NORTHWESTERN UNIVERSITY

Essays on Consumers' and Firms' Forward-Looking Behavior

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Marketing

By

Federico Rossi

EVANSTON, ILLINOIS

June 2008

ABSTRACT

Essays on Consumers' and Firms' Forward-Looking Behavior

Federico Rossi

There is a growing interest in both marketing and economics in the dynamics of decision-making. Researchers have proposed richer behavioral models of consumers and firms with the purpose of learning behavioral primitives that static models cannot capture. In these studies, models move away from stylized views of decision-making by explicitly incorporating intertemporal dynamics. They assume *forward-looking* agents who develop expectations on the future and how it will be affected by their current actions.

This dissertation consists of three such studies. The first essay develops and estimates a dynamic model of consumers' decisions, in order to investigate the effectiveness of reward programs in travel industries. The structure of the dynamic model is necessary to determine the marginal utility that consumers extract from cash spent on travel goods and from rewards. I find that there is a significant portion of customers who get more value from one dollar worth of rewards than from one dollar paid for the good purchased. This result is consistent with the idea that in these industries an important segment of customers are employees who travel for work and make purchase decisions

using the money of their employer. In this situation, the reward scheme can exploit the principal-agent separation between employer and employee by inducing the latter, who is the recipients of the rewards, to choose higher-priced goods. The study also provides counterfactual analyses to quantify the impact of these consumers on the effectiveness of the reward program, and to evaluate the difference in terms of sales generated between reward programs and price reductions.

The second essay focuses on firms' forward-looking behavior, and proposes an approach for extending the analysis of dynamic games to multi-product firms. Ericson and Pakes (1995) propose what has since become the standard framework for dynamic games. The computational tractability of their model, however, limits its application to cases where firms own only a few products. The approach proposed here shows that it is possible to collapse the information generated by multiple products offered by the firms into a few market variables. This result can be used to study dynamics in differentiated-product industries, which have found scarce attention due to the limitations of the original framework.

The third essay is also a study of firms' dynamic behavior, providing an empirical application of the analytical result proposed in the second essay. The demand-supply equilibrium analyses proposed in the literature so far have almost exclusively modeled firms' pricing decisions, ignoring product assortment. In this study I estimate a dynamic supply system where firms make pricing and new product location decisions jointly. I use the model to investigate why companies in the U.S. ready-to-eat cereal market tend to launch new products that are similar to their existing ones. My estimates suggest the existence of an asymmetric fixed-cost structure that prevents firms from launching new products in any segment of the market.

Acknowledgements

I would like to thank all of those people who have taken part in this amazing journey at Northwestern. My wonderful advisor, Eric Anderson, guided me through the process, giving me unconditional support and encouragement in every circumstance. I am forever indebted to my committee members, Eric Anderson, Aviv Nevo, Karsten Hansen, and Lakshman Krishnamurthi, for their invaluable comments that have profoundly shaped my thinking. It would not have been possible to complete this program if not for all the help and training I received from my professors.

I would like to thank my family for their love. I am thankful to my mother, for being such an inspiring figure in my life. Thank you to my father, for his encouragement and his unshakeable optimism. Thank you to my sister, for her letters loaded with such unmistakable humor.

I would like to extend my gratitude to my dear fellows, Gonca Cengiz, Sandeep Conoor, Kanishka Misra, Jong-Hee Song, Manish Tripathi, Lei Wang, and all the special friends who I encountered in the past years and made this period such a joy.

Finally, I would like to thank Valentina, my wife. Without your love, support, and patience I would not have even started this adventure. This is dedicated to you.

A Valentina

Table of Contents

List of Tables

List of Figures

CHAPTER 1

$1 Discount or $1 Reward? The Effect of Consumers' Preferences on Reward Programs

1.1. Introduction

The use of reward schemes has drawn increased scrutiny in the past years. Criticisms have been raised about operational issues, such as the high cost of developing, marketing, and running reward programs (Dowling and Uncles 1997), as well as about the basic effectiveness of these schemes in generating brand loyalty (Shugan 2005; Hartmann and Viard 2008). Despite these claims, today we observe industries where almost every firm ties its sales to a reward program scheme. For example, airline companies have been using frequent-flyer programs to reward travelers with free tickets for many years[1]: currently, every major airline company offers a frequent-flyer program[2]. Similarly, in other industries such as hotels, credit cards, car rental agencies, and gas retailers[3], most companies invest in elaborate trading point schemes to reward customers' repeated purchases.

The widespread prevalence of reward programs in travel industries such as airlines has been claimed by some researchers to be due to their ability to exploit the principal-agent separation between business travelers and their employers (Levine 1987; Borenstein 1996).

[1]The first frequent-flyer program was introduced in 1980 by Western Airlines, soon followed by American Airlines in 1981.

[2]See Tan et al. (2002), page 291

[3]Reward programs are offered by most European gas stations, whereas this is not true for U.S. gas stations. I will discuss later the reasons behind this difference in policy.

When purchasing a travel good, the decision of an employee who travels for work (agent) cannot in most cases be perfectly monitored by her employer (principal): in this scenario the reward scheme functions as a bribe paid to the employee, who is induced to choose a higher-priced alternative by the prospect of the reward. These claims have been embraced in the literature, and a few studies on reward programs have already proposed equilibrium analyses based on the existence of such agency relationship (Cairns and Galbraith 1990; Basso, Clements, and Ross 2007). However there is no empirical evidence showing the existence in the marketplace of consumers who effectively behave as agents of the agency relationship suggested above, nor there is any attempt to quantify the impact of the presence of such consumers on the firm's decision to invest in reward programs versus other policies.

In this paper I investigate whether there are travel industry consumers whose behavior can be related to that of business travelers engaged in an agency relationship with their employer. I identify these consumers through revealed preferences, by observing the trade-off that they make between lower prices and extra points earned while purchasing travel goods. Because business travelers are recipients of rewards and do not pay travel goods with their money, I expect the value they attribute to one dollar of cash spent for the good is lower than one dollar worth of reward. This relationship is inverted for regular consumers, who should find cash more valuable than rewards because of its higher liquidity. I develop and estimate a structural dynamic model of consumer purchase and redemption choice in order to determine the marginal utility that consumers extract from cash spent on travel goods and the marginal utility they extract from rewards. I then compare the individual-specific marginal utilities to distinguish regular consumers, who show higher marginal utility from cash than from rewards, from other customers, who show higher marginal utility from rewards than from cash. I refer to this second type of consumers as agents, because their behavior can be related to that of business travelers described above.

This paper contributes to the current literature in three important ways. First, to my knowledge this is the first study showing that in travel industries there is a significant portion of consumers who behave as do business travelers engaged in agency relationship with their employer. This study quantifies the impact of the presence of such consumers on the firm's decision to invest in reward programs versus other policies. Such empirical evidence helps in understanding why reward programs are such a popular form of investment in travel industries, compared to other industries. Previous empirical studies (Stephenson and Fox 1992; Proussaloglou and Koppelman 1999) have shown that in the airline industry business travelers are less sensitive to prices than other consumers, and that they tend to choose carriers related to the reward programs they are enrolled in. Notice, however, that these studies are based on survey data; besides the potential lack of accuracy in self-reported data, in case of agency relationship revealed information might not necessarily be a reliable source of data because consumers who behave against the interests of their employer have an incentive to hide their identity.

Second, the study provides counterfactual analyses to assess the impact of reward programs on sales. A more intuitive method for this investigation would be to obtain data where the change in a firm's decision to implement the reward program is observed. However, unless this policy change is a deliberate field experiment run by the firm, the analysis could be affected by endogeneity issues, and the result of the comparison could be due to factors other than policy changes that are not observed by the researcher. On the other hand, administering field experiments that discontinue the reward program can be very expensive, if not impractical. In my analysis I do not need to observe policy changes to the program; I instead produce the effects of different policies by relying on the dynamic structure of the consumer choice model specified. This structure is tested in a preliminary analysis. In this study I also determine the level of price reduction which would generate the same level of sales induced by the program. The analysis does not account for competitive reaction to price changes, therefore it only provides the drop in prices necessary, but probably not sufficient, to generate the same level of sales induced by the program. Nonetheless, the results provide the firm with a useful benchmark value for the effectiveness of price reductions that can be compared with the effectiveness of the reward program.

Third, the study proposes a dynamic demand model of consumers' purchase and redemption choice to study reward programs with trading point schemes. In each period a consumer makes purchase and redemption decisions. For each transaction made with the company, the consumer receives a given number of points that she can accumulate and later use to trade rewards offered by the company. These points represent a species of currency for the consumer because she can use them to buy out rewards. The dynamic structure of the model makes explicit the variation in the trade-off between reward points and price paid as the number of points accumulated increases. According to the model, and assuming that rewards have positive values, consumers in each period experience a cost for not earning reward points and advancing toward the reward. This cost, also called a switching cost, depends on two factors: point deadline and time discount.

A deadline implies that a consumer could lose her accumulated points if not redeemed within a certain date; the switching cost is then measured by the value of the points lost because of the forgone purchase opportunity. The time discount is related to the time value of money. Whenever a consumer delays her purchase, she is implicitly delaying the redemption to a later date, which depreciates her final reward because of the time value of money. The time value of money can also be interpreted as the consumer's impatience to reach her reward. The switching cost is represented by the periodic discount times the number of points accumulated. The model derives the value the consumer attaches to rewards by assuming these switching costs and by observing how the trade-off in the purchase decision changes with the number of points accumulated. This model is similar to that proposed by Hartmann and Viard (2008). Unlike theirs, the model proposed here includes the consumer's choice of rewards, which is a feature of reward programs with trading point schemes.[4] Consumers can in fact decide whether or not to redeem and which reward to redeem whenever a new reward becomes available for their accumulated points.

Recent behavioral studies have shown that the accumulation of credits toward a reward can also increase consumers' repeated purchase because of psychological effects induced by reward schemes. Inspired by past experiments on animals (Hull 1932), Kivetz, Urminsky and Zheng (2006) have shown that consumers enrolled in reward programs accelerate their effort as they near their redemption goal. These studies ascribe this phenomenon to

[4]The application presented here also differs for two more aspects that are relevant for the identification: first, in this study consumers can make several purchase choices, each leading to different point earnings; second, the reward program has no deadline, so it becomes an infinite periods' dynamic problem.

the perception of goal proximity evoked by points accumulated, that affects individuals' motivation. In this paper I will focus on the economic incentive of reward programs.[5]

I apply the dynamic model to the market of fuel retailing. I obtain individual purchase data from a leading European firm that administers a reward program over the entire Italian territory. Given the reduced dimensions of this territory, most business travelers find it more convenient to drive rather than fly or use other transportation services. The Italian fuel market is therefore an industry where the number of business travelers is significant. As a matter of interest, all main fuel retailing companies in the market offer a reward program.[6]

Preliminary results support the forward-looking structure of the behavioral model; consumers' purchase decisions are indeed affected by the reward program. The estimation of the full dynamic model reveals that 19.2% of customers show the behavior that is typical of business customers/agents in a principal-agent relationship: these consumers are in fact less sensitive to the value of the dollars spent on fuel than to the value of rewards. More specifically, they consider a dollar of reward to be worth 1.6 times a dollar used to pay gasoline. Also, these consumers are among the heaviest users of gasoline, representing the company's most valuable customers. Overall, the reward program increases the volume of sales by 5.29%. The presence of agents increases the overall effectiveness of the reward

[5] In a preliminary study, I identified the behavioral effects suggested by Kivetz et al. by assuming that consumers extract utility from simply holding points beyond the value of rewards; I found this specific effect to be scarse. Also, the result is sensitive to the choice of the functional form relating points with utility, which is not suggested by the theory.

[6] Similarly, in other European countries most fuel retailers offer a reward program. On the other hand, in the USA only a few of them have one. This could be due to territorial differences: because of limited distances and more homogeneous population density across the territory compared to the U.S., European business travelers might often prefer to drive rather than fly, so in the U.S. fuel retailing market there might not be enough business travelers to justify the use of a reward scheme.

program and decreases the effectiveness of alternative investment in price reductions: to reach the same volume the company would need to cut prices by at least 1.85%. Without agents, a decrease of 1.68% would be sufficient.

1.1.1. Previous Literature

Despite a number of theoretical papers that recognize the role played by the principal-agent separation on the effectiveness of reward programs (Levine 1987; Cairns and Galbraith 1990; Borenstein 1992; Borenstein 1996; Basso, Clements, and Ross 2007), there is hardly any empirical evidence supporting such a claim. The few studies available have used stated preference surveys of airline travelers (Stephenson and Fox 1992; Proussaloglou and Koppelman 1999) or purchase information recorded by the employer (Nako 1992) to show that frequent flier programs can change business travelers' willingness to pay. The scarcity of empirical evidence is mostly due to the researchers' limited ability to observe business travelers' choices. Notice that it is very difficult to overcome this issue because imperfect monitoring represents the reason the agency problem arises in first place: if the employer/principal had access to a mechanism that would allow her to monitor its business travelers, it is not clear if the moral hazard would still take place. In this study I take a different approach. I infer the existence of travelers who behave as agents by observing the behavior of consumers in the marketplace and recovering their preferences for cash spent vs. their preferences for rewards.

This work is also closely related to the economics and marketing literature on the effectiveness of reward programs on repeated purchases. A number of different theoretical studies have proposed reward programs as examples of mechanisms used by firms to induce switching costs for consumers and lessen competition (Klemperer 1987a; Klemperer 1987b; Caminal and Matutes 1990; Kim, Shi, and Srinivasan 2001; Kopalle and Neslin 2003). In these models, firms invest in reward programs in the first period of a multi-period game in order to enjoy a lower level of competition in the successive periods. A prerequisite for these models is the assumption that reward programs effectively 'lock-in' consumers as they accumulate points toward redemption. The effectiveness of reward schemes, however, has been the object of recent ongoing debate. Several empirical works have provided mixed evidence on the influence of reward programs on the pattern of repeated purchases.

The study by Lederman (2008), for example, provides support for the effectiveness of frequent flier programs on total sales. Her results show that improved redemption opportunities due to international partnerships increase airlines' market shares, particularly in their hubs. Lewis (2004) shows that the sales of an internet grocery retailer increase due to the retention effect of the reward program. Similarly, using loyalty data on grocery retailing Lal and Bell (2003) provide evidence that the reward scheme increases the profits of a supermarket chain; interestingly, however, the authors find that only the segment of infrequent customers is responsible for the increase, and that part of the profitability is offset by the revenue loss from the segment of frequent customers. In a recent paper, Hartmann and Viard (2008) find that a large fraction of the demand for golf courses, which was represented by the segment of frequent golfers, was indifferent to the reward scheme administered by the golf company. These studies show that consumers responses to reward programs are heterogeneous, and that the reward schemes might fail to create switching costs for the firm's most valuable customers. Most of these studies however have looked at simple reward schemes, such as "buy n, get the nth+1 free", that offer only one type of prize to all customers. These programs might often be inadequate to increase repeat purchases because by rewarding different customers with the same value, they fail to generate different incentive schemes for different customers. As Hartmann and Viard (2008) suggest, deadlines and discounting have a smaller effect on frequent customers because of the shorter time they spend between the first and the last purchase before the redemption.

Some companies overcome this issue by offering menus of rewards with increasing marginal values, i.e. rewards that can be redeemed with a higher number of purchases also have a higher marginal value per purchase. Besides accounting for consumers' horizontal preferences[7], these programs can also induce switching costs for frequent customers, who now prefer to accumulate more purchases in order to obtain the more valuable rewards. In doing so, these consumers will spend more time between the first and the last purchase accumulated than before. For example, airline companies offer their customers menus of rewards that include fragrances and tickets for international flight: occasional consumers will redeem the fragrance once they accumulate enough purchases, while frequent customers will accumulate more purchases and obtain international flights, in order to extract more value from the program. Usually these more sophisticated programs are based on some sort of a "trading points" scheme, that make them similar to the popular trading stamp programs of the 60's (see Fox (1968)); with their purchases consumers earn and accumulate points that can be later used to trade in rewards chosen from a menu provided by the firm. These reward schemes have received little attention so far, in part because the presence of multiple rewards complicates the analysis of the program's

[7]E.g., preferences for different colors of the same DVD player.

effectiveness. However, it is interesting to investigate if the limited effectiveness of single-reward schemes documented by the research cited above also extends to multiple-reward schemes. In this study I provide some evidence showing that this is not the case.

Methodologically, my study is closely related to the recently emerging literature on structural models of consumers' forward-looking behavior in economics and marketing. In these studies, consumers are forward-looking in that, when making decisions, they take into account the effect of current choices on future outcomes. There is a large body of literature that studies consumers' choice of frequently purchased goods in presence of quality uncertainty. The seminal study by Erdem and Keane (1996) adopts a dynamic structural framework to model the choice of consumers who are uncertain about prices and product quality. Consumers have the opportunity to receive signals from their own consumption and advertising to decrease their uncertainty about brand qualities; when making current choices they include the benefits of more accurate future choices. Similarly, Ackerberg (2003) presents a model where consumers take advantage of learning opportunities in the current period to increase their utility in future choices. He shows that advertising has both an informative and a prestige effect, and evaluates its impact on welfare effects.

Several studies also model consumers' formation of price expectation. Gonul and Srinivasan (1996) study the impact of consumers' expectations of coupons' availability in the future on current decisions. Erdem, Imai and Keane (2003) and Hendel and Nevo (2006) examine how expectations about future prices can induce consumers to stockpile goods. The first paper focuses on the impact of price expectations on demand elasticities. The second documents the potential bias produced by estimates that ignore the dynamic nature of demand.

Dynamic demand models are also applied to markets for durable goods, especially for studying price expectations of high-technology products. Melnikov (2000) and Song and Chitagunta (2003) use purchase data for new high-tech products and show that consumers delay their purchases because of their expectations of future price reductions. Erdem, Keane and Strebel (2003) report estimation biases of static demand due to ignoring the effects of price expectations.

The paper proceeds as follows. The next section presents the dynamic demand model. Section 1.3 explains the details of the estimation. Section 1.2 describes the dataset and shows some preliminary analysis. The results of the full model are in Section 1.5. Section 1.6 concludes.

1.2. The Model

1.2.1. Setup

In each period t, consumer h decides whether and how much fuel to purchase and receives the following utility

$$u\left(q_{ht}, p_t, \tau_{ht}, \nu_{ht}; \theta_h\right), \tag{1.1}$$

where $q_{ht} \in \{0, q_1, q_2, ..., Q\}$ represents the quantity of fuel purchased at any station of the company under analysis in period t, p_t is a vector of fuel prices for all firms in the market, τ_{ht} indicates the number of periods since the consumer purchased fuel from the company, ν_{ht} is a vector of individual- quantity- and time-specific stochastic shocks to utility from purchase. θ_h is a vector of consumer-specific preference parameters. The utility shock ν_{ht} represents the uncertainty about future purchases faced by consumers, who only know the distribution of shocks and their current realizations but not their future realizations. I assume the following specification for (1.1):

(1.2)

$$u\left(q_{ht}, p_t, \tau_{ht}, \nu_{ht}; \theta_h\right) = \begin{cases} \xi_{0h} + \xi_{q_{ht}h} + \alpha_h \left(p_{1t} - p_{0t}\right) q_{ht} + \beta_h log\left(\tau_{ht}\right) + \nu_{hq_{ht}t} & \text{if } q_{ht} > 0; \\ \nu_{h0t} & \text{if } q_{ht} = 0, \end{cases}$$

where p_{1t} is the price set by the company at period t for one liter of fuel, p_{0t} is a weighted average of the prices set by the industry's eight other biggest companies at period t, with respective market shares as weights. $\{\xi_{qh}\}_{q=0}^{Q}$, α_h, β_h are consumer-quantity-specific preference parameters to be estimated.

In the specification provided above, the coefficient λ_{2h} is recovered through the price difference $(p_{1t} - p_{0t})$, and not through the absolute variation of prices. Two important observations follow. First, a specification using absolute levels instead of differences would return unrealistic estimates for the cross-price elasticities with the outside good, which for markets such as gasoline are notoriously low. In contrast, the specification used here implicitly assumes that the outside good does not vary for an increase or a decrease of prices in terms of absolute levels; in other words, when the price of fuel increases for the entire market, consumers are not expected to decrease their use of cars. For the moderate variation in prices observed in the market (see Figure A.1) and the one-year time horizon considered in the empirical application, this seems a realistic assumption. Second, using prices set by the competition instead of just the company's prices provides a better mean of controlling for competition. In the model, the competitive pressure not only is captured through the fixed effect ξ_{0h}, but also through the variation in price differences.

While controlling for competition, the individual coefficient ξ_{0h} also accounts for consumers' inherent (unobserved) preferences for the company network, which could be due to the physical distance from home to the the company's closest gas station compared to the gas stations of competitors, or to the attractiveness of the neighborhood where the stations are located, and so on.

The company administers a reward program, and offers one reward point for each liter of fuel purchased. Therefore, in each period, consumer h earns q_{ht} reward points, which can be traded to redeem any of the prizes listed in the reward program catalog $\mathcal{J} = \{0, 1, ..., J\}$. A consumer decides whether or not to redeem and which reward to redeem. In that case, she must trade in a certain number of points, $\bar{c}_j$ depending on the reward j. In order to be able to make a redemption and receive the reward, a consumer must have a sufficient number of points. Points not traded do not expire: they roll over to the next period and are accumulated for future redemptions. The balance of points available to consumer h in period t, is represented by c_{ht}, which accounts both for points left from the previous period and points earned in the current period.

I denote the redemption of prize j by $d_{hjt} = 1$, where $j = 0$ stands for no redemption and $\bar{c}_0 = 0$, and I assume $\sum_j d_{hjt} = 1$, i.e. only one redemption per period is allowed. I also assume that a consumer makes redemption decisions whenever a new reward becomes available for her point balance. The consumer's problem can therefore be represented as

$$(1.3a)\quad V(s_1) = \max_{q_{ht}(s_t), d_{hjt}(s_t)} \sum_{t=1}^{\infty} \delta^{t-1} E\left[u(q_{ht}, p_t, \tau_{ht}, \nu_{ht}; \theta_h) + \sum_j d_{hjt}(\gamma_h r_j + \epsilon_{hjt}) \quad | s_1 \right]$$

$$(1.3b)\quad \text{s.t. } c_{ht} \geq 0, \ \sum_j d_{hjt} = 1,$$

$$(1.3c)\quad c_{ht+1} = c_{ht} - \sum_j d_{hjt} \bar{c}_j + q_{ht+1},$$

$$(1.3d)\quad d_{h0t} \neq 1 \Rightarrow c_{ht} - q_{ht} < \bar{c}_j \leq c_{ht}, \ j \neq 0$$

where s_t denotes the state in period t, δ is the discount factor, γ_h represents the consumer sensitivity to the value of the reward, ϵ_{hjt} is a random shock which accounts for randomness of the consumer's reward preferences, unobserved to the researcher.

Condition (1.3d) ensures that redemption decisions take place when the balance reaches a new reward. This assumption implies that consumers plan to make redemption decisions only when new rewards become affordable, so they do not plan to dispose of their points during other periods.[8]

1.2.2. State Space and Laws of Motion

In each period, the state s_t in (1.3a) carries the relevant information the consumer relies on to make her decisions: the reward point balance (c_{ht}), the number of periods from the last time the consumer purchased fuel from the company (τ_{ht}), the vector of shocks in the purchase choice (ν_{ht}), and the vector of shocks in the redemption choice (ϵ_{ht}).

[8]From discussion with reward program managers of the company this appears to be a reasonable assumption. It becomes unrealistic only if I believe consumers are very likely to revise their redemption choices before a new reward becomes available.

In each period t, the reward point balance is computed as in (1.3c): the current balance is obtained by subtracting the points traded in the previous period from the balance of the previous period, and adding the points earned from the current purchase. The number of periods from last purchase is determined according to the following law:

$$\tau_{ht} = \begin{cases} \min\left(\tau_{ht-1} + 1, \tau_{max}\right) & \text{if } q_{ht-1} = 0, \\ 1 & \text{otherwise,} \end{cases} \tag{1.4}$$

where τ_{max} is the maximum value that can be taken by variable τ. This value is chosen to keep the state space a reasonable size[9]

The last two state variables represent the two sources of uncertainty faced by the consumers; the uncertainty about future consumption and the uncertainty about future redemptions, respectively. I make the following assumption about their distribution:

Assumption A1: $\nu_{q_{ht}t}$ is independently and identically distributed type 1 extreme value.

Assumption A2: ϵ_{jt} is independently and identically distributed type 1 extreme value.

Both assumptions are made to reduce the computational burden in the estimation process. In particular, by assuming conditional independence I only need to keep track of current shocks, significantly reducing the dimension of the state space. The extreme value

[9]In the application I use $\tau_{max} = 10$.

distribution also allows the probabilities and the expected values to be expressed with convenient analytical closed forms (McFadden 1981). Both assumptions can be relaxed at a high cost in term of computational burden.

In my application I do not explicitly model expectations on prices: consumers assume future prices of companies to be equal to the mean prices observed throughout the year. An alternative route would be to explicitly model consumers' price expectations by imposing a specific pricing process. However, in this market there is not a clear process consumers use to develop expectations about prices; this is partly because companies set their prices using an Every-Day-Low-Price strategy rather than a Hi-Lo pricing strategy. Moreover, the fuel price is highly correlated with the price of crude oil, which consumers cannot predict.

1.3. Estimation

1.3.1. Overview

In this section I describe the procedure used to estimate the parameters of the demand model presented above. Following Rust's algorithm (1987), I solve the dynamic programming problem using a contraction mapping, which yields consumer-specific optimal decisions of purchase and redemption conditional on the state and individual-specific preferences. These decisions generate the expected future value of each choice discounted in the current period, conditional on state and individual-specific preferences. Since I do not observe the current random shock state variables, I use the distributional assumptions made above (A1 and A2) and derive the individual-specific likelihoods of observing each decision, conditional on the observed state variables. These likelihoods are obtained after integrating over the distribution of random coefficients previously defined. Finally, I use a method recently introduced by Ackerberg (2001) to efficiently compute the simulated likelihood of the observed sample and find the parameters that maximize it.

Notice that in Rust's original nested fixed point algorithm the inference of the parameters is derived only by the researcher's uncertainty about current shocks. If the researcher could perfectly observe current states, the model would yield perfect predictions. Clearly, this scenario would very likely create overidentification since we could also expect other sources of the researcher's "ignorance" to prevent her from perfectly predicting observed decisions. The heterogeneity in consumers' preferences introduces a new form of researcher's ignorance in the model through a parametric assumption (the distribution of random coefficients), that increases the accuracy of the model.

I next discuss the identification of the model parameters. In sections 1.3.3 and 1.3.4, I describe in detail the main steps in the estimation process outlined above: computing the value function, and deriving the likelihood of the observed sample to be maximized.

1.3.2. Identification

I informally discuss identification. The coefficient for the reward value is identified through the variation in the number of trading points accumulated and the observed quantities of fuel purchased. Consumers extracting a positive value from rewards will increase the quantities purchased as the number of points accumulated increases. The increase in purchasing is due to the pressure generated by opportunity costs; because of time discount, consumers in each period experience a cost for not earning reward points and advance toward the reward. This value depends on the discount factor, the number of points accumulated, and the value associated with the reward. Since the discount factor is assumed to be constant over periods, the increase in the quantities purchased for a given budget level is indicative of the value associated with the reward by the consumer. In particular the magnitude of this increase determines the level of the value for rewards. However, if a consumer does not extract any value from rewards, she does not perceive any opportunity costs for not purchasing fuel. Therefore, as she accumulates trading points in her budget, the quantities of purchased fuel should not increase. The identification of the other parameters is standard. The price coefficient is identified through the variation in prices with respect to competition and the variation in quantities purchased, the coefficient of the number of periods after last purchase is identified through the variation of such variable and quantities purchased, and the quantity-specific fixed effects are identified from the variation of purchases across quantities.

1.3.3. Value Function

In each period, a consumer can decide which quantity to purchase ($q_{ht} \in \{0, q_1, ...Q\}$) and which redemption to make ($j \in \{0, 1, ..., J\}$) according to the rules (1.3b)-(1.3d). The consumer makes both these decisions by forward-looking because they both affect her point balance in the next period, so they become the control variables of the dynamic problem.

I can rewrite the Bellman equation (1.3) as

$$V(s_t) = \max_{q',j'} \left\{ u\left(q_{ht} = q', s_t; \theta_h\right) + W\left(j_{ht} = j', s_t; \theta_h\right) \right\} \tag{1.5}$$

where $s_t = \{c_t, \tau_t, \nu_t, \epsilon_t\}$ and

$$W(s_t;\theta_h) = Emax\{\gamma_h r_{j'} + \delta E[V(s_{t+1})|s_t, q_{ht}=q', j_{ht}=j'],$$
$$\forall j' \in \mathcal{J},\ j' \neq 0 \Rightarrow c_{ht} - q' < \bar{c}_j \leq c_{ht}\ j \neq 0\} \tag{1.6}$$

Assumption A2 allows me to exploit the inclusive value to represent the nested redemption choice with a closed form expression, thus simplifying the computation of the Emax. To solve the Bellman equation I use policy function iteration (Rust 1996); in particular, I integrate out the utility shocks and I exploit the contraction mapping on probabilities (Aguirregabiria and Mira 2002) to compute the exact value function[10].

1.3.4. Likelihood

As I do not observe the current values of the state variables ν and ϵ, from my perspective consumers' decision rules are stochastic, therefore my model predicts decisions probabilistically. So, by assumption A1 the distribution of preference shock ν is type 1 extreme value, so the probability of purchase predicted by the model becomes:

(1.7)

$$Pr(q_{ht}=0|\, s_t;\theta_h) =$$
$$\frac{exp(M(s_t, q_{ht}=0))}{exp(M(s_t, q_{ht}=0)) + \sum_{q=q_1}^{Q} exp(\xi_{0h} + \xi_{qh} + \alpha_h(p_{1t}-p_{0t})q + \beta_h log(\tau_{ht}) + M(s_t, q_{ht}=q))}$$

[10]Notice that in their paper Aguirregabiria and Mira (2002) propose instead an estimator for the value function.

(1.8)

$$Pr\left(q_{ht}=q'\right|s_t;\theta_h)=$$

$$\frac{exp\left(\xi_{0h}+\xi_{q'h}+\alpha_h\left(p_{1t}-p_{0t}\right)q'+\beta_h log\left(\tau_{ht}\right)+M\left(s_t,q_{ht}=q'\right)\right)}{exp\left(M\left(s_t,q_{ht}=0\right)\right)+\sum_{q=q_1}^{Q}exp\left(\xi_{0h}+\xi_{qh}+\alpha_h\left(p_{1t}-p_{0t}\right)q+\beta_h log\left(\tau_{ht}\right)+M\left(s_t,q_{ht}=q\right)\right)}$$

$$\forall q'\in\{1,...,Q\}$$

where $M\left(s_t,q_{ht}\right)=log\left(\sum_{j=0}^{Q}\exp\left(\gamma_h r_j+\delta E\left[V\left(s_{t+1}|s_t,q_{ht},j\right)\right]\right)\right)$ is the expectation of the future values as a consequence of today's state s_t and action q_{ht}.

The model can probabilistically predict the choice of prize redeemed. This probability ideally could be included into the likelihood to identify additional primitives affecting the redemption choice. However, in this case, the estimation of the parameters would become much harder. Given that the main focus of the paper is on identifying the sensitivity of consumers to the value of rewards (i.e., recovering parameter γ), and the sensitivity to prices, I derive individual-specific likelihoods using only the purchase decisions predicted by the model as in (1.7) and (1.8)

$$L_h\left(q'_h|\theta_h\right)=\prod_{t=1}^{T_h}\prod_{t=0}^{1}Pr\left(q_{ht}=q'_{ht}\right|s_t;\theta_h)^{\left\{q_{ht}=q'_{ht}\right\}} \tag{1.9}$$

where q'_h is the T_h-dimension vector of observed choices made by consumer h. Notice that the laws of motion of the observed state variables presented in section 1.2.2 generate deterministic transition rules between states, therefore they do not affect the likelihood function.

I model heterogeneity using random coefficients. In particular I assume consumers have different preferences, represented by the vector parameter θ_h; these preferences, however, are generated by a common distribution, which I assume to be normal, i.e.

$\theta_h \sim N(\theta, \Sigma)$. Let η_h be independent standard normal distributed draws, and Γ be the Cholesky decomposition of the variance-covariance matrix Σ, then the preferences for individual h are given by

$$\theta_h = \theta + \Gamma \eta_h \ . \tag{1.10}$$

I can therefore express the individual-specific likelihood above as

$$L_h\left(q_h' | \theta, \Sigma\right) = \int \prod_{t=1}^{T_h} \prod_{t=0}^{1} Pr\left(q_{ht} = q_{ht}' \middle| s_t; \eta_h, \theta, \Sigma\right)^{\left\{q_{ht}=q_{ht}'\right\}} f\left(\eta_h\right) d\eta \tag{1.11}$$

If I followed the standard estimation procedure, the next step would be to approximate the integral above by extracting D draws from the random coefficient distribution $N(\theta, \Sigma)$, and find the parameters θ and Σ that maximize the likelihood of the total sample of consumers. However, notice that the choice probabilities in (1.7) and (1.8) depend on the value function that needs to be numerically computed for each parameter draw and each individual. If the optimization of the likelihood required R iterations, I would need to solve the Bellman equation $R * H * D$ times, and my estimation problem would become nearly unfeasible. I instead adopt a technique recently developed by Ackerberg (2001), which uses importance sampling and a change of variable to find consistent estimates for problems like the maximum likelihood above, reducing the number of times the dynamic problem needs to be solved. Let $u_h = \theta + \Gamma \eta_h$ be the change of variables, and $g(u) = f_1(u|\theta_0, \Sigma_0)$ be the importance sampling distribution, where θ_0 and Σ_0 are starting values for the true parameters θ Σ. Following Ackerberg's technique, the individual-specific simulated likelihood becomes

$$L_h\left(q'_h|\theta,\Sigma\right)=\frac{1}{D}\sum_{d=1}^{D}\left[\prod_{t=1}^{T_h}\prod_{t=0}^{1}Pr\left(q_{ht}=q'_{ht}|\,s_t;u_d\right)^{\{q_{ht}=q'_{ht}\}}\frac{f\left(u_d|\theta,\Sigma\right)}{g\left(u_d\right)}\right] \tag{1.12}$$

and the likelihood for the total sample of consumers is

$$L=\prod_{h=1}^{H}L_h \tag{1.13}$$

For each iteration of the algorithm, the new proposed parameters will not require computing $H*D$ Bellman equations, but only updating the numerator of the "weight" $\frac{f(u_d|\theta,\Sigma)}{g(u_d)}$. The choice probabilities are computed only once at the beginning of the algorithm.

1.4. Data and Preliminary Analysis

1.4.1. The Reward Program

The reward program was administered over the entire Italian territory by a leading petroleum company[11] that refines and sells fuel, lubricants, and other petroleum derivatives through its gas station network located in several European countries.[12]

The enrollment in the reward program is free of charge and open to all customers except truck drivers[13]. Consumers receive points on their electronic cards every time they purchase fuel or lubricant during opening hours at any Italian gas station in the company network. For each liter[14] of fuel purchased a consumer receives one reward point. The points are loaded through POS machines that automatically record the transaction on the

[11]The identity of the company is not revealed at this point.

[12]The company directly owns a small fraction of the gas stations in the network; most of them are associated through a franchising contract.

[13]The program rules specify that the vehicle's weight cannot exceed 35 quintals (7716.2 pounds).

[14]One liter corresponds to about 0.264 gallons.

database and release receipts with transaction details and point balance to the consumer. The program also includes the possibility for consumers who collect at least 800 points within three months to earn 10% more points in the three months that follow. For model tractability this scheme is not included in the analysis. However, this omission should not affect the final results, because the consumers who are affected by this scheme are only those earning on average 800 points every three months[15].

Every year the company offers a new collection of rewards, consisting of different types of products, from electronics to kitchen appliances. The number of points required to redeem rewards varies by reward; for most rewards, a monetary contribution is also requested when the reward is delivered. Rewards can be redeemed throughout the year, before a new collection is released; an exception is made for a group of prizes that are replaced twice a year. When a new collection is introduced, the rewards from the previous collection are discontinued. The new collection is categorically similar to the old collection. Points accumulated and not traded roll over to the next periods and do not expire. It is estimated that about 65% of all customers who visit the company stations participate in the program.

The data collected by the company provides transaction-level information on fuel and lubricant purchases at each gas station of the company network over the Italian territory, covering the period from October 2004 to December 2006. Each transaction reports customer card number, type of product bought (regular fuel, premium fuel, mineral lubricant, synthetic lubricant), number of points earned with the current purchase, and total

[15] This belief is shared by the company's management.

number of points accumulated on the card. The database also includes detailed information on other operations that customers can make with their electronic card: redeeming prizes associated with the reward program, pooling points with other cards, and swapping points across other reward programs in partnership with the company. The market values of the redemptions have also been provided by the company.

The database does not provide information on fuel prices. I retrieved this information from the Italian Ministry of Economic Development; in particular, I obtained the fuel prices that the nine biggest petroleum companies[16] suggest daily to their distributors at the national level.[17] I tested for the accuracy of this price data using a dataset with actual retailer-level fuel prices across the Italian territory. This data was self-reported by consumers who are members of the website www.prezzibenzina.it and periodically record on a voluntary basis the actual price paid in their last fuel purchases.[18]. The dataset provides 15,259 observations of fuel price paid, which I regressed on the suggested national-level prices used in this study, obtaining an R-squared of 0.84. This result suggests that the retailer-level price data closely follows the national-level price data that I will use in this study. In figure A.1, I report the national price levels (in euros) of the nine biggest companies during the period from October 2005 to September 2006 for one liter (0.264 gallons) of unleaded fuel with self-service pumping. This is by far the most common type

[16] CR9 $\cong$ 98%.

[17] Prices in the Italian fuel market have been regulated until recently. With the 04/13/1994 deliberation, the CIPE (*Comitato Interministeriale per la Programmazione Economica*, "Intragovernmental Board for Economic Plans") established that petroleum companies are free to set the prices charged to the distributors; however, they cannot contractally specify the prices that distributors should apply to the final consumers - only suggest them. Nevertheless, it is acknowledged that distributors tend to follow these suggested prices closely.

[18] I thank Stefano Bittante for making this dataset available.

of fuel purchased in the market so it is also the definition of fuel price that I will keep for my analysis.

Table A.1 reports statistics on the cards' main activities[19]. Compared to the purchase of regular fuel, the purchase of the other products is marginal. Only a small fraction of cards purchase premium fuel (7.6%), mineral lubricants (10.6%), or synthetic lubricants (2.6%). The table also shows that, of the total number of cards activated, only a small fraction makes a redemption in almost three years - only one customer out of four redeems at least one reward in the 27-month period. The reason for this small number has primarily to do with the low opportunity cost of consumers; the activation of the electronic card is free, so even consumers who rarely purchase at the stations of the company might decide to enroll in the program. Also, because of the cost-free activation and the possibility of pooling points across cards at a later date, many consumers might activate more than one card and later pool card points. During the period for which I have data, I see 17.3% of the cards being pooled together. A final observation is that 8.5% of all cards have their points swapped with points of other reward programs in partnership with the company.

Table A.2 reports statistics of fuel and lubricant purchase. From this Table one can see that when the typical (median) customer decides to buy, she purchases almost 8 gallons of regular fuel, 10.6 gallons of high-quality-Diesel fuel, 59.7 oz. of mineral lubricant, and 67.6 oz. of the more expensive synthetic lubricant. This customer visits 2 different gas stations in 27 months. It is clear from these numbers that compared to the purchase of regular fuel, which is the main business of the company, the purchase of the other products is marginal. In Figure A.2, I report the variation in the quantity of fuel purchased

[19]Due to confidentiality concerns, I only report relative values.

across weekdays for a sample of 500 individuals. The Figure clearly shows that the purchase of fuel is significantly lower on Sundays compared to that every other day. This could be in part due to the opening hours of the distributors. Since controlling for Sundays in the dynamic model would significantly increase the state space and therefore the computational tractability of the problem, to avoid biases in the estimation of price coefficients due to specific week-end policies offered by some companies, I keep Sundays out of the analysis.

Table A.3 presents statistics on the redemption activity for those consumers who redeemed at least once. Over the 27-month period of the database, the typical customer redeems only one reward, spending 1500 points corresponding to 3/4 of her total points accumulated.

1.4.2. Preliminary Analysis

The main purpose of this section is to provide evidence that the dynamic structure of forward-looking behavior assumed in section 1.2 is supported by the data. If consumers are myopic, demand for gasoline will only be affected by consumers' preferences entering the purchase utility (Equation 1.2). If consumers are forward-looking, however, their purchase behavior is also affected by the value of the rewards that will eventually be reached in the future. In this case, they will experience an opportunity cost from delaying purchases proportional to the number of points accumulated.

I present a static random-coefficient model to estimate fuel demand at gas stations. To test for consumers' forward-looking behavior, I include in the utility specification the number of points accumulated, which should affect fuel demand as discussed above. The period of analysis is days. In each period t, consumer h faces utility

$$(1.14) \quad u_{ht} = \begin{cases} -\lambda_{2h} p_{0t} + \nu_{h0t} & \text{if } y_{ht} = 0, \\ \lambda_{1h} - \lambda_{2h} p_{1t} + \lambda_{3h} exp(Points_{ht}) + \lambda_{4h}\tau_{ht} + X_t \lambda_{5h} + \nu_{h1t} & \text{if } y_{ht} = 1; \end{cases}$$

where $y_{ht} \in \{0, 1\}$ is the purchase choice, indicating whether in period t consumer h stopped and purchased fuel at any of the company's stations, p_{1t} is the fuel price set by the company at period t, p_{0t} is a weighted average of fuel price of the other major companies in the market, with respective market shares as weights, $Points_{ht}$ is the number of points (divided by 10,000) accumulated by the customer, τ_{ht} indicates the number of periods from the last time the consumer purchased fuel from the company, X_t is a 4-dimensional (row) vector of dummies indicating the Sundays and the seasons of the year, and ν is a vector of consumer- and time-specific shocks to preferences. I assume the stochastic term ν to be Gumbel distributed and use normal distribution for modeling heterogeneity across individuals.

The estimation was performed using the period when the 2006 reward campaign was active. A sample of 500 consumers was selected among those in the database who purchased at least once every two months, reaching at the end of the year a total quantity of purchased fuel not less than 500 liters (132.1 gallons), with a minimum average of 20 liters (5.3 gallons) per purchase. I exclude from the analysis inactive cards and very small customers, whose purchases have hardly any effect on the firm's profitability. I also exclude consumers who only visited one gas station throughout the year. This selection prevents the inclusion of cards owned by gas station employers who fraudulently use their POS machines to charge points for non-existent purchases; however, it does not affect the final results because consumers who only visit one station in three years are very likely responsible for a tiny portion of the company's total sales. Finally, I exclude from the analysis consumers who pooled points across cards, swapped points across reward programs, or participated in comarketing activities, because I do not have information on the other reward schemes.

The estimates of the random coefficient logit are shown in Table A.4. All coefficients have the expected signs. The coefficient of accumulated points is positive and significant, implying that consumers are more likely to make purchases as they accumulate points. This result suggests that consumers are forward-looking and that the reward program affects purchase decisions. The estimates of the off-diagonal elements from the variance-covariance matrix shows a positive correlation between the intercept and price; given that frequent customers also tend to purchase higher quantities of fuel, this supports the idea that consumers who purchase more frequently are also less sensitive to price. Besides income effects, this result could be due to the fact that consumers who purchase fuel more frequently and travel more are more likely to do that for business, and therefore they are less sensitive to prices because they do not pay fuel out of their pocket. Instead, their travel expenses, including fuel purchases, will be paid by the company for which they work.

The goal of this static analysis is simply to show that the data supports consumers' forward-looking behavior, and that consumers who travel more show less sensitivity to prices. In the next section I develop a fully structural model that allows me to recover the primitives of interest: the marginal utility for price and for rewards. These primitives allow me to distinguish agents from regular customers and quantify the effect of the reward program and price reduction on each consumer type.

1.5. Results

For the estimation of the dynamic model I employ a sample of 500 consumers from the 2006 reward campaign of the company. This sample is the same used in the static demand analysis. In the catalog of the reward program there are in total 58 different redemptions. I select six rewards among those most commonly redeemed. Each reward is selected to represent the redemption that has been most commonly chosen in a given interval of points. These six redemptions chosen account for about 26% of all points redeemed during that campaign.

1.5.1. Parameter Estimates

Table A.5 presents the ML estimates of the parameters for the dynamic model specification. Coefficients are significant and have the expected signs[20]. In particular, the coefficients for price and rewards are negatively correlated; this means that consumers more sensitive to prices are also those who consider rewards to be more valuable. Both coefficients express the intrinsic preference of individuals for money, which might be partly explained by their different income levels. The results also show a positive correlation between price and retailer fixed effect: consumers who purchase more often are also less sensitive to prices. This result can be due to the pricing strategy adopted by the firm, which is therefore more likely to select price-insensitive consumers. This result, however,

[20]For negative values of the reward coefficient the model is not identified because the value function is always zero. Therefore, I set this coefficient to be log-normally distributed, so that $log\left(\gamma_h\right) \sim N\left(\gamma, \Sigma\right)$.

can also be explained with the principal-agent argument: frequent travelers are likely to travel for work, so the reimbursement of their trip expenses could explain the lower price sensitivity of this segment of customers.

1.5.2. You Pay For The Good, I Get The Reward

In this section I identify those consumers on the market whose behavior can be related to that of business travelers engaged in an agency relationship with their employer. For each consumer I determine the marginal utility that consumers extract from dollars spent for fuel at the gas station and the marginal utility they extract from rewards. Because of the structural specification of the dynamic model, these values are the absolute values of the individual-specific coefficients associated with price and reward; these primitives are represented by the individual parameters α_h and γ_h in the value function (1.3a). Such values are found after drawing from the estimated sampling distribution of the parameter vector θ and deriving individual-specific coefficients and standard errors from the conditional distribution of each individual.

By comparing these two values I can distinguish regular consumers from agents: regular consumers are those individuals who extract more or the same value from cash spent at the gas station than from rewards, agents those individuals having this relation reversed, i.e. they extract more value from rewards than from cash. For a regular customer, the value of one dollar spent on fuel cannot be lower than the value she can extract from one dollar worth of reward. Cash is more liquid, and therefore always better than or at least equal to the monetary value of any good rewarded[21]. When a consumer is an agent

[21] An exception to this could be special rewards that do not have a market, although it is not the case of this reward program. Each and every reward offered on the catalog is also sold on the market.

of a principal-agent relationship, however, the value of a dollar spent on fuel decreases because it is money of the employer that she cannot dispose of. Since these primitives are estimates, I perform a t-test for each consumer to determine if the utility from rewards is statistically greater than the utility from cash[22].

In Figure A.3, I plot consumers based on their preferences, to show their intrinsic heterogeneity. To make the graph easier to read, I only plot 30 customers randomly selected from the sample used in my estimation analysis. The horizontal axis measures the individual marginal utility from money paid to purchase fuel. The vertical axis measures the individual marginal utility from rewards. The size of the bubble represents consumer's usage, i.e. the simulated quantity of gasoline purchased in one year by the individual at the company's stations. The diagonal line represents the indifference line, where consumers extract the same value from cash and from rewards. This line divides the graph into two areas, populated by two types of consumers: those who value more cash than rewards or are at least indifferent between them (regular customers), and those who value strictly more rewards than cash (agents). This graph suggests three important observations supporting the principal-agent interpretation of the results. First, the number of bubbles representing agents are few compared to those representing regular customers, suggesting that the agents represent a small portion of total customers. Second, despite their size, agents are among the heaviest product users, and therefore the firm's most valuable customers. Third, the standard deviation of the individual coefficients for price is greater than the standard deviation of the individual coefficients for reward. This observation is also confirmed in the whole sample of individuals (0.04 versus 0.028). Such

[22] More specifically, $H_0 : \alpha_h \geq \gamma_h$, $H_1 : \alpha_h < \gamma_h$, using 95% confidence interval.

difference, although not huge, could suggest that individuals turn out to be agents just because of their low utility with respect to cash. In other words, both regular customers and agents value rewards quite similarly. After all, except from intrinsic preferences for the products offered in the catalog, they are both recipients of the reward. The difference between them is mostly due to the value they can extract from the money spent at the pump station.

To better quantify agents' gasoline usage and their contribution to total sales, I simulate the total quantity of gasoline that every consumer purchases in one year when the reward program is on. I divide consumers into five groups based on the percentiles of the distribution of gasoline quantity purchased (0 to 20th, 21st to 40th, 41st to 60th, 61st to 80th, 81st to 100th). This quantity was simulated using the same initial level of state variables (zero points and one period after purchase) for each consumer. The results are reported in Figure A.4.

In the first column of the graph I report the percentage of total sales due to agents and the percentage of agents in the total population of consumers. Overall, 19.2% of the individuals are agents, who contribute to 21.2% of company's total volume of sales. In the next columns I break down consumers into their total quantity of fuel purchased. Two measures are reported: the impact of each group of customers on total sales, and the percentage of agents in the group. The results show that the percentage of agents in the group increases from light users to heavy users. The groups of light users, i.e. customers in the 0 to 20th percentile and in the 21st to 40th percentile, contribute to total sales on average by 17% or less; the percentage of agents among these consumers is low, 1% and 12%, respectively. This percentage increases to 29% for medium users, i.e. customers between the 41st and the 60th percentile, who are responsible for almost 20% of total volumes sold. The percentage of agents decreases slightly (21%) when we consider the groups between the 61st and the 80th percentile, but increases to one third in the group of heavy users. The latter group accounts for about 27% of total sales. Overall, these results lead to two important findings. First, there is a considerable group of consumers, among the population of the company's consumers, who value one dollar of reward more than one dollar spent on fuel. Second, most of these customers are heavy users. Not surprisingly, consumers who travel for business are likely to purchase fuel on a regular basis, and therefore purchase more often than people traveling for pleasure.

1.5.3. Agents and Reward Programs

So why should companies care about the presence or absence of agents among their customers? Compared to others, agents tend to be heavy users and less sensitive to price changes. Their choice of gas station will not be influenced by prices as much as by the reward program. In this scenario, the reward scheme represents a mechanism for companies to attract with a monetary benefit valuable consumers who are insensitive to price. The reward program can become a valid monetary incentive, an alternative to pricing, as long as consumers extract a positive value from rewards. More specifically, reward programs are a more efficient monetary incentive whenever one dollar spent on the good is worth less than one dollar of reward.

In Figure A.5, I report how much a dollar spent on fuel is worth for consumers, grouped by their type. In the first column I present the value for all customers. The average consumer finds cash and rewards to have a similar value. This relation however changes significantly depending on the type of customer. In particular, regular customers extract from the reward only 86% of its market value. In contrast, agents value a dollar of reward 60% more than a dollar of cash; this means that these customers are willing to pay up to $1.60 more for one dollar of fuel in order to get the reward. The reward schemes allows the firm to offer a bribe to extract these 60 cents, and it becomes a more efficient investment than price reductions. I will quantify this statement in the next section.

1.5.4. The Value of the Reward Program

In this section I evaluate the investment in the reward program by comparing it with a price reduction policy. I first use the demand estimates of the dynamic model to investigate the effects of the reward program on sales. In order to do that, I analyze two scenarios. In the first scenario I simulate consumers purchases using the demand estimates and holding everything else constant; this simulation generates the total sales produced in one year when the firm runs the reward program using the current design. In the second scenario I simulate a counterfactual scenario where the company does not run the reward program. By comparing the sales from the two scenarios I can assess the value of the reward program in terms of extra sales generated. Then, I run a set of counterfactuals where the company does not run the reward program, and instead invests in a price reduction policy: I run different price reduction policies in order to determine what price reduction should be undertaken to generate the same level of sales generated by the reward program.

I report the results in Figure A.6. The dark gray columns represent the level of sales generated by the reward program, and the light gray columns represent the sales generated by pricing policies with different price reductions. I normalize to 0 the sales generated when there is no reward program or price reduction. The first columns on the left are sales for the total population of customers. The reward program increases total volume sales by 5.29%. In the right columns I report sales by consumer type. The program produces a similar increase in sales for the group of regular customers. For agents, on the other hand, the increase is higher - 6.3%. There is a substantial difference also in the reaction of consumers to changes in prices. As expected, regular consumers are more elastic to price cuts. For this group the same additional sales generated by the reward program would be obtained by a price reduction of less than 1.8%. Agents, on the other hand, are less likely to react to price changes. In this group the same level of sales generated by the program could be reached only after cutting prices by 2.8%. Overall the company could achieve the same benefits in terms of repeated purchases decreasing prices by 1.85%. Such a price decrease does not reach the same level of sales as the reward program among agents; however the increase in sales among regular customers compensates for it. Notice that in this analysis I do not account for competitive reaction. Despite this limitation, this analysis represents a useful benchmark for comparing the effectiveness of reward programs versus price reductions: a price decrease policy of 1.85% will decrease the company's total revenue by the same percentage. Assuming that competitors will very likely respond to such policy decreasing their prices as well, the company will need to cut further its prices to maintain the same level of sales. Therefore, the pricing policy cost outlined above represents a lower bound estimate of the actual cost of this policy. If the cost for running the reward program is lower than such lower bound, the company should maintain the program.

1.6. Conclusions

Previous theoretical research has claimed that reward programs applied to travel industries are particularly effective because they exploit the separation between business travelers, who travel for work, and their employers, who pay for the travel expenses. The imperfect monitoring of the employee's traveling decisions allows the traveler to apply a trade-off between price and personal traveling benefits that is not necessarily optimal for the employer. The reward scheme exploits this trade-off and induces the employee to pay a higher price for a reward. Despite these claims, however, the empirical evidence is scarce. To my knowledge there are no studies showing the existence of this type of consumer in the marketplace or trying to quantify their impact on the firm's investment decisions. This paper is an attempt to fill this gap. I develop and estimate a dynamic model of purchase and redemption choice to show that there are consumers on the market who trade higher prices for extra reward points. According to their revealed preferences, these consumers find one dollar worth of reward more valuable than one dollar of cash spent on travel goods. Their behavior is consistent with that of business travelers engaged in an agency relationship with their employers.

I apply my demand model to the market of gasoline retailing. I use data from a leading company operating in Italy. The estimates suggest that 19.2% of its customers are agents. According to their revealed preferences, these customers value a dollar worth of reward 60% more than a dollar spent to fill their car tanks; this means that they are willing to pay up to $1.60 more for one dollar of fuel in order to get a reward. Also, these customers tend to be the heaviest users of gasoline, and therefore the company's most valuable customers. In such scenario, reward schemes become a potentially more effective investment than price reduction, because they represent a bribing scheme to extract these 60 cents. I investigate whether this is the case by using the estimates of the dynamic model to determine the volumes of sales generated by the reward program policy, and comparing it with a scenario where the company decides to reduce prices instead of investing on the reward program. Although competitive reaction is not modeled, this analysis provides a lower bound of the total costs for the pricing policy. The reward program generates an increase of 5.29% in total volume of gasoline sold. This increase is more pronounced for agents, who tend to purchase 6.3% times more gasoline. The presence of these consumers also makes the pricing policy a less effective investment alternative. To reach the same volume the company would need to cut prices by at least 1.85%. Without agents, a decrease of 1.68% would be enough.

The analysis reported in this paper also provides a normative value. Implementing field experiments to assess alternative investment policies could be very costly. Managers administering reward programs should apply this analysis first, in order to assess the profitability of their program. Using counterfactual experiments can be helpful in understanding the impact of the reward program and the viability of alternative investments.

CHAPTER 2

An Approach for Extending Dynamic Structural Models to Settings with Multi-Product Firms (joint with Aviv Nevo)

2.1. Introduction

Dynamic questions have long been a central part of the study of markets. However, for many years modeling and computational constraints limited our ability to structurally study dynamics. Ericson and Pakes (1995) outlined what has since become the standard framework for dynamic oligopolistic games. In principle, the parameters of the model, such as investment costs or sunk costs, can be estimated by comparison between the observed choices and those predicted by the model, following the "nested algorithm" (Rust 1987), which has been used successfully in single agent models. In practice, however, this approach is not computationally feasible when studying dynamic games because of the need to solve the equilibrium many times. More recently several alternatives have emerged based on ideas proposed in Hotz and Miller (1993) and Hotz et al. (1994).[1] A common feature of these new methods is that they avoid the use of computationally intense dynamic programming techniques to compute the equilibrium strategies, and instead estimate strategies directly from the choices observed in the data (Aguirregabiria and Mira 2007; Bajari, Benkard, and Levin 2007; Pakes, Ostrowsky, and Berry 2005; Pesendorfer and Schmidt-Dengler 2003).[2]

[1]See also Manski (1993) and more recently Aguirregabiria and Mira (2007).

[2]For a review of structural estimation of dynamic games see Ackerberg et al. (2005).

While our ability to estimate the dynamic model has significantly improved, in order to study counterfactual situations the equilibrium of the model still needs to be computed. Therefore, the original Ericson-Pakes model is somewhat limited in its application to cases where the state space is relatively small. In particular, the model is limited in its ability to study markets with multi-product firms. Consider, for example, the setup of Pakes-McGuire (1994). In this setup, single product firms compete each period by setting prices for differentiated products. Firms can invest in improving the quality of the products, where the outcome of the investment is stochastically increasing in the investment amount. In Pakes-McGuire each firm has a single product and the state variable is given by the "quality" of this product. If each firm produces many products, as is the case in essentially every differentiated products industry, then the state of each firm is a vector of quality of each of its products. So even though the model might be tractable for single product firms, it quickly becomes intractable for multi-product firms.[3] In this paper we propose a feasible approach to modeling multi-product firms in dynamic games.

2.2. The Model

We focus on the differentiated product version of the Ericson-Pakes model, detailed in Pakes-McGuire (1994).

2.2.1. Static Flow Profits

On the demand side, we assume that consumers choose either one of the J products offered in the market, or the outside good. A consumer will choose the product that gives

[3] Extending the core version to allow multi-product firms is also a concern listed in the agenda outlined by Pakes (2000), pg. 22.

the highest utility. The utility that consumer i derives from purchasing brand j at time t is

$$U_{ijt} = \delta_{jt} + \epsilon_{ijt},\ \delta_{jt} = x'_{jt}\beta - \alpha p_{jt} + \xi_{jt} \tag{2.1}$$

where δ_{jt} is the mean utility of product j in t; α and β are taste and price parameters, respectively; x_{jt} is a vector of observable characteristics of product j; p_{jt} is the price for brand j in period t; the term ξ_{jt} captures product- and time-specific shocks which are correlated across consumers; ϵ_{ijt} is the idiosyncratic error term. The consumers might also decide not to purchase any of the goods, in which case they choose the outside option that has a mean utility normalized to zero.

There are F firms in the market. Each firm f sells a subset of the J products denoted with $\mathcal{F}_f$. We define the quality, or efficiency level, of a product as $\omega_{jt} = x'_{jt}\beta + \xi_{jt}$, and the market structure of the industry at time t is characterized by a J-dimensional vector $s_t = (\omega_{1t}, ..., \omega_{Jt})$. The quantity sold and the optimal price will be a function of the efficiency levels of all of the firms' products and the state of the industry (i.e., the competitors' quality). Therefore, the static profit function of firm f can therefore be written as (dropping subscripts t):

$$\pi_f\left(\{\omega_j\}_{j\in\mathcal{F}_f}, s\right) = \sum_{j\in\mathcal{F}_f}\left[p_j\left(\{\omega_j\}_{j\in\mathcal{F}_f}, s\right) - mc_j\right]\mathcal{M}\sigma_j\left(\{\omega_j\}_{j\in\mathcal{F}_f}, s\right) - C_f, \tag{2.2}$$

where p_j and σ_j denote the price and market share of product j, mc_j and C_f are the marginal cost to produce product j and the fixed cost of production, $\mathcal{M}$ is the size of the market. We assume that firms set prices to maximize their profits and we assume the existence of a pure-strategy Bertrand-Nash equilibrium in prices.

2.2.2. Dynamic Decisions

In addition to pricing, in each period the firms decides whether to invest in each product and if so how much. Let x_j denote the investment in product j. Investment costs c per unit of investment, and its outcome is stochastic. Examples of investment are advertising or investment in research that is aimed in improving the quality of the product. We do not consider for now entry or exit, either at the firm or at the product level. Below, once we present our main results we show how to add these features to the model.

The investment decisions are made to maximize the value of the firm, given by

$$V_f\left(\{\omega_{j,1}\}_{j\in\mathcal{F}_f}, s_1\right) = \max_{(x_{j,t}>0, j\in\mathcal{F}_f)} \sum_{t=1}^{\infty} \beta^{t-1} E\left[\pi_f\left(\{\omega_{j,t}\}_{j\in\mathcal{F}_f}, s_t\right) - c\sum_{j\in\mathcal{F}_f} x_{j,t}\right] \tag{2.3}$$

where β is the discount rate.

The expectations are taken with respect to uncertainty about future quality levels, and competitors' actions. Let the CDF $P\left(\{\omega_{j,t+1}\}_{j\in\mathcal{F}_f}, s_{t+1} | \{x_{jt}\}_{j\in\mathcal{F}_f}, \{\omega_{jt}\}_{j\in\mathcal{F}_f}, s_t\right)$ represent firm f's beliefs about next period efficiency levels ($\omega_{j,t+1}$) and market structure (s_{t+1}), given current investments (x_{jt}), efficiency levels (ω_{jt}), and market structure (s_t). In every period each product's efficiency moves to the next level

$$\omega_{j,t+1} = \omega_{j,t} + (\nu_{j,t} - \zeta_t) \ , \tag{2.4}$$

where ν_{jt} and ζ_t are two independent, non-negative random variables. The first has a distribution that comes from a family $\{P(\cdot|x), x \in \mathcal{R}^+\}$ that is stochastically increasing in the investment level for that product, x_{jt}, and such that $\nu_{jt} = 0$ if $x_{jt} = 0$. The second is an exogenous random variable with probability $\mu(\zeta)$; in our setup it represents the efficiency value of the outside good, therefore it is a demand shock which is common across products.

The value of the firm is a function of its own state and the state of its competitors. Even if each firm has a single product, and there is a small number of firms, and the efficiency levels can take on a small number of values, then solving for the value function is subject to the curse of dimensionality. Pakes and McGuire (1994) propose to mitigate this problem somewhat by assuming exchangeability of the profit function such that the identity of the firms is not important. Therefore, only the number of firms at each efficiency level matters, not their identity. This significantly reduces the state space.

However, with multi-product firms an exchangeability assumption across products is both problematic and potentially not sufficient. To see why it might be problematic to assume exchangeability at the product level consider the following case. Suppose two firms produce an identical product. However, suppose one of them only produces this product, while the other firm produces other products as well. In general, the pricing of these firms will be different, because the second firm will internalize the substitution to its other products, and therefore assuming the products are exchangeable in the profit function is problematic. Further, even if we are willing to assume exchangeability, the state space will still be very large and probably not computationally tractable.

2.3. Results

In the section we propose an approach that makes the model tractable. Our solution we rely heavily on what we will call an *adjusted inclusive value* (henceforth AIV) defined as

Definition 1. *Let* $i_f = log\left[\sum_{r\in\mathcal{F}_f} exp\left(\omega_r - \alpha mc_r\right)\right]$ *be the adjusted inclusive value of firm f.*

The adjusted inclusive value is the difference between the quality of each product, defined by the characteristics, and the marginal cost needed to produce the quality level of each product. It can be therefore interpreted as the net quality level that the firm is able to produce in the market. The AIV is closely related to the inclusive value (McFadden,1978), which captures the expected utility for a consumer from several products prior to observing the random variables ϵ_{ij}'s. From the firm's perspective this inclusive value needs to be adjusted to take account of different marginal costs of production. Indeed, as we will now show under some assumptions, the AIV is all that we need to compute the static profits.

Assumption A1: The idiosyncratic error term e_{ijt} in (2.1) is identically and independently distributed type I extreme value.

Assumption A1 implies that the demand is given by the Logit (McFadden 1974). In particular it implies market shares of the form

$$\sigma_j\left(p; \{\omega_j\}_{j\in\mathcal{F}_f}, s\right) = \frac{\exp(\omega_j - \alpha p_j)}{1 + \sum_{k\in\mathcal{F}_f}\exp(\omega_k - \alpha p_k)}.$$

It is well-known that this model has several unattractive features (for example, see McFadden (1978); or Berry Levinsohn and Pakes (1995). However, this assumption will turn out to be extremely useful for us. Below we discuss ways to relax this assumption somewhat. We note that this assumption is very similar to the assumptions made by both Pakes-McGuire and the literature cited in the Introduction.

Lemma 1. *Under Assumption A1* $\pi_f\left(\{w_j\}_{j\in\mathcal{F}_f}, s\right) = \pi_f\left(i_f, sf\right)$*, where* $sf = (i_1, ..., i_F)$.

Proof: Taking the first-order condition of the profit function for firm f, as defined in (2.2), with respect to product j's price, we get

$$p - mc = \Omega^{-1}\sigma\left(p; \{w_j\}_{j\in\mathcal{F}_f}, s\right) \tag{2.5}$$

where $\sigma(\cdot)$, p, and mc are $J \times 1$ vectors of market shares, prices, and marginal cost, respectively, and Ω is a $J \times J$ matrix with the element Ω_{jr} equal to $-\partial\sigma_r/\partial p_j$ if j and r are produced by the same firm, 0 otherwise. Given Assumption A1, the derivatives of the share equations are $\partial\sigma_j/\partial p_j = -\alpha\sigma_j(1-\sigma_j)$ and $\partial\sigma_r/\partial p_j = \alpha\sigma_j\sigma_r$. Plugging these back into equation (2.5) yields

$$(p - mc)_f = \frac{1}{\alpha\left(1 - \sum_{r\in\mathcal{F}_f}\sigma_r\right)} = \frac{1}{\alpha\left(1 - \bar{\sigma}_f\right)}. \tag{2.6}$$

where $\bar{\sigma}_f = \sum_{r\in\mathcal{F}_f}\sigma_r$ is firm f's total share. This equation implies that each firm applies the same markup to all of its products. In order to compute the profits we need to compute the share of each firm.

We now show that, given this pricing rule, the share of firm f can be computed knowing only the firms' AIV.

$$\bar{\sigma}_f = \sum_{j\in\mathcal{F}_f} \sigma_j = \sum_{j\in\mathcal{F}_f} \frac{exp\left(\delta_j\right)}{1+\sum_{r=1}^{J} exp\left(\delta_r\right)} =$$

$$= \sum_{j\in\mathcal{F}_f} \frac{exp\left(-\alpha\left(p_j - mc_j\right)\right) exp\left(\omega_j - \alpha mc_j\right)}{1+\sum_{r=1}^{J} exp\left(-\alpha\left(p_r - mc_r\right)\right) exp\left(\omega_r - \alpha\, mc_r\right)}.$$

Since firms apply the same markup on each of their own products,

$$= exp\left(-\alpha\, markup_f\right) \sum_{j\in\mathcal{F}_f} \frac{exp\left(\omega_j - \alpha\, mc_j\right)}{1+\sum_{g=1}^{F} exp\left(-\alpha\, markup_g\right) \sum_{r\in\mathcal{F}_g} exp\left(\omega_r - \alpha\, mc_r\right)}$$

$$= \frac{exp\left(i_f - \alpha\, markup_f\right)}{1+\sum_{g=1}^{F} exp\left(i_g - \alpha\, markup_g\right)}$$

The result above shows that firms' shares are function of the AIV. Therefore, substituting the markup computed in equation (2.6) into the profit defined in equation (2.2) we get

$$\pi_f\left(\{w_j\}_{j\in\mathcal{F}_f}, s\right) = \mathcal{M}\frac{\bar{\sigma}_f\left(i_f, s_f\right)}{\alpha(1-\bar{\sigma}_f\left(i_f, s_f\right))} - C_f = \pi_f\left(i_f, sf\right) \qquad Q.E.D.$$

The Lemma shows that under Assumption A1 the static flow profits can be written as a function of firm level AIV, and do not require the product specific quality levels. In order to show that the firm's dynamic problem also does not require the product level quality we need to make an additional assumption.

Assumption A2: $P\left(i_{f,t+1}, sf_{t+1} |\, \{x_{jt}\}_{j\in\mathcal{F}_f}, \{\omega_{jt}\}_{j\in\mathcal{F}_f}, s_t\right) = P\left(i_{f,t+1}, sf_{t+1} | x_{ft}, i_{ft}, sf_t\right)$.

This assumption restricts the stochastic evolution of the states. It also restricts the way that investment decisions can be made. We now can write our main result.

Proposition 1. *Under Assumption A1 and A2* $V_f\left(\{w_{j,1}\}_{j\in\mathcal{F}_f}, s_1\right) = V_f\left(i_{f,1}, sf_1\right)$, $\forall f \in \{1, ..., F\}$.

Proof. Substituting the result of Lemma 1 into equation (2.3) we get

$$\begin{aligned} V_f\left(\{\omega_{j,1}\}_{j\in\mathcal{F}_f}, s_1\right) &= \\ &= \max_{x_{f,t}>0}\sum_{t=1}^{\infty}\beta^{t-1}E\left[\pi_f\left(i_{f,t}, sf_t\right) - cx_{f,t}\right] \\ &= \max_{x_{f,t}>0}\sum_{t=1}^{\infty}\beta^{t-1}\int\left[\pi_f\left(i_{f,t}, sf_t\right) - cx_{f,t}\right]dP\left(i_f, s_t \middle| \{x_{j1}\}_{j\in\mathcal{F}_f}, \{\omega_{j1}\}_{j\in\mathcal{F}_f}, s_1\right) \end{aligned}$$

By Assumption A2

$$= \max_{x_{f,t}>0}\sum_{t=1}^{\infty}\beta^{t-1}\int\left[\pi_f\left(i_{f,t}, sf_t\right) - cx_{f,t}\right]dP\left(i_{f,t}, s_t | x_{f1}, i_{f1}, sf_1\right) = V_f\left(i_f, sf_1\right) \qquad Q.E.D.$$

What we have shown is that given our assumptions the state variables of the problem include only firm level variables and do not require knowing, and keeping track of, the product-level variables. This result allows us to consider firms that produce many brands without carrying the demand of each single brand, which would make the dynamic multi-product firm problem unfeasible.

2.4. Extensions and Application

There are several ways to relax Assumptions A1 and A2 and still get some of the benefits of our approach. Assumptions A1 can be relaxed by assuming a generalized extreme value distribution. As a special case, in the next section we consider the Nested Logit model. In this case we will need one state variable per firm per nest to compute the flow profits. Our approach will not work for the more general Mixed Logit model. In this case, the markup of a product depends on its share for each consumer relative to other consumers. Since different products are likely to generate different shares across consumers, the firm will not apply the same mark-up for each product, and equation (2.6) will not hold. Assumption A2 can also be somewhat relaxed by allowing for other variables to enter the transition probabilities.

2.4.1. Nested Logit Demand

In this section we extend the results to the case when we allow consumers' tastes to be correlated across products according to a *a priori* specification: products are grouped into L exhaustive and mutually exclusive locations, $l = 1, 2, ...L$, where $\mathcal{L} = \{1, ..., L\}$; each location identifies the subset of products in the market with similar combination of attributes. Let $\Im_l$ be the set of products in location l, where $\#\Im_l = J_l$. Also, let product $j \in l$.

Definition 2. *Let* $i_{fl} = log\left[\sum_{r\in\mathcal{F}_f\cap\Im_l} exp\left(\omega_r - \alpha mc_r\right)\right]$ *be the adjusted inclusive value of firm* f *in location* l*.*

Assumption A1N: The idiosyncratic error term in (2.1) is $\epsilon_{ijt} = \zeta_{ilt} + (1-\lambda)\,\upsilon_{ijt}$, where υ_{ijt} is identically and independently distributed type I extreme value.

In particular, ζ_{ijt} is a consumer taste's shock, which is common across products in location l, and λ is a parameter of the density of ϵ_{ijt}, $f(\cdot, \lambda)$.

Let s_l be the market structure within location l, i.e. $s_l = (\omega_{1l}, ..., \omega_{J_l l})$. Assumption A1N implies market shares of the form

(2.7)
$$\begin{aligned}\sigma_j\left(p; \{\omega_{jl}\}_{\substack{j\in\mathcal{F}_f\\ l\in\mathcal{L}}}, \{s_l\}_{l\in\mathcal{L}}\right) &= \sigma_l\left(p; \{\omega_{jl}\}_{\substack{j\in\mathcal{F}_f\\ l\in\mathcal{L}}}, \{s_l\}_{l\in\mathcal{L}}\right)\cdot\sigma_{j|l}\left(p; \{\omega_j\}_{j\in\mathcal{F}_f\cap\Im_l}, s_l\right)\\ &= \frac{exp\left(\frac{R_l}{\mu_1}\right)}{\sum_{s=1}^{L} exp\left(\frac{R_m}{\mu_1}\right)}\cdot\frac{exp\left(\frac{\omega_j-\alpha p_j}{\mu_2}\right)}{1+\sum_{k\in l} exp\left(\frac{\omega_k-\alpha p_k}{\mu_2}\right)},\end{aligned}$$

where $R_m = \mu_2 ln \sum_{r\in m} exp\left(\frac{\omega_r-\alpha p_r}{\mu_2}\right)$ is the inclusive value of location m.

Lemma 2. *Under Assumption A1N* $\pi_f\left(\{w_j\}_{j\in\mathcal{F}_f}, s\right) = \pi_f\left(\{i_{fl}\}_{l\in\mathcal{L}}, \{sf_l\}_{l\in\mathcal{L}}\right)$, *where* $sf_l = (i_{1l}, ..., i_{Fl})$.

Proof: Solving for the first-order condition of the profit function with respect to product j's price we can compute the optimal mark-up for each firm in each location

(2.8)
$$(p-mc)_{fl} = \frac{\mu_2 + \left(\frac{\mu_2}{\mu_1}\right)\sum_{\substack{m\in\mathcal{L}\\ m\neq l}}(p-mc)_{fm}\,\sigma_m\left(p; \{\omega_{jm}\}_{\substack{j\in\mathcal{F}_f\\ m\in\mathcal{L}}}, \{s_m\}_{m\in\mathcal{L}}\right)\sigma_{f|m}\left(p; \{\omega_j\}_{j\in\mathcal{F}_f\cap\Im_m}, s_m\right)}{1-\left[\frac{1}{\mu_2}-\frac{1}{\mu_1}+\frac{1}{\mu_1}\sigma_l\left(p; \{\omega_{jl}\}_{\substack{j\in\mathcal{F}_f\\ l\in\mathcal{L}}}, \{s_l\}_{s\in\mathcal{L}}\right)\right]\sigma_{f|l}\left(p; \{\omega_j\}_{j\in\mathcal{F}_f\cap\Im_l}, s_l\right)}$$

where $\sigma_{f|l}(\cdot) = \sum_{r\in\mathcal{F}_f\cap\Im_l}\sigma_{r|l}(\cdot)$ is the share of firm f conditional on location l. This equation implies that each firm applies the same markup to all of its products that are in the same location.

Now we show that the share of firm f can be computed knowing only the firms' AIV in each segment. Notice that the inclusive value of a location can be rewritten as $R_m = \mu_2 ln \sum_{f=1}^{F} \sigma_{f|m}(\cdot)$. The share of all products in location l can be written as

$$\sigma_l\left(p;\{\omega_{jl}\}_{\substack{j\in\mathcal{F}_f\\ l\in\mathcal{L}}},\{s_l\}_{l\in\mathcal{L}}\right)=\frac{\left[\sum_{f=1}^{F}\sigma_{f|l}\left(p;\{\omega_j\}_{j\in\mathcal{F}_f\cap\Im_l},s_l\right)\right]^{\frac{\mu_2}{\mu_1}}}{\sum_{m\in\mathcal{L}}\left[\sum_{f=1}^{F}\sigma_{f|m}\left(p;\{\omega_j\}_{j\in\mathcal{F}_f\cap\Im_m},s_m\right)\right]^{\frac{\mu_2}{\mu_1}}}; \tag{2.9}$$

the share of firm f conditional on location l is

(2.10)

$$\sigma_{f|l}\left(p;\{\omega_j\}_{j\in\mathcal{F}_f\cap\Im_l},s_l\right)=\sum_{j\in\mathcal{F}_f\cap\Im_l}\frac{exp\left(\omega_j-\alpha p_j\right)}{1+\sum_{g=1}^{F}\sum_{r\in\mathcal{F}_g\cap\Im_l}exp\left(\omega_r-\alpha p_r\right)}$$

$$=-exp\left(\alpha\cdot\text{markup}_{fl}\right)\sum_{j\in\mathcal{F}_f\cap\Im_l}\frac{exp\left(\omega_j-\alpha mc_j\right)}{1+\sum_{g=1}^{F}-exp\left(\alpha\cdot\text{markup}_{gl}\right)\sum_{r\in\mathcal{F}_g\cap\Im_l}exp\left(\omega_r-\alpha mc_r\right)}$$

$$=\frac{exp\left(i_{fl}-\alpha\text{markup}_{fl}\right)}{1+\sum_{g=1}^{F}exp\left(i_{gl}-\alpha\text{markup}_{gl}\right)}.$$

Substituting into the profit function we get

(2.11)

$$\pi_f\left(\{w_{jl}\}_{\substack{j\in\mathcal{F}_f\\ l\in\mathcal{L}}},\{s_l\}_{l\in\mathcal{L}}\right)=$$

$$=\mathcal{M}\sum_{l\in\mathcal{L}}\frac{\mu_2+\left(\frac{\mu_2}{\mu_1}\right)\sum_{\substack{m\in\mathcal{L}\\ m\neq l}}\left(p_m-mc_m\right)\sigma_m\left(\{i_{fm}\}_{m\in\mathcal{L}},\{sf_m\}_{m\in\mathcal{L}}\right)\sigma_{f|m}\left(i_{fm},sf_m\right)}{\frac{1}{\sigma_{f|l}\left(i_{fl},sf_l\right)}-\left(\frac{1}{\mu_2}-\frac{1}{\mu_1}+\frac{1}{\mu_1}\sigma_l\left(\{i_{fl}\}_{l\in\mathcal{L}},\{sf_l\}_{l\in\mathcal{L}}\right)\right)}$$

$$=\pi_f\left(\{i_{fl}\}_{l\in\mathcal{L}},\{sf_l\}_{l\in\mathcal{L}}\right)\qquad Q.E.D.$$

Under assumption A1N and A2, it can be shown analogously to Proposition 1 above that $V_f\left(\{w_{j,1}\}_{j\in\mathcal{F}_f},s_1\right)=V_f\left(\{i_{fl,1}\}_{l\in\mathcal{L}},\{sf_{l,1}\}_{l\in\mathcal{L}}\right),\forall f\in\{1,...,F\}$.

CHAPTER 3

An Empirical Model of Dynamic Attribute-Space Competition

3.1. Introduction

Every year multi-product firms launch in the market thousands of new products. Novel combinations of attributes and adequate marketing strategies help position the new products in specific locations in the space defined by product attributes. One of the most important decisions that managers face when launching a new product is which combination of attributes to choose and where to locate the product in the multi-attribute space. Several factors come into play in this decision, such as marginal costs associated with producing the new combination of attributes, fixed (sunk) costs of launching the new product, potential demand for the specific combination of attributes offered, and competitive reaction. Multi-product firms face a further issue that has to do with the interdependence between the new product and existing ones.

To extend their product line multi-product firms can follow two different strategies. The can opt for an *offensive* or a *defensive strategy.* In the first case, firms introduce a new product with a combination of attributes such that its location in the attribute space is far from the location of the other existing products of the firm. The distance in terms of attributes determines a weak substitution effect between the new and the existing products that decreases the risk of cannibalization. This strategy, also called *interlacing strategy* ((Bhatt 1987)), leads to an expansion effect (Shaked and Sutton 1990): by expanding its presence in the attribute-space area, the firm offers a wider variety of products and reaches a larger demand spectrum.

According to the second strategy, firms introduce a new product with a combination of attributes such that its location is close to that of the other existing products of the firm. The downside of cannibalization in demand, which now is very likely to occur between sales of new and existing products, is compensated for by advantages in terms of production costs due to increasing economies of scale. Because of location proximity, the new product can also benefit from the spillover of the reputation and consumer loyalty previously created by existing products. Furthermore, consumer tastes are better known, so uncertainty in demand is reduced. Finally, this strategy provides a means of defending the profitability of existing brands from the competition, as the introduction of new products might preempt the surrounding multi-attribute space or reinforce the brand awareness of the firm.[1]

Previous research in marketing and economics has mostly considered product attributes as exogenously given, ignoring firms' product assortment choice: few theoretical studies have moved away from such an assumption, investigating the equilibria arising from multi-product firms' spatial competition. Brander and Eaton (1984) have modeled competition between two firms offering two products, and have shown that both offensive and defensive strategies can emerge as optimal strategies. Martinez-Giralt and Neven (1988) extend this result by showing that if the attribute-space is continuous instead of

[1]This motivation has been explored theoretically by a number of papers (see Schmalensee (1978), Hay (1976), Prescott and Visscher (1977), Judd (1985)), and has also recently been the focus of several empirical works (see Ellison and Ellison (2000), Dafny (2003), Goolsbee and Syverson (2004).

discrete, the defensive strategy will prevail over the offensive. Recently, a number of empirical papers in the literature on new product entry have applied spatial competition to uncover firms' optimal location decision. Mazzeo (2002) estimates a static location model where firms optimally decide the quality of their service and shows that, on U.S. highways, motels use vertical differentiation to lessen price competition. Using the incomplete information framework to deal with the size of the choice set, Seim (2005) assesses the importance of geographic differentiation in video rental industry. Adopting a similar model, Zhu, Singh, and Manuszak (2005) investigate the determinants of store format choice in the retail discount store industry. These studies show that by endogenizing firms' location choice researchers can develop a better understanding of the strategies used in the market; firms may trade off the opportunity for high demand, so the introduction of a product into a less profitable location may be beneficial for the profit of the whole product portfolio. Including attribute-space competition in the model, also improves the reliability of counterfactual experiments: instead of simply reoptimizing the price of their product, firms can also optimally reallocate the new products in the attribute space. In the studies above, however, the profit functions are expressed in reduced form, so the firm's behavior is not explicitly modeled and the counterfactual exercise is therefore limited.

In this paper, I propose an empirical model to investigate new product assortment strategies. I estimate a model that combines pricing decisions with attribute-space location choice into a single structural framework; each multi-product firm competes in price and chooses the best attribute-space location for its new products. This model combines and extends two well-known frameworks: (1) I use a static discrete choice demand system and assume an oligopolistic Bertrand-Nash equilibrium to estimate the parameters of demand and marginal costs (Bresnahan 1987); (2) I use an oligopolistic model of industry dynamics to estimate the fixed costs of product positioning (Ericson and Pakes 1995). By using a structural dynamic model of multi-product firms that compete in price and introducing new products in the attribute space, I can evaluate the impact of market primitives on the firm's location choice. After I recover the parameters from the static and dynamic model and identify all the primitives in the market, I focus on a counterfactual exercise to evaluate the impact of firms' positioning costs on competitors profitability and market competition level.

I apply the model to the U.S. ready-to-eat cereal industry. This market is particularly interesting for two reasons: first, in the past firms have been accused of exercising market power, and strategically using new product entry to defend this competitive advantage.[2] Using this framework, I will assess if there is indeed a *defensive strategy* and the conditions in the market that lead to this strategy. Second, the high number of new products that are typically offered in this category facilitates the estimation of the dynamic model.

The paper is organized as follows: in the next paragraph I briefly present some stylized facts from this industry to motivate the study of new product location choice. In section 3.2, I describe the model. In section 3.3 I review the estimation strategy I will follow to estimate the parameters of the dynamic model. Finally, I discuss the results and conclude.

[2]See (Scherer 1982) and (Nevo 2001) for details.

3.1.1. Product Launch and Positioning in the U.S. Ready-to-Eat Breakfast Cereal Industry

In 1988 there are six major competitors in the market: Kellogg and General Mills are the leaders holding more than 60% of the market. Together with Post, the third biggest company, they hold almost 75% of the total market share. Then Quaker, Nabisco and Ralston Purina follow. From 1988 to 1997, Kellogg introduces 16 new brands, General Mills 18, and Post eight[3]. Quaker launches one product only, and Nabisco two. Ralston introduces 25 new products, but almost all of them are limited editions and are scrapped soon after their introduction. Also, Nabisco and Ralston will soon be merged with Post (1993) and General Mills (1996), respectively. Therefore, in what follows I will consider only the three leading firms.

By observing new entries in the market during the period between 1988 and 1997, two main observations can be made. First, it is clear that product launches represent an important activity for this industry. This is confirmed by the following observations. Figures B.2-B.4 report the shares in each quarter for each firm with (solid line) and without (dashed line) the inclusion of the new brands launched after 1988. On average the share of the new brands across time represent 6.7% of the total share for Kellogg, 17.0% for General Mills, and 11.8% for Post. Second, when firms launch new products, they do not cover the attribute space evenly, rather they tend to populate the areas where their presence with existing brands is stronger, following the defensive strategy discussed above. To show this, I partition the market into three segments, depending

[3] I keep out of the analysis oatmeal muesli and granola cereals.

on the ingredients of their products[4]: cereals with added sugar (Kids' Segment), cereals with whole grain (Family/Wholesome Segment), cereals with added fruits and nuts (Taste Enhanced Segment). I follow this segmentation to emphasize the asymmetric positioning of existing brands in the segments by the leading firms. In particular, Figures B.5-B.6 shows that Kellogg and General Mills' product portfolios have a stronger relative presence in the first and second segments as compared to the third segment: Kellogg owns more than 50% of the shares in both of the first two segments and only 20% in the third segment; General Mills has roughly 35% in the first two segments and 20% in the third segment. For Post, the situation is reversed. Compared to its competitors it has a stronger presence in the third segment (with 60% as contrasted with 10% for Kellogg and General Mills). It is clear, by looking at the positioning of the products launched (Figure B.7), that Kellogg and General Mills, as contrasted with Post, tend to prefer the first and second segments. To confirm this observation, I ran a simple logit where I conditioned the segment choices on the firm share level in that segment: the results in Table B.1 report a significant correlation, that confirms the defensive strategy of firms in this market. In what follows, I will investigate *why* firms use a defensive strategy. To answer this question I will estimate a dynamic model of new product location, presented in the next section.

3.2. The Model

3.2.1. Overview

To investigate the impact of competition on new product launch, I will adopt the theoretical framework for dynamic oligopoly proposed by Ericson and Pakes (Ericson and Pakes

[4]To obtain information on the ingredients of the cereals I referred to Gitlin and Ellis (2005).

(1995), henceforth EP). In oligopolistic markets, firms' strategies are defined over a set of different actions, like pricing, advertising, product portfolio management, and other investments; a dynamic model that accounts for all these decisions simultaneously would be computationally intractable. One key idea in the EP approach is to distinguish static decisions from dynamic decisions and solve the static problem before proceeding with the dynamic game; the decisions that enter the static game are then passed to the dynamic problem only through their optimal values, so that they do not add to the computational burden of the dynamic program. In this application I follow the model presented in Chapter 2 of this dissertation. According to this approach, firms in each period play a static Bertrand-Nash equilibrium in prices, so by estimating price elasticities from static demand I can recover the marginal costs of the industry. Then these results are passed into the dynamic game where firms decide whether to launch and how to position new products in the market. Using this approach I can relax the EP assumption that firms can only sell one product, and instead allow for firms to sell multiple products in the same period.

Bringing the EP theoretical framework to the data is known to be difficult because the numerical solution of the dynamic game is computationally intensive.[5] Instead of computing the equilibrium strategies, I follow a novel approach recently introduced in the literature. Hotz and Miller (1993) and Hotz, Miller, Sanders and Smith (1994) propose a method for single-agent dynamic models where optimal strategies are estimated directly from the observed choices. Recently, several papers such as Aguirregabiria and

[5]The numerical solution of the dynamic game is nested into the estimation process, so it must be computed at each iteration of the estimation algorithm (Rust 1987). Also, another potential problem of the EP framework is that the equilibrium might not exist or be unique. Doraszelski and Satterthwaite (Doraszelski and Satterthwaite 2003) overcome nonexistence by introducing firms' private information.

Mira (2007), Bajari, Benkard and Levin (2007), Pakes, Ostrovsky and Berry (2005), and Pesendorfer and Schmidt-Dengler (2003), have extended this idea to the estimation of dynamic games. Using the approach suggested by Aguirregabiria and Mira, I can perform a search for the fixed-point solution on the conditional choice probability space; their algorithm overcomes the potential bias of the two-stage estimation used by the other methods listed above.

In the next section I report the results from the static Bertrand-Nash pricing equilibrium framework. In section 3.2.3 I present each component of the dynamic game in detail.

3.2.2. Static environment

There are F firms in the market, indexed by $f \in \{1, 2, ..., F\}$. Each firm f sells J_f products[6], indexed by $j \in \mathcal{F} = 1, ..., J_f$; $J = \{J_f\}_{f=1}^{F}$. The market is exogenously partitioned into L locations, indexed by $l \in \mathcal{L} = \{1, 2, ...L\}$; each location represents the set of products with a specific combination of attributes. The probability that consumer i chooses brand j at time t is

$$U_{ijt} = x_j'\beta - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt} \tag{3.1}$$

[6] I use the word product and brand interchangeably without any distinction.

where x_j is a vector of observable characteristics of product j[7], p_{jt} is the price for brand j in period t, the term ξ_{jt} captures product- and time-specific shocks which are correlated across consumers, and ϵ_{ijt} is the idiosyncratic error term. The utility derived from choosing the outside good is $U_{i0t} = \epsilon_{i0t}$. Assuming the distribution of ϵ is $G(\epsilon)$, and the ϵ's are i.i.d. across consumers, the market share of brand j in period t is

$$\sigma_{jt} = \int_{\mathcal{A}} dG(\epsilon) \tag{3.2}$$

where $\mathcal{A}$ is the set of values for ϵ that induce the choice of product j in t. For my application, I assume that the ϵ's follow an extreme value distribution.

Given this assumption about the idiosyncratic errors, which corresponds to assumption 2.3 in Chapter 2, the profit equation can be rewritten as a function of the AIV of each firm in the market, as shown in Lemma 1:

$$\pi_f\left(\{\omega_f\}_{f\in F}\right) = \mathcal{M}\frac{\bar{\sigma}_f\left(\{\omega_f\}_{f\in F}\right)}{\alpha(1-\bar{\sigma}_f\left(\{\omega_f\}_{f\in F}\right))} - C_f \tag{3.3}$$

where ω_f is the adjusted inclusive value (henceforth AIV) of firm f as defined in Definition 1 of Chapter 2, C_f is the fixed cost of production, and $\mathcal{M}$ is the size of the market. Each AIV is derived from the difference between the quality, defined by the brand-specific characteristics, and the marginal cost needed to produce that quality level. It can also be interpreted as the net quality level that the firm is able to offer on the market. It is a

[7] A constant term in x_j captures the average valuation that the consumer assigns to all unobserved (by the econometrician) product components relative to the outside good. It is well known that the choice problem of an individual depends on the difference in utility rather than on the absolute levels; so the value of an alternative entering the utility specification can only be defined in relation to the value of another alternative. It is because of this and the fact that we include the outside good, i.e. the alternative of not buying any product, that we can model the aggregate demand for products as a function of prices and products' characteristics.

sufficient statistic in the sense that at any point in time by solely knowing the vector ω, I can recover the profit function of the industry. This result is crucial because it allows me to consider firms that produce many brands without carrying the demand for every single brand, which would make the dynamic problem unfeasible.

3.2.3. Dynamic environment

3.2.3.1. State Space.

At the beginning of each period t, with $t \in \{1, 2, ...\}$, every firm f is characterized by two vectors, x_{ft} and ϵ_{ft}, that determine firms' profitability. The first vector represents state variables which are common knowledge among firms and can be seen by the researcher; in this application they are represented by the firm's current AIV, i.e. $x_{ft} = \omega_{ft}$. The current levels of AIV, which come from the static pricing Bertrand-Nash equilibrium, determine the profit of each firm in the market; there are in total one AIV for each firm in the market[8]. The second vector, ϵ_{ft}, represents firm f's private information: this second component captures all those idiosyncratic variables that impact on firms' profits but cannot be seen by firms nor by the researcher. An example would be particularly effective market research for the launch of a new product, that idiosyncratically informs the management on the true quality value of the new product. Let $x_t = \{x_{1t}, ..., x_{Ft}\}$ be the vectors of common knowledge and $\epsilon_t = \{\epsilon_{1t}, ..., \epsilon_{Ft}\}$ the vectors of private information: the state variable $s_t = \{x_t, \epsilon_t\}$ represents the state of the market at time t.

[8] If a correlation in the errors across brands is assumed, and therefore a nested logit demand system is considered, I would have instead one AIV per nest per firm (see section 2.4.1 in chapter 2).

3.2.3.2. Actions and Transition Probabilities.

In each period t, firms simultaneously decide in which location of the market to launch a new brand, so firm f's choice in period t is represented by $a_{ft} \in \mathcal{A} = \{1, 2, ...L\}$, and the vector a_t is the profile of firms' decisions at time t.

It is clear from the static demand analysis above that the choice of a firm f to launch a product in location l has effects on its AIV level, w_f. A new brand introduced into the market, be it successful or not, will always have a positive market share, either by stealing sales from products in the same category or by expanding the category sales. As a result, by launching a new product in location l, a firm moves the market from the current state s to a new state s', where the AIV level for that firm is higher; in the new state s' the firm reaches higher profits.

Let $\Delta\omega_{fl}$ be the increase in a firm AIV, ω_f, due to a new brand launched in location l by firm f. This value is not perfectly known because firms are uncertain about how successful a new product is, and therefore how much increase in share they can realize from launching a new product. Instead, they develop expectations about the increase for each location in the market, that are common knowledge in the market. My ability to recover the fixed costs of entry depends on these expectations. In order to estimate their values, I first compute for each firm the AIV generated by each new product separately, and then I average across products of the same firm in the same segment:

$$\Delta\omega_{fl} = \frac{\sum_{\substack{k \in \mathcal{F} \\ k \in l}} \hat{\omega}_{kl}}{\sum_k 1\left(k \in \mathcal{F}, k \in l\right)} \tag{3.4}$$

so the value $\Delta\omega_{fl}$ is the increase in AIV that firm f expects to realize by launching a new product in segment l.

When brands leave the market, firms decrease their AIV value. I consider the product exit as exogenous, because its inclusion into the action space would increase the number of possible choices from L to $2 \cdot L$.[9] I assume that firms incur a stochastic decrease in their AIV levels.

The probability of moving from one state to another is commonly known among firms and is determined by a first-order Markov process as follows[10]:

$$\omega_{ft} = \omega_{ft-1} + \nu_{ft} \ , \tag{3.5}$$

where

$$\nu_{ft} = \begin{cases} \left(\Delta\omega_{fl} + \Delta\tilde{\omega}_f\right) & \text{with probability } \frac{\Delta\omega_{fl}}{\Delta\omega_{fl}+\Delta\tilde{\omega}_f} \\ -\left(\Delta\omega_{fl} + \Delta\tilde{\omega}_f\right) & \text{otherwise} \end{cases}$$

and

$$\Delta\tilde{\omega}_f = \frac{1}{L}\sum_{l=1}^{L} \Delta\omega_{fl} \cdot \hat{P}_f\left(exit\right) \tag{3.6}$$

and $\hat{P}_f\left(exit\right)$ is the probability that a brand of firm f is scrapped. $\Delta\tilde{\omega}_f$ accounts for the decrease in AIV that on average firms incur periodically by scrapping their products.

3.2.3.3. Timing of the Game.

In each period, the timing of the game is as follows:

(1) firms observe the state variable s_t, i.e. they observe the common knowledge vector state x_t and privately observe the idiosyncratic information ϵ_t.

[9] Moreover, many times the decision to scrap an existing brand is not associated with the decision to launch a new product, so that would require also modeling the choice of no entry.

[10] The result of the firms' choices is stochastic because I need to consider the exit of brands that occurs exogenously.

(2) firms statically set prices according to the Bertrand-Nash equilibrium, and choose in what location to launch the new product;

(3) consumers choose the cereal brands that maximize their utility, and demand evolves;

(4) firms receive profits.

Notice that firms make the decision of new product location in t and introduce the new product in $t+1$. Figure B.1 shows the timing of the model:

3.2.3.4. Equilibrium Concept.

A problem with dynamic games is that the set of Nash equilibria is generally unbounded. In the literature, researchers usually refine the set of equilibria by invoking the Markov Perfect Equilibrium concept (see Maskin and Tirole (2001) for more details), where agents base their decisions only on past information related to current payoffs and this information is summarized in state variables. In my model, firm f's strategy is a function $\sigma_f : X \times \mathcal{R}^{L+1} \rightarrow \mathcal{A}$. The use of Markovian strategies is also useful because it can significantly reduce the size of the state space. Because of the Markovian assumption, we can drop the period notation and denote the old and new state with s and s'. The value function of firm f, conditional on firms playing strategy σ, is:

$$V_f\left(s|\sigma\right) = \sup_{a_f \in \mathcal{A}} \left\{ \Pi_f\left(a_f, s\right) + \delta \int V_f\left(s'\right) Pr\left(ds'|s, \sigma\right) \right\} \tag{3.7}$$

where the static profit of firm f is $\Pi_f\left(a_{ft}, s_t\right) = \pi_f\left(s_t\right) + \sum_{l=1}^{L} I\left\{a_f t = l\right\} FC_l$; $I\left\{\cdot\right\}$ is the indicator function, and FC_l is the fixed cost of launching the product in location l and is a parameter to be estimated. Notice that the parameters that need to be estimated in the dynamic model enter linearly in the profit function, so the model can benefit from the *Separability in Dynamic Parameters* property (Aguirregabiria and Mira (2007), Bajari Benkard and Levin (2007)). In particular, the maximization of the likelihood of the dynamic model is equivalent to the maximization of a simple logit model, where the likelihood function is globally concave and the maximum can be easily found. Also, with this property the algorithm for finding the value function is faster because the inversion of the transition matrix, which is the most intense operation, is required only once for each given vector of parameters.

3.3. Data and Empirical Strategy

3.3.1. Data

I use scanner panel data from two different sources: IRI and Dominick's Finer Foods (henceforth DFF). The first database provides aggregate information on prices and quantities sold in 65 U.S. cities from the first quarter of 1988 to the last quarter of 1992. I use data on 24 brands with the highest average national market share and 20 new brands introduced during the five-year period. The second database provides aggregate information on prices and quantities sold from the DFF chain in the Chicago area from 1989 to 1997.[11] I aggregate the weekly data across stores up to the quarter level, and I retrieve information for the same 24 leading brands in IRI and for 24 new brands that are introduced between 1993 and 1997.

[11]Two quarters of data are missing in 1995.

3.3.2. Conditional Choice Probabilities

In a seminal work, Hotz and Miller (1993) showed that to estimate the structural parameters of a dynamic model one can avoid using the value function to fully solve for the equilibrium strategies, and exploit instead the one-to-one mapping between normalized value functions and conditional choice probabilities. The conditional choice probability of Hotz and Miller is $P : X \times \mathcal{A} \rightarrow [0,1]$, i.e. the probability that, conditional on being an observable state x_t, firms choose action profile a_t. The probability of firm f is given by

$$P_f\left(a_t|x_t\right) = \int I\left\{\sigma_f\left(x_t, \epsilon_{ft}\right) = a_{ft}\right\} g\left(\epsilon_{ft}\right) d\epsilon_{ft} \tag{3.8}$$

where $I\left\{\cdot\right\}$ is the indicator function, and $g\left(\cdot\right)$ is the probability density function of ϵ. In order to identify the parameters of the dynamic model, I place the following restrictions on the unobserved states of the primitives[12]:

***Assumption 1*:** *(Additive Separability)* Private information enters the static profit function additively, i.e. $\Pi\left(a_{ft}, \omega_t, \epsilon_{ft}\right) = \pi\left(a_{ft}, \omega_t\right) + \epsilon_{fs}\left(a_{ft}\right)$.

***Assumption 2*:** *(Conditional Independence)* The transition probability can be expressed as $P\left(\omega_{t+1}, \epsilon_{ft+1}|a_{ft}, \omega_t, \epsilon_{ft}\right) = P_\epsilon\left(\epsilon_{t+1}\right) f\left(\omega_{t+1}|a_t, \omega_t\right)$. This implies that we can separate the evolution of the private information from that of the observed states. It also means that private information is independent and identically distributed over time. The assumption of serial independence could represent a problem in my application, and bias my estimates, because of my lack of information on firms' AIV. For example, it is possible that the unobserved component in ν that contributes to preventing a firm for launching a new product in some

[12]These assumptions are similar to those used in Rust (1987) for the single-agent problem.

period will also be present in the future periods before the launch. The estimation of the fixed cost of entry could then pick up this unobserved component.

***Assumption 3*:** *(Independent Private Information)* The state variable ϵ_t is independently distributed across firms and locations: $P_\epsilon(\epsilon_t) = \prod_{f=1}^{F}\prod_{l=1}^{L} g(\epsilon_{flt})$, where $g(\cdot)$ is distributed according to the next assumption.

***Assumption 4*:** *(Logit Distribution)* The state variable ϵ_t is generated from a type 1 extreme value distribution.

3.3.3. Algorithm

The estimation of dynamic models can generally be separated into two main parts: the first part requires obtaining the continuation values for a given parameter value, θ, the second part uses the continuation values obtained in the first part to maximize an objective function with respect to the parameter θ. Notice that the first part, i.e. the search for the continuation values, is the source of most of the computational burden of the estimation. In order to find the parameters that maximize the objective function in the second part, we need to obtain continuation values for many different values of θ. The nested fixed point approach (NFP), a logical extension of the method of Rust (1987) to games, provides an algorithm for the estimation: the search for continuation values is nested within the search for the parameter value that minimizes the distance (with respect to some metric) between predictive and observed choices. The characteristic of this approach is that the first step of the procedure does not use data, and computes the value functions without sampling error. However, because of the severe computational burden, this approach has found limited application[13]. As I mentioned above, several new approaches have recently been offered in the empirical literature which differ mainly in that the continuation values are estimated nonparametrically from the data. When control variables are discrete, such as a new product entry decision, I use the method of Aguirregabiria and Mira (2007). The estimation algorithm is the following:

(0) Compute the set of conditional choice probabilities
$\hat{P}^0 = \left\{ \hat{P}\left(a_f = l|x\right), \forall l, f \right\}$
from the data using a consistent estimator. Guess the parameter estimate $\hat{\theta}^0$.

(1) Compute the value function $V\left(x|\hat{P}^0, \hat{\theta}^0\right)$ given parameter $\hat{\theta}^0$ and conditional choice probabilities $\hat{P}\left(a|x\right)$, according to equation (3.7). The resulting value is used to predict agents' optimal behavior; in particular, since the error is type 1 extreme value, the optimal choice has the well known logit form:

$$\Psi\left(a_j|x, \hat{P}^o, \hat{\theta}^o\right) = \frac{exp\left[V\left(x|a_j, \hat{P}^o, \hat{\theta}^o\right)\right]}{\sum_{a_k \in \mathcal{A}} exp\left[V\left(x|a_k, \hat{P}^o, \hat{\theta}^o\right)\right]} \tag{3.9}$$

(2) Now that the model has a prediction of the behavior given the parameter $\hat{\theta}^o$ and the optimal conditional choice probabilities $\hat{P}\left(a|x\right)$, I can minimize the distance between predicted and actual choices. I apply the maximum likelihood:

$$\hat{\theta}' = \arg\max_{\theta} \prod_{n=1}^{N} \Psi\left(a_l|x_l, \hat{P}^o, \hat{\theta}^o\right) \tag{3.10}$$

where n indexes observations from 1 to N. The estimate $\hat{\theta}'$ is the Hotz-Miller parameter estimate. Although their estimator is much faster than the nested

[13] Some extensions of this approach have been proposed to lessen the computational burden. See Judd (1998).

fixed point approach, it relies on the estimate of $\hat{P}^o$ in Step 0, which is often inaccurate due to the lack of data. Aguirregabiria and Mira avoid this problem by searching for a fixed point in the probability space, proposing a new set of conditional choice probabilities (in the next step), and restarting the algorithm all over again. After a few iterations, their estimator overcomes the problems inherent the two-step approach.

(3) Update the estimates of conditional choice probabilities used by agents:

$$\hat{P}'(a_j|x) = \Psi\left(a_j|x, \hat{P}^o, \hat{\theta}'\right) \ \forall j \in \mathcal{A} \tag{3.11}$$

(4) Given a maximum tolerance value Tol, if $\sum_{a_j \in \mathcal{A}} \left|\hat{P}'(a_j|x) - \hat{P}^o(a_j|x)\right| >$ Tol, set $\hat{P}^o = \hat{P}'$ and $\hat{\theta}^o = \hat{\theta}'$, and go to Step 1. Otherwise, stop. $\hat{P}'$ are the conditional choice probabilities associated with the Markov Perfect Equilibrium, and $\hat{\theta}'$ are the estimated parameters.

3.4. Results

3.4.1. Static Equilibrium

I report the results from the logit model in Tables B.2 and B.3. I regress $ln(S_{jt}) - ln(S_{0t})$ on price, brand and time dummies. To account for endogeneity, I use a two-stage least square regression, where as instruments for brand price I use the price of the brand in the other cities of the same region in each quarter[14]. Notice that the price coefficient is similar to the mean of the price coefficient for the mixed logit model in Nevo (2001). For

[14]For more discussion on the choice of such instruments, see Nevo (2001), page 319.

the products that are introduced in the period for which only DFF data are available, I compute their brand dummies by using the price coefficient estimated with IRI data.

Table B.4 reports the average brand dummies in each of the three segments. This value represents the level of quality that each firm is able to produce by launching a new product into each particular segment, and denotes the benefit from new product launch in terms of demand. It is clear that both Kellogg and General Mills have higher demand opportunities in segment 3 than in segments 1 and 2, whereas Post faces higher demand in segment 2. However, according to the defensive strategy that firms seem to be using, Kellogg and General Mills launch new products mostly in segments 1 and 2, and Post in segment 3. These results imply that the defensive strategy is not due to demand.

Table B.4 shows the average marginal costs in each of the three segments. The estimates show that brands in segment 3 have the highest marginal costs, followed by brands in segment 2. Notice that this result is consistent across firms, and that Kellogg and General Mills seem to benefit from cost advantages in segment 1 and 2, whereas Post has lower marginal costs in segment 3. Given the way shares are distributed across segments (Figures B.5-B.6), this seems to suggest that firms take advantage of economies of scale for producing brands within the same segment. Therefore, marginal costs could be a reason for firms' defensive strategies. For a better understanding of the role of marginal costs, in Table B.6 I report the average AIV in each of the three segments; this value represents the net contribution to the total share that each firm is able to obtain by introducing a new product in a particular location, after accounting for marginal costs. It appears that such firms do not enjoy higher AIVs in segments with higher relative shares; Kellogg and Post reach higher inclusive values in the second segment, and General Mills in the first segment, although the differences across segments are quite low. This implies that the differences in marginal costs across segments shown in the previous Table cannot explain firms' defensive strategies. Indeed, the marginal cost advantage of producing a new brand for the segment where the firm is stronger is compensated for by the lower demand obtained for that segment.

3.4.2. Dynamic Parameters

The parameters recovered in the dynamic model are the firm-specific entry costs θ_{fl} for each location l, i.e. the initial fixed costs that a firm faces to launch a product into a location. Each firm can enter in one of three possible locations: (1) added sugar/kids cereal, (2) whole grain/family cereal, (3) added fruit/nuts/enhanced cereal. For identification purpose, the cost of location 3 is set to zero for each firm, so the parameters that are estimated are to be interpreted as differences from this cost. If the estimate of the parameter θ_{fl} is positive, it means that for firm f the fixed cost in location l is bigger than the fixed cost in location 3, i.e. it costs more for firm f to enter in segment l than to enter in segment 3. On the other hand, if the estimate is negative it means that it costs less for firm f to enter in segment l than to enter in segment 3.

The results of the estimation are reported in Table B.7. The parameters of Kellogg are both positive and significant, which means that, for Kellogg, introducing a new brand in segment 3 is more costly than introducing a brand in segments 1 and 2. The same is true for General Mills, although the estimate for segment 2 is significant only at the 10% confidence level. The parameters for Post show an opposite direction, but the estimates are not significant, due to the small number of entries observable by Post. In conclusion, given the entry choice of firms, the dynamic model associates an asymmetric fixed cost structure in the market for Kellogg and General Mills, who find it significantly less expensive to introduce new products in attribute-space locations where their share is bigger compared to that of the other firms. The same seems to be true for Post, for which however I do not have enough observations on new product entry to confirm this. This asymmetry in fixed costs of entry across segments is the ultimate cause of a firm's defensive strategy. The cost of developing and launching a new brand into a segment where the firm has few brands is much higher compared to the cost of introducing a new brand where the firm is already strong. A firm deciding to introduce a new product will significantly diminish its costs by following a defensive rather than an offensive strategy. This result is in line with the tendency of practitioners to focus on cost advantages, and give up future higher profits to save on initial sunk investments[15].

3.5. Counterfactuals

In order to assess the importance of asymmetry in fixed costs of entry across segments in this market it is interesting to observe how the profit of the firms would change by assuming no asymmetry in fixed costs. I limit the counterfactual analysis to a short period of data, from the second quarter of 1993, right after Post launches its successful Banana Nut Crunch brand in segment 3, to the second quarter of 1995. During this period Kellogg launches two brands in segment 1 and one in segment 3, and General Mills launches four brands in segment 1 and two in segment 2. Keeping the number of entries fixed, I let both Kellogg and General Mills re-choose the location of their new products

[15]I thank Betsy Holden at Kraft for this and several other insightful comments on brand managers' practices in consumer packaged goods' markets.

when the fixed cost is the same in each segment, and re-optimize their price according to a Bertrand Nash equilibrium. To compute the expected counterfactual, I compute the new equilibrium for each possible choice combination and then I associate each combination with its associated probability. Since there are 9 brands and 3 possible locations, there exist $3^9 = 19,683$ possible scenarios.

The results are shown in Table B.8. For the period under analysis, a change in Kellogg's fixed costs causes General Mills and Post to lose, respectively, 0.3% and 0.7% of their profits. This is due to a reoptimization of the location of three new products from Kellogg and a new pricing equilibrium. A change in General Mills' fixed costs implies a corresponding profit loss for Kellogg and Post of 1% and 1.3%, respectively. A higher loss in this case is partly due to the fact that General Mills reoptimizes the positioning of five new products.

3.6. Conclusions

When introducing new products, multi-product firms need to evaluate the relationship between the new brand and the existing brands, and opt for a defensive or an offensive entry strategy. In this paper, I estimated an empirical model of competition to study firms' locationing strategy over the attribute space. I applied the framework to the U.S. ready-to-eat cereal market to answer the question of why firms seem to undertake defensive strategies instead of offensive strategies. After recovering the primitives of the market I showed that the asymmetry in fixed entry costs across segments is the main cause of firms' defensive behavior. Marginal costs are also asymmetric across segments, but their difference from one location to another is neutralized by demand opportunities which also differ from one location to another. Using counterfactual analysis I also showed that the asymmetry in fixed cost of entry is more likely to benefit Kellogg's and Post's profits. Allowing for General Mill's symmetric costs would in fact decrease both profits of Kellogg and Post by 1% and 1.3%, respectively.

References

[Ackerberg2001] Ackerberg, Daniel A. 2001. "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation." *mimeo, UCLA*.

[Ackerberg2003] ———. 2003. "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination." *International Economic Review* 44 (3): 1007.

[Ackerberg et al.2005] Ackerberg, Daniel A., Lanier C. Benkard, Steve Berry, and Ariel Pakes. 2005. In *Econometric Tools for Analyzing Market Outcomes*, edited by forthcoming. Elsevier.

[Aguirregabiria and Mira2002] Aguirregabiria, Victor, and Pedro Mira. 2002. "Swapping the Nested Fixed Point Algorithm: A Class of Estimators For Discrete Markov Decision Models." *Econometrica* 70 (4): 1519–1543 (July).

[Aguirregabiria and Mira2007] ———. 2007. "Sequential Estimation of Dynamic Discrete Games." *mimeo* 75 (1): 1–53 (January).

[Bajari, Benkard, and Levin2007] Bajari, Patrick, Lanier C. Benkard, and Jonathan Levin. 2007. "Estimating dynamic models of imperfect competition." *mimeo* 75 (5): 1331–1370 (September).

[Basso, Clements, and Ross2007] Basso, Leonardo J., Matthew T. Clements, and Thomas W. Ross. 2007. "Moral Hazard and Customer Loyalty Programs." *mimeo.*

[Berry, Levinsohn, and Pakes1995] Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841–890.

[Bhatt1987] Bhatt, Swati. 1987. "Strategic Product Choice in Differentiated Markets." *Journal of Industrial Economics* 36 (2): 207–216 (December).

[Borenstein1992] Borenstein, Severin. 1992. "The Evolution of U.S. Airline Competition." *The Journal of Economic Perspectives* 6 (2): 45–73 (Spring).

[Borenstein1996] ———. 1996. "Repeat-Buyer Programs in Network Industries." In *Networks, infrastructure, and the New Task for Regulation*, 137–162. Ann Arbor, Michigan: The University of Michigan Press.

[Brander and Eaton1984] Brander, James A., and Jonathan Eaton. 1984. "Product Line Rivalry." *American Economic Review* 74 (3): 323–334 (June).

[Bresnahan1987] Bresnahan, Timothy F. 1987. "Competition and Collusion in the American Automobile Market: The 1955 Price War." *Journal of Industrial Economics* 35:457–482.

[Cairns and Galbraith1990] Cairns, Robert D., and John W. Galbraith. 1990. "Artificial Compatibility, Barriers to Entry, and Frequent-Flyer Programs." *The Canadian Journal of Economics* 23 (4): 807–816 (November).

[Caminal and Matutes1990] Caminal, Ramon, and Carmen Matutes. 1990. "Endogenous Switching Costs in a Duopoly Model." *International Journal of Industrial Organization* 8:353–373.

[Dafny2003] Dafny, Leemore S. 2003. "Entry Deterrence in Hospital Procedure Markets: A Simple Model of Learning-By-Doing." *mimeo*.

[Doraszelski and Satterthwaite2003] Doraszelski, Ulrich, and Mark A. Satterthwaite. 2003. "Foundations of Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity." *mimeo*.

[Dowling and Uncles1997] Dowling, Grahame R., and Mark Uncles. 1997. "Do Customer Loyalty Programs Really Work?" *Sloan Management Review* 38 (4): 71–82.

[Ellison and Ellison2000] Ellison, Glenn, and Sarah Fisher Ellison. 2000. "Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbents Prior to Patent Expiration." *mimeo*.

[Erdem, Imai, and Keane2003] Erdem, Tulin, Susumu Imai, and Michael P. Keane. 2003. "Brand and Quantity Choice Dynamics Under Price Uncertainty." *Quantitative Marketing and Economics* 1 (1): 5–64.

[Erdem and Keane1996] Erdem, Tulin, and Michael P. Keane. 1996. "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets." *Marketing Science* 15 (1): 1–20.

[Erdem, Keane, and Strebel2003] Erdem, Tulin, Michael P. Keane, and Joel Strebel. 2003. "Learning about Computers: An Analysis of Information Search and Technology Choice." *Quantitative Marketing and Economics.*

[Ericson and Pakes1995] Ericson, Richard, and Ariel Pakes. 1995. "Markov-perfect Industry Dynamics: A Framework for Empirical Work." *The Review of Economic Studies* 62 (1): 53–82 (January).

[Fox1968] Fox, Harold W. 1968. *The Economics of Trading Stamps.* Washington D.C.: Public Affair Press.

[Gitlin and Ellis2005] Gitlin, Marty, and Topher Ellis. 2005. *Milking that Crazy Cow. A Century of Cereals.* Lulu.

[Gonul and Srinivasan1996] Gonul, Fusun, and Kannan Srinivasan. 1996. "Estimating the Impact of Consumer Expectations of Coupons on Purchase Behavior: A Dynamic Structural Model." *Marketing science* 15 (3): 262–279.

[Goolsbee and Syverson2004] Goolsbee, Austan, and Chad Syverson. 2004. "How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines." *mimeo.*

[Hartmann and Viard2008] Hartmann, Wesley R., and V. Brian Viard. 2008. "Do Frequency Reward Programs Create Switching Costs? A Dynamic Structural Analysis of Demand in a Reward Program." *Quantitative Marketing and Economics* 6 (2): 109–137 (June).

[Hay1976] Hay, Donald A. 1976. "Sequential Entry and Entry-Deterring Strategies in Spatial Competition." *Oxford Economic Papers* 28 (2): 240–257.

[Hendel and Nevo2006] Hendel, Igal, and Aviv Nevo. 2006. "Measuring the Implication of Sales and Consumer Stockpiling Behavior." *Econometrica* 74 (6): 1637–1673 (November).

[Hotz and Miller1993] Hotz, Joseph V., and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Review of Economic Studies* 60 (3): 497–529.

[Hotz et al.1994] Hotz, Joseph V., Robert A. Miller, Seth Saunders, and Jeffrey Smith. 1994. "A Simulation Estimator for Dynamic Models of Discrete Choice." *Review of Economic Studies* 61 (2): 265–289 (April).

[Hull1932] Hull, Clark L. 1932. "The Goal-Gradient Hypothesis and Maze Learning." *Psychological Review* 39 (1): 25–43.

[Judd1985] Judd, Kenneth L. 1985. "Credible Spatial Preemption." *RAND Journal of Economics* 16 (2): 153–166.

[Judd1998] ———. 1998. *Numerical methods in economics.* Cambridge, MA: MIT Press.

[Kim, Shi, and Srinivasan2001] Kim, Byung-Do, Mengze Shi, and Kannan Srinivasan. 2001. "Reward Programs and Tacit Collusion." *Marketing Science* 20 (2): 99–120 (Spring).

[Kivetz, Urminsky, and Zheng2006] Kivetz, Ran, Oleg Urminsky, and Yuhuang Zheng. 2006. "The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention." *Journal of Marketing Research* 43 (1): 39–58 (February).

[Klemperer1987a] Klemperer, Paul. 1987a. "The Competitiveness of Markets with Switching Costs." *RAND Journal of Economics* 18 (1): 138–150.

[Klemperer1987b] ———. 1987b. "Markets with Consumer Switching Costs." *The Quarterly Journal of Economics* 102 (May): 375–394.

[Kopalle and Neslin2003] Kopalle, Praveen K., and Scott A. Neslin. 2003. "The Economic Viability of Frequency Reward Programs in a Strategic Competitive Environment." *Review of Marketing Science* 21, no. 1.

[Lal and Bell2003] Lal, Rajiv, and David R. Bell. 2003. "The Impact of Frequent Shopper Programs in Grocery Retailing." *Quantitative Marketing and Economics* 1 (1): 179–202.

[Lederman2008] Lederman, Mara. 2008. "Do Enhancements to Loyalty Programs Affect Demand? The Impact of International Frequent Flyer Partnerships on Domestic Airline Demand." *RAND Journal of Economics*, vol. forthcoming.

[Levine1987] Levine, Michael E. 1987. "Airline Competition in Deregulated Markets: Theory, Firm Strategy, and Public Policy." *Yale Journal on Regulation* 4:393–494.

[Lewis2004] Lewis, Michael. 2004. "The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention." *Journal of Marketing Research* 41:281–292.

[Manski1993] Manski, Charles. 1993. "Dynamic Choice in a Social Setting: Learning from the Experience of the Others." *Journal of Econometrics* 58 (1-2): 121–136.

[Martinez-Girault and Neven1988] Martinez-Girault, Xavier, and Damien J. Neven. 1988. "Can Price Competition Dominate Market Segmentation?" *Journal of Industrial Economics* 36 (4): 431–442 (June).

[Maskin and Tirole2001] Maskin, Eric, and Jean Tirole. 2001. "Markov Perfect Equilibrium." *Journal of Economic Theory* 100 (2): 191–219 (October).

[Mazzeo2002] Mazzeo, Michael J. 2002. "Product Choice Oligopoly and Oligopoly Market Structure." *RAND Journal of Economics* 33 (2): 1–22.

[McFadden1981] McFadden, Daniel. 1981. "Econometric Models of Probabilistic Choice." In *Structural analysis of discrete data*, edited by Charles Manski and Daniel McFadden. MIT Press.

[McFadden1974] McFadden, Daniel L. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers of Econometrics*, edited by Paul Zarembka. New York Academic Press.

[McFadden1978] ———. 1978. "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, edited by A. Karlgvist, Lundqvist L., and Weibull J.W. Amsterdam, North Holland.

[Melnikov2000] Melnikov, Oleg. 2000. "Demand for Differentiated Durable Products: The Case of the U.S. Computer Printer Market." *mimeo, Yale University.*

[Nako1992] Nako, Stephen M. 1992. "Frequent Flyer Programs and Business Travellers: An Empirical Investigation." *Logistics and Transportation Review* 28 (4): 395–414 (December).

[Nevo2001] Nevo, Aviv. 2001. "Measuring market power in ready-to-eat cereal industry." *Econometrica* 69 (2): 307–342 (March).

[Pakes2000] Pakes, Ariel. 2000. "A Framework for Applied Dynamic Analysis in I.O." *NBER Paper*, no. 8024.

[Pakes and McGuire1994] Pakes, Ariel, and Paul McGuire. 1994. "Computing Markov-Perfect Nash equilibria: Numerical Implications of a Dynamic Differentiated Product Model." *The RAND Journal of Economics* 25 (4): 555–589 (Winter).

[Pakes, Ostrowsky, and Berry2005] Pakes, Ariel, Michael Ostrowsky, and Steven T. Berry. 2005. "Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit Examples)." *mimeo.*

[Pesendorfer and Schmidt-Dengler2003] Pesendorfer, Martin, and Philipp Schmidt-Dengler. 2003. "Identification and Estimation of Dynamic Games." *mimeo.*

[Prescott and Visscher1977] Prescott, Edward C., and Michael Visscher. 1977. "Sequential Locations among Firms with Foresights." *Bell Journal of Economics* 8 (2): 378–393.

[Proussaloglou and Koppelman1999] Proussaloglou, Kimon, and Frank S. Koppelman. 1999. "The Choice of Air Carrier, Flight, and Fare Class." *Journal of Air Transport Management* 5:193–201.

[Rust1987] Rust, John P. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55 (5): 999–1033.

[Rust1996] ———. 1996. "Numerical Dynamic Programming in Economics." In *Handbook of computational economics*, edited by Hans M. Amman and Daviv A. Kendrick, Volume 1. North-Holland.

[Scherer1982] Scherer, Frederic M. 1982. In *The Breakfast Cereal Industry*, edited by W. Adams. Macmillan, New York.

[Schmalensee1978] Schmalensee, Richard. 1978. "Entry Deterrence in the Ready-To-Eat Breakfast Cereal Industry." *Bell Journal of Economics* 9:305–327.

[Seim2005] Seim, Katja. 2005. "An Empirical Model of Firm Entry with Endogenous Product–Type Choices." *RAND Journal of Economics*, vol. forthcoming.

[Shaked and Sutton1990] Shaked, Avner, and John Sutton. 1990. "Multiproduct Firms and Market Structure." *RAND Journal of Economics* 21 (1): 45–62 (Spring).

[Shugan2005] Shugan, Steven M. 2005. "Brand Loyalty Programs: Are They Shams?" *Marketing Science* 24 (2): 185–193.

[Song and Chintagunta2003] Song, Inseong, and Pradeep K. Chintagunta. 2003. "A Micromodel of New Product Adoption with Heterogeneous and Forward-Looking Consumers: Application to the Digital Camera Category." *Quantitative marketing and economics* 1:371–407.

[Stephenson and Fox1992] Stephenson, Frederick J., and Richard J. Fox. 1992. "Corporate Strategies for Frequent-Flier Programs." *Transportation Journal*, Fall.

[Tan et al.2002] Tan, Chyi Woan, Greg Tower, Phil Hancock, and Ross Taplin. 2002. "Empires of the Sky: Determinants of Global Airlines Accounting-Policy Choices." *International Journal of Accounting* 37:277–299.

[Zhu, Singh, and Manuszak2005] Zhu, Ting, Vishal Singh, and Mark Manuszak. 2005. "Market Structure and Competition in the Retail Discont Industry." *mimeo.*

APPENDIX A

Tables and Figures for Chapter 1

	Percentage	Total
Total number of electronic cards	100.0%	Some Millions
Purchased *regular fuel* at least once	98.8%	
Purchased *premium fuel* at least once	7.6%	
Purchased *mineral lubricant* at least once	10.6%	
Purchased *synthetic lubricant* at least once	2.6%	
Redeemed at least one reward	26.3%	
Pooled points with other cards	17.3%	
Swapped points with partner programs	8.5%	

Table A.1. Statistics for cards' main activities on the entire database. The table shows that in the database most cards are not used to redeem any reward. Only a small amount of cards redeem at least one reward. Due to confidentiality concerns, I only report relative values.

	% Cust.	Mean	Median	Std	Min	Max
Purchase of *regular fuel*						
Quantity (gallons)	98.8%	9.16	7.93	5.47	0.26	29
Duration (days)		15.78	9.91	22.29	0.03	794
Purchase of *premium fuel*						
Quantity (gallons)	7.6%	11.46	10.57	6.51	0.26	29
Duration (days)		338.97	257	291.08	0.20	822
Purchase of *mineral lubricant*						
Quantity (fl. oz)	10.6%	78.81	59.74	58.92	2.71	13,661
Duration (days)		450.30	404	273.58	0.25	822
Purchase of *synthetic lubricant*						
Quantity (fl. oz)	2.6%	94.38	67.63	73.90	33.81	324
Duration (days)		459.44	408.50	281.93	0.33	822
# different distributors visited		3.13	2.00	3.86	1.00	209

Table A.2. Statistics of fuel and lubricant purchase.

	Mean	Median	Std	Min	Max
Pts used per redemption date (#)	2164.48	1500	2308.29	200	61700
Pts used per redemption date (used/available)	0.68	0.74	0.28	0.01	1.00
Rewards redeemed on same date (#)	1.24	1	0.71	1	59
Rewards redeemed per customer (#)	2.24	1	2.26	1	85

Table A.3. Points used and prizes redeemed. Note: to compute these statistics I selected from the database cards that redeemed at least one prize, and never pooled points across cards or swapped points across partner programs.

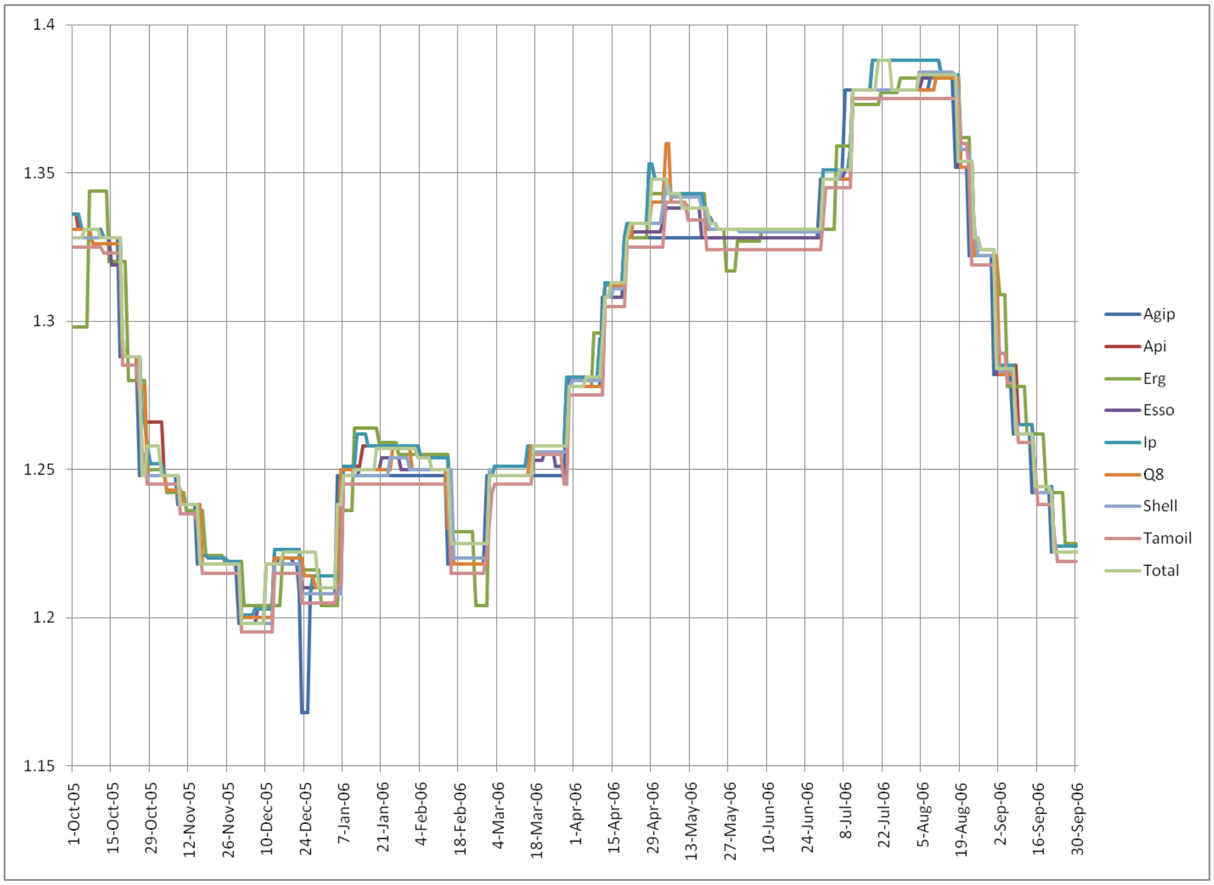

Figure A.1. National price levels (in euros) for one liter (0.264 gallons) of unleaded fuel with self-service pumping in the Italian market, during the period from October 2005 to September 2006.

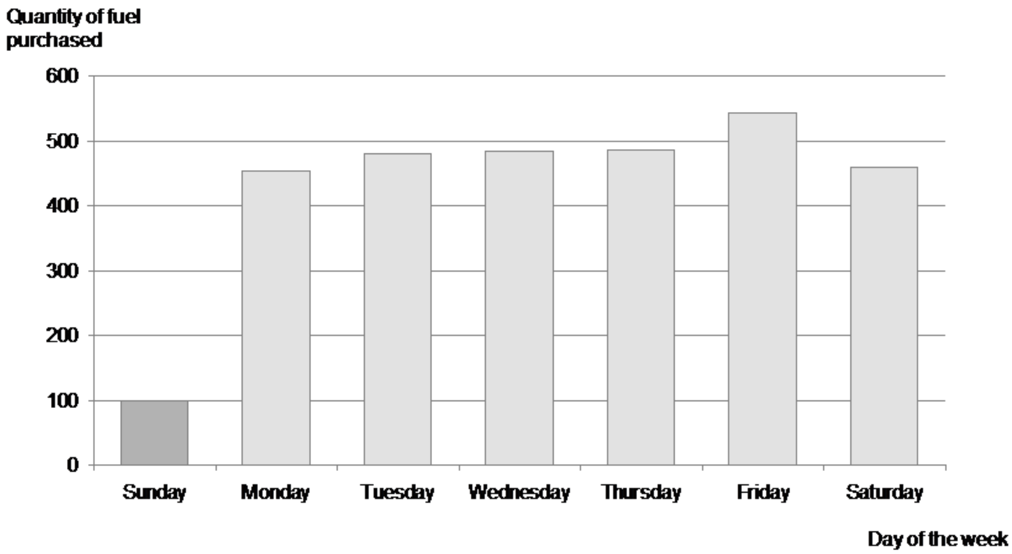

Figure A.2. Variation in the quantity of fuel purchased across weekdays for a random sample of 500 individuals. The sales quantity of Sundays is set as benchmark level (100). The graph clearly shows that the purchase of fuel is significantly lower on Sundays compared to every other day of the week.

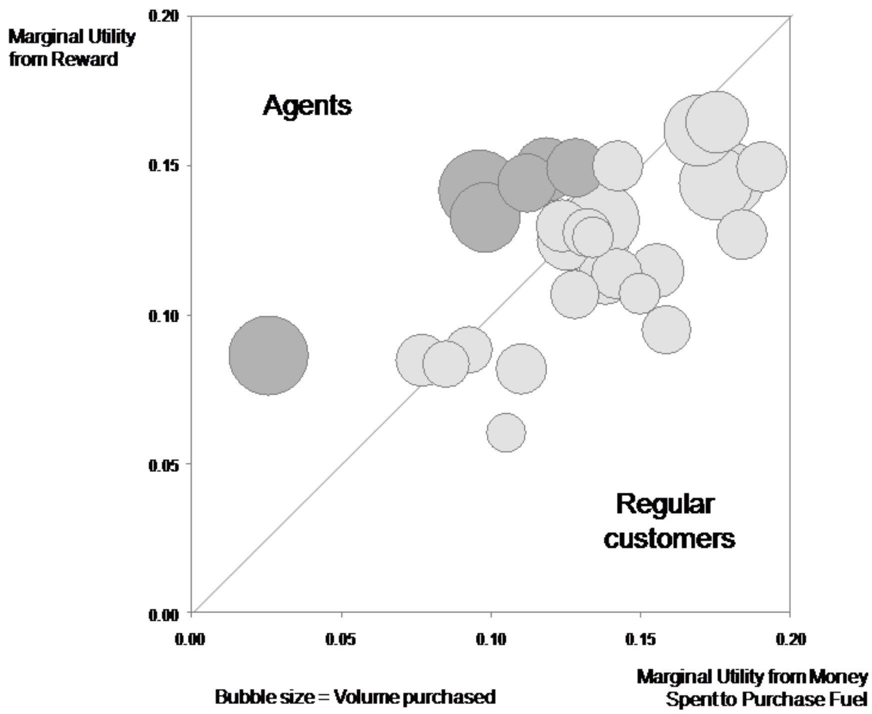

Figure A.3. Sensitivity to price and to reward for 30 consumers randomly drawn from the sample used in the estimation. Note: the horizontal axis measures the individual marginal utility from money spent for fuel at the gas station; the vertical axis measures the individual marginal utility from rewards; the size of the bubble represents consumer's usage, i.e. the simulated volume of sales generated in one year when the reward program is on. The diagonal line represents the indifference line where money and rewards provide the same level of marginal utility. This line distinguishes the group of agents from the group of regular customers.

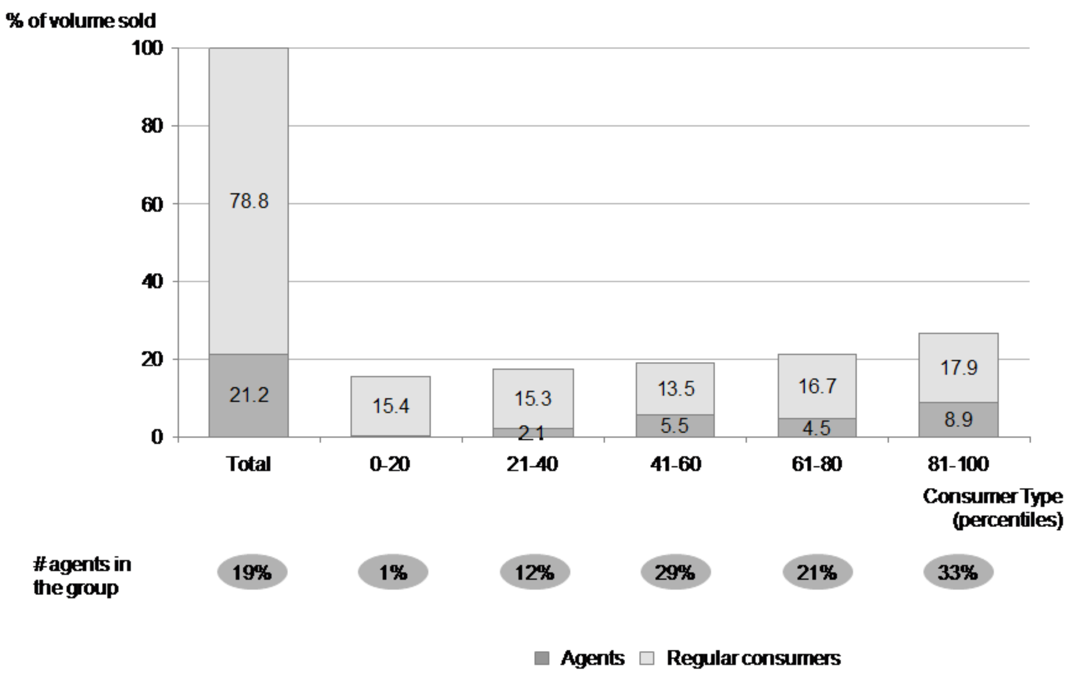

Figure A.4. Percentage of agents in the population by usage type. Note: in the first column on the left I report the sales of all customers. In the next columns consumers are divided into five groups based on the percentiles of the distribution of total liters of gasoline purchased in one year (0 to 20th, 21st to 40th, 41st to 60th, 61st to 80th, 81st to 100th). The liters of gasoline purchased are simulated sales generated in one year when the reward program is active.

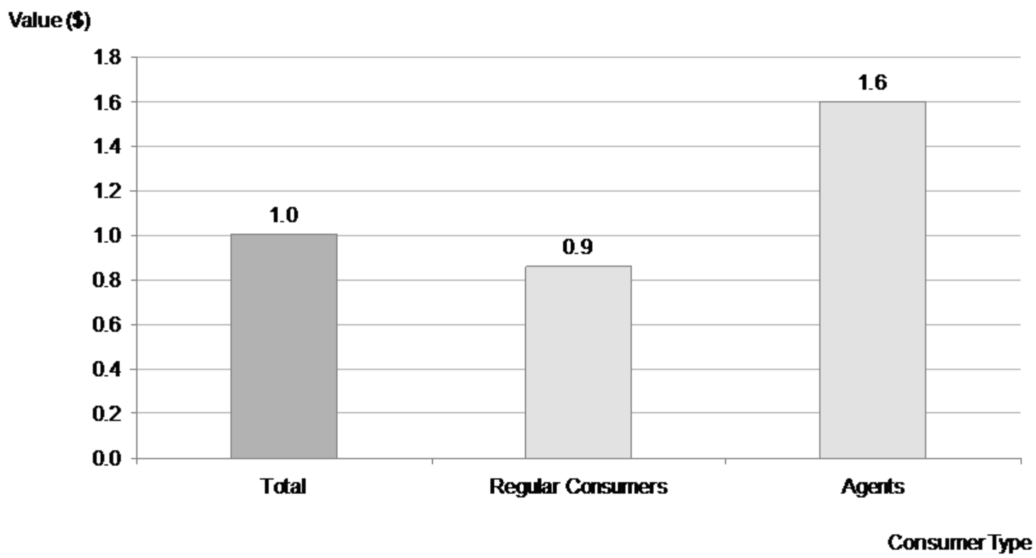

Figure A.5. Value of one dollar worth of reward, expressed in terms of dollars spent on fuel. Note: in the first column on the left I report the value of one dollar spent on fuel for all customers. In the next columns I distinguish by type of consumers: regular consumers and agents.

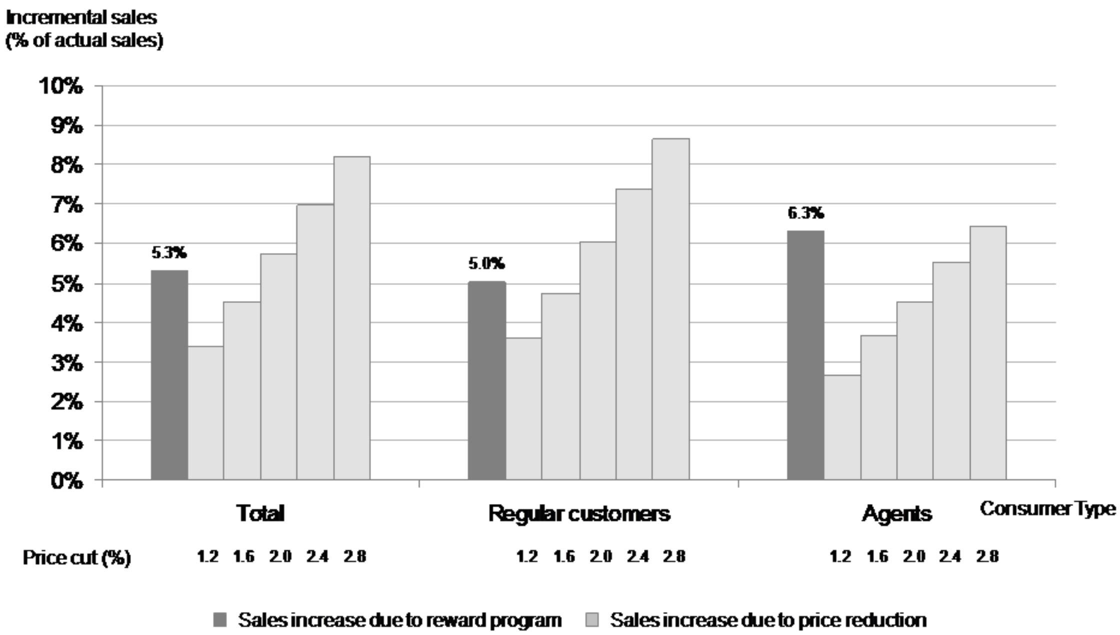

Figure A.6. Impact of reward program and price reduction policies on sales volumes. Overall, the reward program is responsible for 5.29% of incremental sales volumes. The same level of sales can be obtained by decreasing prices by 1.85%. This result, however, differs depending on consumer type. Note: the dark gray columns represent the level of sales generated by the reward program; the light gray columns represent the sales generated by pricing policies with different price reductions; the sales generated when there is no reward program or price reduction are normalized to 0. The first columns on the left represent sales of the whole population. In the next columns I distinguish by type of consumers: regular consumers and agents.

	Random-Coefficient Model								
	Mean	Variance-Covariance Matrix							
Fixed Effect Retailer	-2.089	0.895							
	(0.034)	(0.025)							
Price Coefficient	-5.316	0.3940	3.249						
	(1.515)	(0.032)	(1.498)						
Points Accumulated	0.576	0.047	-0.030	0.001					
	(0.065)	(0.010)	(0.025)	(0.081)					
Days Since Last Purchase	0.266	0	0	0	0.523				
	(0.021)				(0.024)				
Sunday Indicator	-2.195	0	0	0	0	0.926			
	(0.069)					(0.060)			
Fall Indicator	0.290	0	0	0	0	0	0.057		
	(0.027)						(0.039)		
Winter Indicator	0.238	0	0	0	0	0	0	0.013	
	(0.023)							(0.032)	
Spring Indicator	0.248	0	0	0	0	0	0	0	0.015
	(0.023)								(0.033)
N. customers	500								
N. obs	162,937								
LogLik	-58,559.51								

Table A.4. Estimates of the static demand model. Parameters are ML estimates. Standard errors are in parentheses. The period is days. The outside good's intercept is normalized to zero. The zero elements in the Variance-Covariance matrix have not been estimated.

	Random-Coefficient Model									
	Mean	Variance-Covariance Matrix								
Price Coefficient	-0.126	0.010								
	(0.005)	(0.003)								
Reward Coefficient	-2.140	-0.047	0.4086							
	(0.037)	(0.019)	(0.189)							
Days Since Last Purchase	0.729	0	0	0.0911						
	(0.016)			(0.036)						
Fixed Effect Retailer	-5.832	0.023	-0.043	-0.254	0.953					
	(0.045)	(0.002)	(0.008)	(0.087)	(0.024)					
Fixed Effect $q = 20$	1.257	0	0	0	0	0.918				
	(0.053)					(0.363)				
Fixed Effect $q = 30$	1.235	0	0	0	0	0	0.002			
	(0.005)						(0.001)			
Fixed Effect $q = 40$	1.170	0	0	0	0	0	0	0.046		
	(0.015)							(0.065)		
Fixed Effect $q = 50$	0.839	0	0	0	0	0	0	0	0.001	
	(0.002)								(0.043)	
Fixed Effect $q = 60$	0.836	0	0	0	0	0	0	0	0	1.282
	(0.053)									(0.785)
Discount Factor	0.9995									
N. customers	500									
N. obs	139,409									
LogLik	-85,741.70									

Table A.5. Estimates of the dynamic demand model. Parameters are ML estimates. Standard errors are in parentheses. The period of the model is days. The outside good's fixed effect and the coefficient for fixed effect for $q = 10$ are normalized to zero. The reward coefficients are distributed $log\left(\gamma_h\right) \sim N\left(\gamma, \Sigma\right)$. The discount factor is not estimated.

APPENDIX B

Tables and Figures for Chapter 3

Parameter	Coefficient	Std. Error
Intercept Segment 1	-1.483	0.855
Intercept Segment 2	-2.251	0.980**
Own Segment Share	17.302	5.557**
LogL	-25.560	
N	30	

Table B.1. Reduced-Form Analysis: Conditional Logit Model of Segment Choice. (**.01 significance)

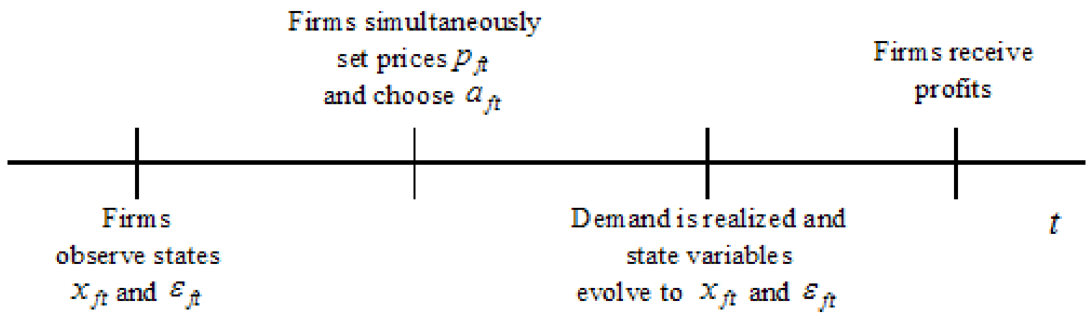

Figure B.1. Timing of the game in period t.

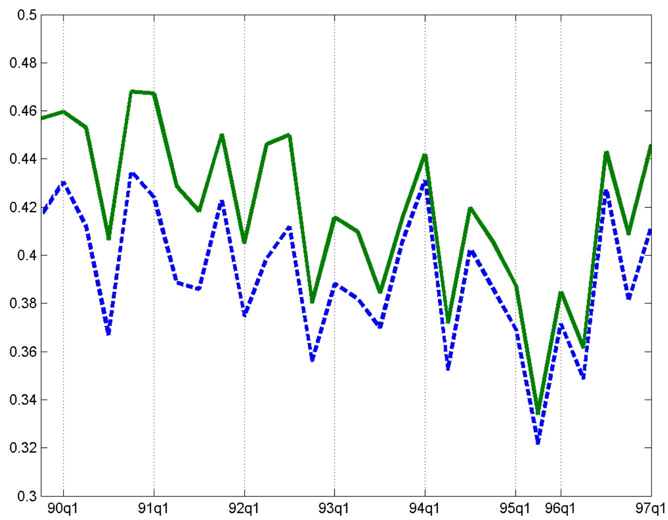

Figure B.2. Shares per quarter generated by all products (solid line) and by incumbent products only (dashed line) of Kellogg's during 1988-1997. Incumbent products are products that have been launched in the market before 1989. Note: the second and third quarters of 1995 are missing.

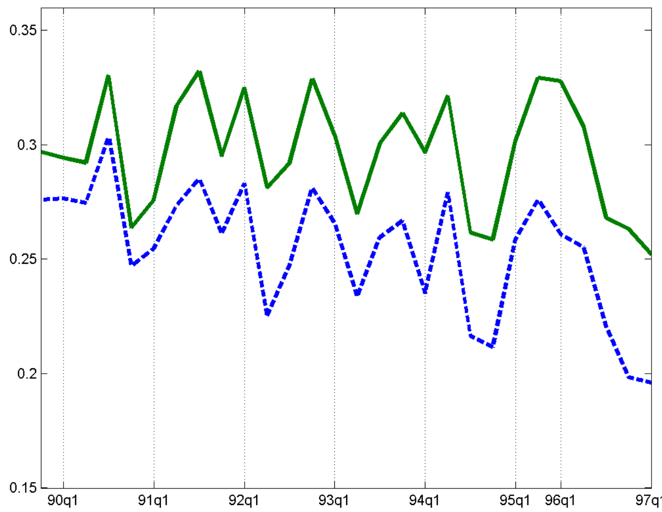

Figure B.3. Shares per quarter generated by all products (solid line) and by incumbent products only (dashed line) of General Mills during 1988-1997. Incumbent products are products that have been launched in the market before 1989. Note: the second and third quarters of 1995 are missing.

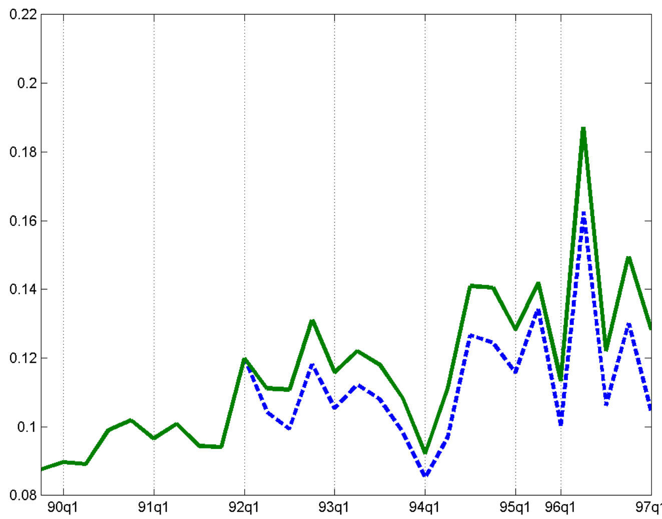

Figure B.4. Shares per quarter generated by all products (solid line) and by incumbent products only (dashed line) of Post during 1988-1997. Incumbent products are products that have been launched in the market before 1989. Note: the second and third quarters of 1995 are missing.

Figure B.5. Shares per firm across segments in the last quarter of 1989.

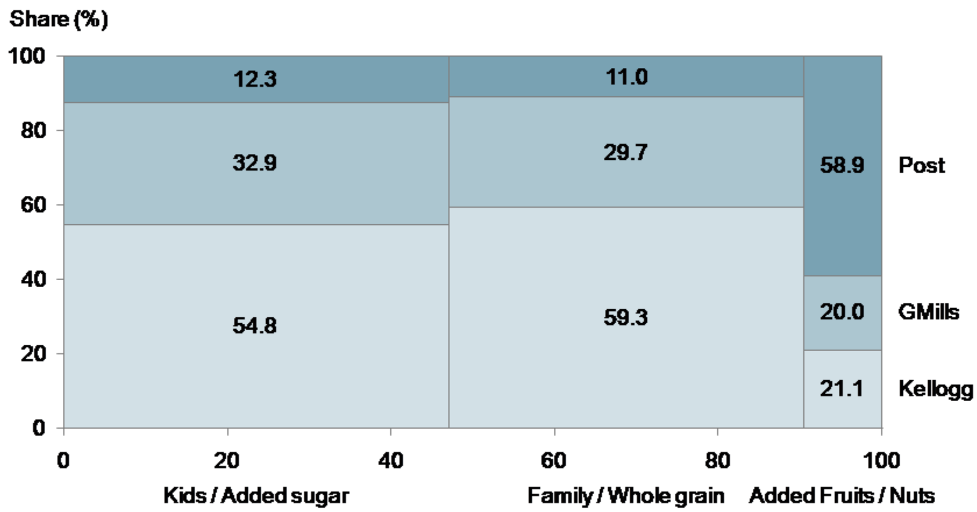

Figure B.6. Shares per firm across segments in the first quarter of 1997.

	Kellogg			General Mills			Post		
Period	Sugar	Family	Fruits/Nuts	Sugar	Family	Fruits/Nuts	Sugar	Family	Fruits/Nuts
1988q1		nutrific							
1988q2							smurfb. crunch		
1988q3							croonchy stars		
1988q4		common sense		fruity y. mummy					
1989q1				a.cinn. cheerios					
1989q2		s.w. graham						honey b.of oats	
1989q3	nut honey cr. os	all bran ex.fiber			benefit				
1989q4		hearthwise							
1990q1									
1990q2			kenmei						
1990q3		bigmixx							
1990q4									great grains
1991q1	cinn. mini buns					basic4			
1991q2					triples				
1991q3	doub.dip crunch								
1991q4					wheaties h.gold				
1992q1		frosted bran							
1992q2					cheerios m.grain				
1992q3									
1992q4									
1993q1									ban. nut crunch
1993q2					ripple crisp				
1993q3									
1993q4				sprink. spangles					
1994q1	a.cinn. krispies			bunuelitos					
1994q2			healthy choice	hidden treasures					blueb. morning
1994q3				reeses					
1994q4					sun crunchers				
1995q1	temptations				frosted cheerios				
1995q2									
1995q3									
1995q4									
1996q1							waffle crisp		
1996q2									
1996q3	h.crunch corn			honeyn. clusters		bc streusel			
1996q4				french toast					
1997q1	cocoa frosted			jurassic crunch					cranb. almond

Figure B.7. New product introductions in the U.S. RTE cereal market between 1988 and 1997.

	Segment	Coefficient	Std. Error	Avg. Marg. Cost
Price	-	-26.56	0.58	-
K Frosted Flakes	1	0.37	0.09	0.11
K Froot Loops	1	0.24	0.11	0.14
K Cinn. Mini Buns*	1	-1.20	0.11	0.14
K Corn Pops	1	0.25	0.12	0.16
K N.Hon. Crunch Os*	1	-1.98	0.11	0.14
K Double Dip Crunch*	1	-5.05	0.36	0.16
K App.Cinn. Krispies*	1	-4.32	0.38	0.14
K Temptations*	1	-3.79	0.41	0.17
K H.Crunch Corn*	1	-2.29	0.59	0.10
K Cocoa Fr.Flakes*	1	-2.50	0.99	0.09
K Frosted Bran*	1	-4.22	0.30	0.12
GM Hon.N. Cheerios	1	0.47	0.10	0.13
GM Lucky Charms	1	0.39	0.12	0.15
GM Trix	1	0.85	0.14	0.18
GM Cinn.Toast Crunch	1	0.02	0.12	0.16
GM Kix	1	0.28	0.13	0.17
GM Ice Cream Cones*	1	-3.54	0.13	0.16
GM Fruity Y. Mummy*	1	-0.63	0.15	0.18
GM App.Cinn. Cheerios*	1	-0.06	0.11	0.15
GM Triples*	1	-1.50	0.09	0.12
GM Wheaties H.Gold*	1	-1.63	0.09	0.11
GM Sprinkle Spangles*	1	-4.49	0.39	0.14
GM Bunuelitos*	1	-7.53	0.61	0.20
GM Hidden Treasures*	1	-3.53	0.53	0.16
GM Reeses*	1	-3.26	0.37	0.16
GM Hon.N. Clusters*	1	-3.19	0.59	0.12
GM French Toasts*	1	-2.23	0.71	0.12
GM Jurassic Crunch*	1	-3.28	0.99	0.13
GM Frosted Cheerios*	1	-1.89	0.41	0.12
P Smurfb. Crunch*	1	-0.66	0.14	0.18
P Croonchy Stars*	1	-1.39	0.13	0.16
P Waffle Crisp*	1	-3.89	0.43	0.14
Q Life	1	-0.85	0.10	0.12
Q CapN Crunch	1	-0.29	0.09	0.11

Table B.2. Results of static equilibrium: demand parameters and marginal costs. (* = new product)

	Location	Coefficient	Std. Error	Avg. Marg. Cost
K Corn Flakes	2	-0.97	0.06	0.06
K Raisin Bran	2	0.50	0.11	0.14
K Rice Krispies	2	-0.28	0.08	0.09
K Mini Wheats	2	2.45	0.17	0.24
K Crispix	2	-0.10	0.12	0.15
K Nutrific*	2	-0.98	0.13	0.16
K Common Sense*	2	-1.03	0.11	0.14
K S.W. Graham*	2	-1.43	0.12	0.14
K All Bran Ex.Fib*	2	-2.62	0.09	0.11
K Hearthwise*	2	4.17	0.25	0.37
K Big Mix*	2	-2.78	0.08	0.09
GM Cheerios	2	1.09	0.11	0.14
GM Wheaties	2	-1.00	0.09	0.11
GM Raisin Nut Bran	2	2.55	0.19	0.27
GM Benefit*	2	1.92	0.21	0.29
GM Triples*	2	-4.84	0.30	0.13
GM Cheerios M.grain*	2	-1.93	0.28	0.16
GM Ripple Crisp*	2	-4.28	0.38	0.13
GM Sun Crunchers*	2	-3.72	0.44	0.14
P Raisin Bran	2	-0.23	0.11	0.14
P Hon. B.of Oats*	2	-0.59	0.11	0.14
P 100N Shredded Wheat	2	0.86	0.14	0.20
K Special K	2	0.42	0.12	0.16
K Kenmei*	3	1.30	0.18	0.25
K Healthy Choice*	3	-1.45	0.35	0.17
GM Total	3	1.32	0.13	0.17
GM Basic 4*	3	3.65	0.22	0.34
GM B.C. Streusel*	3	-2.16	0.59	0.13
P Grape Nuts	3	1.32	0.14	0.19
P Great Grains*	3	2.32	0.20	0.30
P Banana Nut Crunch*	3	-2.18	0.30	0.14
P Blueb. Morning*	3	-2.01	0.35	0.16
P Cranb. Almond*	3	-1.90	0.99	0.14

Table B.3. (Continue) Results of static equilibrium. (* = new product)

	Segment 1	Segment 2	Segment 3
Kellogg	-3.17	-0.78	-0.07
	(0.14)	(0.05)	(0.19)
General Mills	-2.83	-2.57	0.74
	(0.11)	(0.14)	(0.29)
Post	-1.98	-0.59	-0.94
	(0.13)	(0.11)	(0.23)

Table B.4. Average brand dummies.

	Segment 1	Segment 2	Segment 3
Kellogg	0.133	0.154	0.209
General Mills	0.148	0.171	0.214
Post	0.157	0.165	0.187

Table B.5. Average marginal costs.

	Segment 1	Segment 2	Segment 3
Kellogg	-6.46	-5.95	-6.01
General Mills	-5.56	-7.40	-5.89
Post	-7.21	-4.34	-5.82

Table B.6. Average adjusted inclusive values for each firm and segment.

	Segment 1	Segment 2	Segment 3
Kellogg	-1.26**	-1.18**	0
	(.58)	(.58)	-
General Mills	-1.60**	-1.26*	0
	(.73)	(.76)	-
Post	0.29	1.41	0
	(.79)	(1.07)	-

Table B.7. Differences between fixed costs in the segment and fixed costs in segment 3. (**.05 significance, *.10 significance)

Change fixed costs Kellogg		Change fixed costs GMills	
Firm	$\Delta\pi$	Firm	$\Delta\pi$
G.Mills	-0.3%	Kellogg	-1.0%
Post	-0.7%	Post	-1.3%

Table B.8. Change in profits with no asymmetry in fixed costs of entry.