolist will always set $\psi = 1$. We will thus drop the advertising intensity ψ for the rest of the proof. As before, define the demand for unmotivated and motivated consumers consumers as $\underline{D}$ and $\overline{D}$, respectively. The total demand for the good is $D = \underline{D} + \overline{D}$. Evaluated at time $\tau = 0$, the monopolist's overall profits without a policy intervention can be written as $\pi = D(p)p$. Denote π^{int} and p^{int} monopolist's profits and price, respectively, in case of a policy intervention. These profits now turn to $\pi^{int} = \delta D(p^{int})p^{int}$. In absence of the policy intervention, the consumer surplus becomes

$$CS(p) = \overline{D}(p)\left[u(c,\overline{o}) + \sum_{j=1}^{\infty}\delta^j u(c) - p\right] + \underline{D}(p)\left[u(c,\underline{o}) + \sum_{j=1}^{\infty}\delta^j u(c) - p\right]. \quad (4.27)$$

Denote CS^{int} the consumer surplus in case of a policy intervention. This can then be written as

$$CS^{int}(p) = \delta\left[\overline{D}(p)\left[u(c,\overline{o}) + \sum_{j=1}^{\infty}\delta^j u(c) - p\right] + \underline{D}(p)\left[u(c,\underline{o}) + \sum_{j=1}^{\infty}\delta^j u(c) - p\right]\right]. \quad (4.28)$$

Therefore, the change in overall consumer surplus due to the intervention can be written as

$$\pi^{int} - \pi + CS^{int} - CS = D(p^{int})(\frac{\delta^2}{1-\delta}u(c)) - D(p)\frac{\delta}{1-\delta}u(c) - \overline{D}(p)u(c,\overline{o}) - \underline{D}(p)u(c,\underline{o}). \tag{4.29}$$

We firstly consider the case $\theta \notin (\frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}, \sqrt{\frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}})$. In such a case the demand for the good does not change due to an intervention. One can then see from equation 4.29 that the intervention will always hurt overall welfare.
Now consider the case $\theta \in (\frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}, \sqrt{\frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}})$. This yields $D(p) = \theta$ and $D(p^{int}) = 1$. Equation 4.29 can now be rewritten as $\frac{\delta}{1-\delta}u(c) - \theta[\frac{\delta}{1-\delta}u(c) + u(c,\overline{o})]$. This is positive as long as $(1-\theta)\frac{\delta}{1-\delta}u(c) > \theta u(c,\overline{o})$. This can be rewritten as $\frac{\delta}{1-\delta}u(c) > \theta[u(c,\overline{o}) + \frac{\delta}{1-\delta}u(c)]$. Rearranging this term then yields $\theta < [1 + \frac{1-\delta}{\delta}\frac{u(c,\overline{o})}{u(c)}]^{-1}$. This completes the proof. □

Proof of Proposition 4.6

Proof. First consider the case $\theta < \frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}$ In such a case, the monopolist will always sell the good to all consumers at a price $p = p^{int} = u(c,\underline{o}) + \frac{\delta}{1-\delta}\tilde{u}(c,\underline{o})$. From equations 4.27 and 4.28, the change in consumer surplus due to the intervention can be written as, $CS^{int}(p^{int}) - CS(p) = (\delta - 1)\frac{\delta}{1-\delta}[u(c) - \tilde{u}(c,\underline{o})]$, which is always negative.
Now consider the case $\theta \in (\frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}, \sqrt{\frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}})$. In this case, we obtain $p = u(c,\overline{o}) + \frac{\delta}{1-\delta}\tilde{u}(c,\overline{o})$, $p^{int} = u(c,\underline{o}) + \frac{\delta}{1-\delta}\tilde{u}(c,\underline{o})$, $D(p) = \theta$, and $D(p^{int}) = 1$. The change in overall consumer welfare can be written as $CS^{int}(p^{int}) - CS(p) = \delta[u(c) - u(c,\underline{o}) + \frac{\delta}{1-\delta}(u(c) - u(c,\underline{o}))] - \theta\frac{\delta}{1-\delta}[u(c) - \tilde{u}(c,\overline{o})]$. From our assumptions, it follows directly that $\delta[u(c) - u(c,\underline{o}) + \frac{\delta}{1-\delta}(u(c) - u(c,\underline{o}))]$ is always positive, and $\theta\frac{\delta}{1-\delta}[u(c) - \tilde{u}(c,\overline{o})]$ is always negative, such that the policy intervention will always increase consumer surplus in this case.
Finally, consider the case $\theta > \sqrt{\frac{\tilde{U}(c|\underline{o})}{\tilde{U}(c|\overline{o})}}$. This yields $p = p^{int} = u(c,\overline{o}) + \frac{\delta}{1-\delta}\tilde{u}(c,\overline{o})$, $D(p) = D(p^{int}) = \theta$. The change in consumer surplus can now be written as $CS^{int}(p^{int}) - CS(p) = (\delta - 1)\frac{\delta}{1-\delta}[u(c) - \tilde{u}(c,\overline{o})]$, which is always positive. This completes the proof.

□

Chapter 5

General Discussion

In 1989, Timothy Bresnahan concluded in his article in the first Handbook of Industrial Organization: "We know essentially nothing about the causes, or even the systematic predictors, of market power, but have come a long way in working out how to measure them.[1]" Surprisingly, even if the empirical literature in industrial organization has made significant progress over the last 25 years, this statement is still true. The framework developed in chapter 2 can be used to jointly estimate marginal costs and the form of industry competition in a specific industry. The insights from such studies can in principle be used to gain further knowledge about the predictors of market power. To do so, it is necessary to study several industries in order to see the relationship between price-cost margins and competition across different industries, and potentially also across time.

The way the theoretical organizational economics literature has been evolving, structural econometric methods seem to be a very promising tool for testing hypotheses of within-firm behavior. The framework in chapter 2 makes use of industry data in a horizontal setting to identify the form of cooperation between horizontally aligned divisions. Inferring both frictions and the form of cooperation between different vertically aligned division seems to be a natural next step. Ideally, such work would also include within-firm data, for example data on divisional contracts. From the empirical side, these topics still seem to be unanswered.

[1] Bresnahan (1989).

Chapter 2 also briefly touches the issue of estimating synergies that result from mergers. Such synergies are very often brought forward as reasons for a merger. However, both the scope as well as the origins of the synergies are not well-explored. One needs a general framework that can distinguish demand side synergies in terms of an increase in joint brand evaluation from synergies arising due to savings in fixed costs and synergies arising due to savings in marginal costs of production. Such work also requires detailed cost data on the industry side.

From both a behavioral economics and consumer policy viewpoint, the empirical literature seems to be even more underdeveloped. Identifying consumer biases from a fully rational explanation to me seems to be an equally important task as detecting market power. The theoretical incorporation of consumer biases into an industrial organization framework, as done in chapters 3 and 4, can already give insights about the effects of potential consumer biases on market outcomes. However, I believe that for the near future the key question is whether empirical research in this field will be able to distinguish biased decision making from rational behavior. While some research has already tried to estimate the magnitude of consumer biases, even very basic econometric identification issues have not yet been solved. From a normative viewpoint, one key difficulty that looms is the imposed difference between revealed choices and revealed preferences in behavioral research, see also Caplin and Schotter (2008). Much more ground work in this respect seems necessary to meet the increased interest in behavioral questions inside and outside of academic research.

Bibliography

Agarwal, S., Driscoll, J. C., Gabaix, X., and Laibson, D. (2009). The age of reason: Financial decisions over the life cycle and implications for regulation. *Brookings Papers on Economic Activity*, 2:51–118.

Alonso, R., Dessein, W., and Matouschek, N. (2008). When does coordination require centralization? *The American Economic Review*, 98(1):145–179.

Appel, M., Labarre, R., and Radulovic, D. (2004). On accelerated random search. *SIAM Journal on Optimization*, 14(3):708–731.

Armstrong, M. (2008). Interactions between competition and consumer policy. *Competition Policy International*, 4(1):97–147.

Armstrong, M. and Chen, Y. (2009). Inattentive consumers and product quality. *Journal of the European Economic Association*, 7(2-3):411–422.

Bagwell, K. (2007). The Economic Analysis of Advertising. *Handbook of Industrial Organization*, 3:1701–1844.

Becker, G. and Murphy, K. (1993). A Simple Theory of Advertising as a Good or Bad. *The Quarterly Journal of Economics*, 108(4):941–964.

Bernheim, B. and Rangel, A. (2004). Addiction and cue-triggered decision processes. *American Economic Review*, 94(5):1558–1590.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Bjoernerstedt, J. and Verboven, F. (2012). Does merger simulation work? a "natural experiment" in the swedish analgesics market. *mimeo*.

Bloch, F. and Manceau, D. (1999). Persuasive advertising in hotelling's model of product differentiation. *International Journal of Industrial Organization*, 17(4):557–574.

Bloom, N., Genakos, C., Sadun, R., and Van Reenen, J. (2012). Management practices across firms and countries. *The Academy of Management Perspectives*, 26(1):12–33.

Bonnet, C. and Dubois, P. (2010). Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance. *The RAND Journal of Economics*, 41(1):139–164.

Bower, G. (1992). How might emotions affect learning. In Christianson, S., editor, *The Handbook of Emotion and Memory: Research and theory*, pages 3–31. Lawrence Erlbaum Associates.

Braun, K. (1999). Postexperience Advertising Effects on Consumer Memory. *Journal of Consumer Research*, 25(4):319–334.

Bresnahan, T. (1982). The oligopoly solution concept is identified. *Economics Letters*, 10(1-2):87–92.

Bresnahan, T. (1989). Empirical studies of industries with market power. *Handbook of industrial organization*, 2:1011–1057.

Brito, D., Pereira, P., and Ramalho, J. (2012). Mergers, coordinated effects and efficiency in the portuguese non-life insurance industry. *mimeo*.

Caplin, A. and Schotter, A. (2008). *The Foundations of Positive and Normative Economics: A Handbook*. Oxford University Press, USA.

Chebat, J., Davidow, M., and Codjovi, I. (2005). Silent voices: Why some dissatisfied consumers fail to complain. *Journal of Service Research*, 7(4):328–342.

Choi, B. and Ishii, J. (2010). Consumer perception of warranty as signal of quality: An empirical study of power train warranties. Technical report, Working Paper, Amherst College.

Ciliberto, F. and Williams, J. (2010). Does multimarket contact facilitate tacit collusion? inference on conjectural parameters in the airline industry. *mimeo*.

Coase, R. H. (1972). Durability and monopoly. *Journal of Law and Economics*, 15(1):143–49.

Conlin, M., O'Donoghue, T., and Vogelsang, T. (2007). Projection bias in catalog orders. *American Economic Review*, 97(4):1217–1249.

Corts, K. (1995). *The ready-to-eat breakfast cereal industry in 1994 (A)*. Harvard Business School.

Corts, K. (1999). Conduct parameters and the measurement of market power. *Journal of Econometrics*, 88(2):227–250.

Cotterill, R. and Franklin, A. (1999). An estimation of consumer benefits from the public campaign to lower cereal prices. *Agribusiness*, 15(2):273–287.

DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*, 47(2):315–372.

DellaVigna, S. and Malmendier, U. (2004). Contract design and self-control: theory and evidence. *Quarterly Journal of Economics*, 119(2):352–402.

Dessein, W., Garicano, L., and Gertner, R. (2011). Organizing for synergies. *American Economic Journal: Microeconomics*, 2(4):77–114.

Dhar, S. and Hoch, S. (1996). Price discrimination using in-store merchandising. *The Journal of Marketing*, 60(1):17–30.

Dixit, A. and Norman, V. (1978). Advertising and welfare. *The Bell Journal of Economics*, 9(1):1–17.

Drago, F. and Kadar, D. (2006). Rebate or bait? A model of regret and time inconsistency in consumer behaviour. *CEPR Discussion Paper 5757*.

Eliaz, K. and Spiegler, R. (2006). Contracting with diversely naive agents. *The Review of Economic Studies*, 73(3):689–714.

Ellison, G. (2006). Bounded Rationality in Industrial Organization. In Blundell, R., Newey, W., and Persson, T., editors, *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*. Cambridge University Press.

Emons, W. (1988). Warranties, moral hazard, and the lemons problem. *Journal of Economic Theory*, 46(1):16–33.

Erdem, T., Katz, M., and Sun, B. (2010). A simple test for distinguishing between internal reference price theories. *Quantitative Marketing and Economics*, 8(3):303–332.

Eysenck, M. E. and Keane, M. T. (1993). *Cognitive Psychology - A Student's Handbook*. Taylor and Francis, Erlbaum.

Farrell, J. and Shapiro, C. (1990). Horizontal mergers: An equilibrium analysis. *The American Economic Review*, 80(1):107–126.

Fauli-Oller, R. and Giralt, M. (1995). Competition and cooperation within a multidivisional firm. *The Journal of Industrial Economics*, 43(1):77–99.

Gabaix, X. and Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics*, 121(2):505–540.

Glicksberg, I. (1952). A Further Generalization of the Kakutani Fixed Point Theorem, with Application to Nash Equilibrium Points. *Proceedings of the American Mathematical Society*, 3(1):170–174.

Hitsch, G. (2006). An empirical model of optimal dynamic product launch and exit under demand uncertainty. *Marketing Science*, 25(1):25–50.

Ho, T., Lim, N., and Camerer, C. (2006). Modeling the Psychology of Consumer and Firm Behavior with Behavioral Economics. *Journal of Marketing Research*, 43(3):307–331.

Huppertz, J. (2007). Firms' complaint handling policies and consumer complaint voicing. *Journal of Consumer Marketing*, 24(7):428–437.

Inderst, R. and Ottaviani, M. (2013). Sales Talk, Cancellation Terms, and the Role of Consumer Protection. *Forthcoming in Review of Economic Studies.*

Jolson, M., Wiener, J., and Rosecky, R. (1987). Correlates of rebate proneness. *Journal of Advertising Research*, 27(1):33–43.

Knittel, C. and Metaxoglou, K. (2011). In search of the truth: Merger simulations using random coefficient logit models. *UC Davis Department of Economics working paper.*

Lafontaine, F. and Slade, M. (2012). Inter-firm contracts: Evidence. In Gibbons, R. and Roberts, J., editors, *The Handbook of Organizational Economics.* Princeton University Press.

Lau, L. (1982). On identifying the degree of competitiveness from industry price and output data. *Economics Letters*, 10(1-2):93–99.

Loewenstein, G., O'Donoghue, T., and Rabin, M. (2003). Projection bias in predicting future utility. *Quarterly Journal of Economics*, 118(4):1209–1248.

Lutz, N. and Padmanabhan, V. (1998). Warranties, extended warranties, and product quality. *International Journal of Industrial Organization*, 16(4):463–493.

Mann, D. and Wissink, J. (1990). Money-back warranties vs. replacement warranties: A simple comparison. *American Economic Review*, 80(2):432–436.

McClure, S., Li, J., Tomlin, D., Cypert, K., Montague, L., and Montague, P. (2004). Neural Correlates of Behavioral Preference for Culturally Familiar Drinks. *Neuron*, 44(2):379–387.

McElheran, K. (2010). Delegation in multi-establishment firms: The organizational structure of it purchasing authority. *US Census Bureau Center for Economic Studies Paper No. CES-WP-10-35.*

Michel, C. (2007). A Projection-Bias Approach towards Persuasive Advertising. *Diploma Thesis, University of Bonn.*

Michel, C. (2009). Essays in Behavioural Industrial Organisation. *Master of Philosophy Thesis, University of Oxford.*

Nelson, P. (1974). Advertising as information. *The Journal of Political Economy*, 82(4):729–754.

Nevo, A. (1998). Identification of the oligopoly solution concept in a differentiated-products industry. *Economics Letters*, 59(3):391–395.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 31(3):395–421.

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica*, 69(2):307–342.

Nocke, V. and Whinston, M. (2010). Dynamic merger review. *Journal of Political Economy*, 118(6):1201–1251.

OFT (2002). *Extended warranties on domestic electrical goods.* Office of Fair Trading, Fleetbank House, 2-6 Salisbury Square, London, EC4Y 8JX.

OFT (2012). Extended warranties on domestic electrical goods. *Office of Fair Trading market study report.*

Oliveira, A. (2011). Estimating market power with a generalized supply relation: Application to an airline antitrust case. *mimeo.*

Rivers, D. and Vuong, Q. (2002). Model selection tests for nonlinear dynamic models. *The Econometrics Journal*, 5(1):1–39.

Rubinfeld, D. (2000). Market definition with differentiated products: the post / nabisco cereal merger. *Antitrust Law Journal*, 68:163–185.

Schmalensee, R. (1978). Entry deterrence in the ready-to-eat breakfast cereal industry. *The Bell Journal of Economics*, 9(2):305–327.

Spence, M. (1977). Consumer misperceptions, product failure and producer liability. *The Review of Economic Studies*, pages 561–572.

TARP (1996). *TARP's Approach to Customer Driven Quality: Moving from Measuring to Managing Customer Satisfaction.* White House Office of Consumer Affairs, Washington DC.

Thomas, C. (2011). Too many products: Decentralized decision making in multinational firms. *American Economic Journal: Microeconomics*, 3(1):280–306.

Vickers, J. (2004). Economics for consumer policy. *Proceedings of the British Academy*, 125:287–310.

Weinberg, M. C. and Hosken, D. (2012). Evidence on the accuracy of merger simulations. *Review of Economics and Statistics, forthcoming.*

Yoshimoto, H. (2011). Reliability examination in horizontal-merger price simulations: An ex-post evaluation of the gap between predicted and observed prices in the 1998 hyundai–kia merger. *mimeo.*

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig angefertigt und mich keiner anderen als der in ihr angegebenen Hilfsmittel bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Mannheim, 28.4.2012

Christian Felix Michel

Curriculum Vitae

2009-2013	Doctoral Studies in Economics, Graduate School for Economic and Social Sciences, University of Mannheim, Germany
2007-2009	Master of Philosophy in Economics, University of Oxford, United Kingdom
2003-2007	Diploma in Economics, University of Bonn, Germany
2005-2006	Visiting Student, Ecole Nationale de la Statistique et de l'Administration Economique, Malakoff–Paris, France
2002-2003	Mandatory Social Service at Kreiskrankenhaus Waldbröl, Germany
2002	"Abitur" (High School Diploma), Kopernikus Gymnasium Wissen, Wissen, Germany