When Demand Projections Are Too Optimistic:

A Structural Model Of Product Line And Pricing Decisions

Andrés Musalem*

February 2015

*Andrés Musalem is an assistant professor at the University of Chile. I thank Carl Mela, Michaela Draganska, Naufel Vilcassim, Kanishka Misra, Bruce Hardie, Wes Hartmann, Che-Lin Su, Iaria Alessandro, Chris Nosko, Chris Conlon, Peter Arcidiacono, Bryan Bollinger and Carlos Noton for very useful discussions about this paper. I am also extremely grateful to Wochoel Shin for his participation and contributions to earlier versions of this paper and to Greg Montalvo and Mauricio Giadach for outstanding programming and R.A. support. All remaining errors are my own. Partial support from FONDECYT (Grant 1130412) is acknowledged. Please address all correspondence on this manuscript to Andrés Musalem, Republica 701, Santiago, Chile, amusalem@dii.uchile.cl.
When Demand Projections Are Too Optimistic:

A Structural Model Of Product Line And Pricing Decisions

Abstract

A methodology is proposed to estimate structural models of product line competition. This methodology enables researchers to estimate demand systems accounting for the endogeneity of the mix of products available in each market, an issue which is typically ignored in the empirical literature. It is observed that not accounting for this endogeneity leads to overoptimistic estimates of demand due to a sample selection bias. These biased estimates of demand can generate misleading managerial recommendations.

Under the proposed model consumer demand is characterized by utility maximization with unobserved heterogeneity in preferences. Price decisions are assumed to be the outcome of a Bertrand-Nash game while product line decisions are modeled using a Bayesian-Nash equilibrium concept. The estimation approach relies on parallelization decomposing some of the most computationally intensive steps into a series of independent and much smaller problems.

The methodology is illustrated using simulated and real data. The results show that ignoring this form of endogeneity leads a researcher to overestimate the demand, prices and profits for products that have not yet been introduced in the market, while for existing products market shares and profits are substantially underestimated when performing policy simulations involving the addition of products to current assortments.

Key Words: Structural models, product line decisions, assortment, pricing, competition, demand estimation, sample selection, industrial organization.

1 Introduction

Determining the best mix of product varieties for every market is one of the most difficult challenges for firms. This problem is relevant for a large number of diverse industries. For example, airline carriers must decide how many options will be made available to consumers to fly between each pair of cities during each day. Manufacturers of portable media players (e.g., Apple Inc.) introduce several models with different capabilities (e.g., the Shuffle, Nano, Mini, Classic and Touch versions of the iPod) targeting customers with different needs. Fast moving consumer goods (FMCG) manufacturers such as The Coca-Cola company make available a large number of products with different flavors, sizes and packaging options to their customers. Automobile manufacturers offer several different models and produce many different versions (trims) of each of these models. Telecommunications companies give consumers multiple choices in terms of cellular phone plans with different numbers of included minutes of each type (e.g., peak and weekend minutes), overage charges and internet services.

There at least three fundamental motivations for firms to offer multiple varieties which are related to: i) customer segmentation, ii) price discrimination and iii) intensity of competition. First, offering multiple varieties enables firms to reach segments of consumers with very different preferences. Some of these segments might choose not to buy any products if only a few varieties were offered. This is the motivation for companies such as Unilever to introduce, for example, sachets of shampoo, detergent and hair oil in India. With these formats, the firm is able to reach a substantial fraction of the market that corresponds to lower-income consumers who are not able to purchase larger quantities of these products on a regular basis (Business Week 2008).

Second, given that consumer preferences for certain product attributes (e.g., size) can sometimes be correlated with price sensitivity, offering multiple products can be potentially used as a price discrimination device. A classic example is the use of yield management techniques in the airline industry that take advantage of the correlation between a customer’s willingness to buy a ticket in advance and her price sensitivity. Third, having multiple products in a market can also have an impact on the intensity of competition. For instance, brand proliferation may serve as a tool to deter further entry in a product category, as it was argued in the case of the ready-to-eat breakfast cereal industry (Schmalensee (1978); Nevo (2003)). In distribution channels, introducing private label products can be used to
intensify competition within a product category leading to more attractive options for a retailer’s customer base (Raju et al. (1995)). Considering these three central motivations (segmentation, price discrimination and intensity of competition), it is evident that the analysis of product line (assortment) decisions is of great importance to managers and to regulators interested in performing consumer welfare analysis, given that these decisions affect the product varieties available to consumers and the prices they pay.

In this paper a methodology is proposed to estimate structural models of product line competition. This methodology enables researchers to estimate demand systems accounting for the endogeneity of the mix of products available in each market, an issue which is typically ignored in the empirical literature (e.g., Berry et al. (1995); Villas-Boas and Winer (1999)). In particular, ignoring this form of endogeneity leads to overoptimistic estimates of total demand. This can be explained by the fact that, ceteris paribus, a product is more likely to be introduced if it generates sufficient demand from consumers in the market. Therefore, if there are unobservable factors contributing to consumer utility, the utility contribution of these unobserved characteristics for introduced products is likely to be higher than the one corresponding to the non-introduced products. Consequently, when ignoring the endogeneity of the observed assortments in each market, only the introduced products will be considered and this particular form of sample selection (Heckman (1979)) may cause a researcher to overestimate the true demand for the product category. More importantly, biased estimates of demand can lead to misleading managerial recommendations and inaccurate inferences about consumer welfare. In addition, the use of a structural model enables a researcher to consider policy simulations aimed at studying the consequences of changes in a number of structural factors, such as consumer preferences, cost structure, ownership and capacity.

The structural model in this paper jointly considers the interaction between consumer preferences, pricing and product line decisions. Specifically, consumer demand is characterized by a utility maximization process with unobserved heterogeneity in consumer preferences. Prices are assumed to be the outcome of a Bertrand-Nash game. Product line decisions are modeled using a Bayesian-Nash equilibrium concept where firms form beliefs about the profits of their competitors and anticipate the prices and demand they would observe for any given set of products that could be introduced in the market. Other setups are also possible to be modeled, such as games of complete information. Nevertheless, a Bayesian-Nash framework is chosen given that, in the limit, it approximates a complete
information case as the uncertainty on the firms’ types approaches zero.

In relation to the extant literature, Draganska et al. (2009) address a similar problem. They propose a framework to jointly estimate firms’ product assortment decisions, pricing and demand. Their approach, however, does not consider product characteristics unobserved to the econometrician. These unobservables play a crucial role in the estimation of demand from aggregate data as they prevent the model from becoming deterministic. Moreover, the inclusion of these demand unobservables introduces an econometric estimation challenge as pricing and assortment decisions could be made by firms with knowledge of these unobserved demand conditions. In this paper, this endogeneity issue will be explicitly addressed by formulating a demand model that includes these unobserved demand characteristics and by estimating the relationship between these characteristics and both pricing and product line decisions.

Another related paper is Misra (2008), which considers a setup where demand is characterized by a logit model with homogenous preferences and assortment decisions are represented by a complete information game. This setup leads to very simple properties of the optimal price and assortment policies. In particular, equilibrium retail margins are constant across products and the set of introduced products can be determined by ranking the products of each firm using a product score that can be computed using the utility coefficients and the marginal cost of the product. This greatly simplifies the estimation procedure given that it is not necessary to compute the profitability of each possible assortment. Nevertheless, a limitation of this approach is that it ignores interactions among different products in the assortment. In order to illustrate the implications of this result, consider the following example. Suppose a store carries three products in the carbonated soft drinks product category: Coke, Diet Coke and Pepsi. The retailer relies on a homogenous logit model to characterize consumer demand and uses this model to determine which additional product should be introduced. Furthermore, assume that Diet Pepsi and Sprite, which are currently not in the assortment, have the same score and the same cost for the retailer. As shown in Misra (2008) and based on the assumptions of the homogenous logit demand system, the introduction of each of these two products (Diet Pepsi and Sprite) would be equally profitable for the retailer. In contrast, if consumers have heterogeneous preferences such that the first segment prefers cola products, the second diet products and the third lemon-lime products, then even though Diet Pepsi and Sprite have the same score, introducing the latter is likely
to lead to higher incremental profits. This is explained by the fact that adding Sprite to the product assortment would enable the retailer to appeal to the third segment, which is not able to find any of its favorite products in the current assortment. Consequently, although the results in Misra (2008) provide great simplifications for the study of competitive product line decisions, it is still necessary to develop tools that address more flexible and richer models of consumer demand. In particular, in this paper a random-coefficients logit model is used which at the aggregate level avoids the limitations related to the independence of irrelevant alternatives (IIA) assumption and, therefore, explicitly accounts for the interactions among all of the alternatives under consideration. Moreover, in the empirical literature in marketing and economics, random-coefficient choice models have become one of the most prevalent random-utility frameworks to characterize consumer demand.

More recently, Eizenberg (2014) models product introduction decisions in the context of the U.S. home PC market. In that paper firms only observe demand shocks after products are introduced. Therefore, these shocks are not only unknown to the econometrician but also to firms, which by construction eliminates selection issues from the demand estimation. Wollmann (2014) studies product introduction decisions in the commercial vehicles industry and also assumes that demand shocks are observed after product introduction. That assumption is perhaps most appropriate when studying the introduction of new products as opposed to when modeling the choices of product varieties to be offered by a firm across multiple markets, as is the focus of this paper. In addition, Iaria (2014) models the entry of big-box retailers. In order to account for the endogeneity of these decisions, exclusion restrictions are employed that rely on having a set of instruments correlated with market entry but that do not directly affect demand. In this paper instead a full information approach is used instead where the probability of making a set of products available in the market is explicitly modeled as a function of the primitives that determine consumer utility and firm’s costs and profits.

Interestingly, the study of product line decisions has also been extensively considered in the operations management literature. For example, Kök and Fisher (2007) and Gallego and Topaloglu (2014) analyze a case where a single firm faces demand from consumers whose utility can be characterized by a logit or a nested logit model, respectively. Rusmevichientong et al. (2014) consider a case with mixed logit demand. Besbes and Saure (2014) and Kök and Xu (2011) consider competitive settings for a logit and nested logit model of demand,
respectively. Overall, the focus of this stream of research is to obtain and characterize optimal or equilibrium assortment decisions given an assumed demand and supply model. In contrast, the focus of this paper is to estimate a demand and supply model conditioning on data about the observed choices of firms and consumers (i.e., prices, products introduced and market shares in each market).

Finally, this paper is also related to the entry literature (see Aguirregabiria and Suzuki (2015) for a survey of empirical methods for estimating entry games). The connection is evident since if a firm can only introduce one product in a market, then the assortment problem can be modeled using an entry formulation. In terms of the extant literature on entry models, it is often the case that the profits of entering a market are modeled using (linear) functions of market characteristics and competitors’ entry decisions. In the structural approach proposed in this paper, the profit functions are instead formulated by modeling consumer demand and the firms’ pricing decisions. This provides an explicit link between the profits under each entry/assortment option and the corresponding equilibrium demand and prices. This is particularly relevant when performing policy simulations since the consequences of a structural change (e.g., related to cost functions or firm ownership) can be decomposed in terms of changes in prices, market shares and profits.

In summary, this paper makes four contributions. First, a structural model is proposed, which jointly considers product line and pricing decisions using a demand system that accounts for both unobserved heterogeneity in consumer preferences and unobserved product characteristics, two central elements in demand estimation based on aggregate data that have not been simultaneously considered in the extant research. Second, this paper describes and documents the extent by which overoptimistic demand inferences are obtained when product line endogeneity is ignored using both a Monte Carlo simulation study and empirical results. Third, it proposes a method to address this problem, which takes advantage of recent econometric advances to estimate structural models based on the use of a mathematical programming with equilibrium constraints approach (MPEC, see Su and Judd 2008). Fourth, the proposed method adds to the literature on two-step and iterative methodologies to estimate discrete games of incomplete information (e.g., Bajari et al. (2007); Ellickson and Misra (2008); Aguirregabiria and Mira (2007)). These methods are successful at reducing the complexity of the estimation task by using the data to obtain consistent estimates of firms choice probabilities (e.g., probability of entering a market), which are then employed
in a second (or successive) step(s) to estimate the model parameters. This paper considers
the case in which the firms’ choice probabilities may depend on product characteristics that
are unobserved to the econometrician, but observed by the firms and, consequently, pro-
vides an iterative method to account for this dependency. This is accomplished by relying
on parallel computing, since it is shown that it is possible to decompose some of the most
computationally intensive steps into a series of independent and much smaller problems that
can be solved in parallel. Finally, this paper shows how the estimated model can be used
to perform policy experiments that are relevant for managers and regulators interested in
assessing market responses to product line changes.

The rest of this paper is structured as follows. Section 2 describes the model of prod-
uct line and pricing decisions and characterizes consumer demand. Section 3, documents
the magnitude of the assortment endogeneity bias using a Monte Carlo simulation study.
This section also introduces an estimation strategy to correct this bias. Section 4 applies
this methodology to a real data set on liquid laundry detergent market shares, prices and
assortment. Section 5 concludes the paper with a discussion of directions for future research.

2 Model

This section introduces the model used to characterize consumer demand and describes the
game theoretic formulation employed to represent the strategic interaction among firms in
terms of their pricing and product line decisions.

2.1 Consumer Demand Model

First, it is necessary to introduce some notation about the set of alternatives that a consumer
may face. Let $J_f$ denote the number of products that firm $f$ may consider introducing in
any given market. For instance, in the case of ready-to-eat breakfast cereal, each of the $J_f$
products may have a different box size, package design or ingredients. In the case of mp3
players, these $J_f$ alternatives may have different storage capacities and functionalities.

Using this definition of $J_f$, the number of all possible combinations of products that firm
$f$ may introduce in a market corresponds to $2^{J_f}$. For example, if firm $f$ can produce $J_f = 2$
versions of a product (small, large), then there are $2^2$ possible combinations: a) introducing
the small version, b) introducing the large version, c) introducing both versions and d) not introducing any of the products. Furthermore, let $\mathcal{A}_f$ be the set of all the possible product lines that can be offered by firm $f$ in any given market (i.e., the cardinality of $\mathcal{A}_f$ is equal to $2^{J_f}$). In the previous example, $\mathcal{A}_f = \{\{\text{small}\}, \{\text{large}\}, \{\text{small}, \text{large}\}, \emptyset\}$, where $\emptyset$ denotes the empty set.

In terms of the demand model, each consumer in a given market $m$ chooses an option among a set of alternatives that includes all products introduced by the firms and an outside good. In the case of the purchase options, a consumer $i$ in market $m$ derives utility $U_{ifjm}$ from alternative $j$ from firm $f$:

$$U_{ifjm} = V_{ifjm} + \epsilon_{ifjm},$$

$$V_{ifjm} = \alpha'_{im}X_{fjm} + \beta_{im}p_{fjm} + \xi_{fjm},$$

where $V_{ifjm}$ is defined as the sum of $\alpha'_{im}X_{fjm}$, $\beta_{im}p_{fjm}$ and $\xi_{fjm}$; $\alpha_{im}$ is a vector of utility coefficients for consumer $i$ in market $m$; $X_{fjm}$ is a vector of product attributes for alternative $j$ offered by firm $f$ in market $m$; $\beta_{im}$ denotes consumer $i$'s price coefficient; $p_{fjm}$ is the price of alternative $j$ offered by firm $f$ in market $m$; $\xi_{fjm}$ is an unobserved product characteristic (also referred to as a unobservables for brevity) that affects the utility of all consumers in market $m$ for alternative $j$ offered by firm $f$; and $\epsilon_{ifjm}$ is an individual-specific demand shock for alternative $j$ offered by firm $f$ in market $m$. In addition, the utility of the outside good is given by: $U_{i0m} = \epsilon_{i0m}$.

As is common in the empirical literature, assuming each $\epsilon_{ifjm}$ is distributed according to an extreme value distribution $\text{EV}(0,1)$, the probability that consumer $i$ will choose alternative $j$ offered by firm $f$ in market $m$ is given by:

$$s_{ifjm} = \frac{\exp(\alpha'_{im}X_{fjm} + \beta_{im}p_{fjm} + \xi_{fjm})}{1 + \sum_{f'=1}^{F} \sum_{k \in a_{f'm}} \exp(\alpha'_{im}X_{pf'km} + \beta_{im}p_{f'km} + \xi_{f'km})},$$  

(1)

where $a_{f'm} \in \mathcal{A}_f$ is the set of products offered by firm $f$ in market $m$. For notational convenience, define $a_m$ as the set of products introduced by all firms (i.e., $\bigcup_{f=1}^{F} a_{f'm}$). Furthermore, if each vector of utility coefficients $\theta_{im} \equiv (\alpha'_{im}, \beta_{im})'$ is distributed according to a multivariate
normal distribution with mean $\overline{\theta}_m$ and variance $\Lambda$, expected market shares are given by:

$$s_{fjm} = \int \frac{\exp(\alpha'X_{fjm} + \beta p_{fjm} + \xi_{fjm})}{1 + \sum_{f'=1}^{F} \sum_{k \in a'_{f'm}} \exp(\alpha'X_{f'km} + \beta p_{f'km} + \xi_{f'km})} \phi(\theta; \overline{\theta}_m, \Lambda) d\theta,$$

(2)

where $\phi(\cdot; \overline{\theta}_m, \Lambda)$ denotes the probability density of a multivariate normal random variable with mean $\overline{\theta}_m$ and variance $\Lambda$. Finally, it is assumed that unobserved product characteristics $\xi_{fjt}$ are Gaussian with zero mean and variance $\varphi^2_\xi$. Note that Jiang et al. (2009) provide Monte Carlo evidence showing that demand estimation results are not very sensitive to departures from this assumption (e.g., heteroskedasticity, autocorrelation and asymmetry of the unobserved product characteristics).

### 2.2 Supply Model

Now consider the two decisions faced by each firm in our model: i) the products that will be made available in a market and ii) the prices that will be charged for each alternative. Accordingly, this section begins by defining the profit function of each firm, which will be used to characterize both pricing and product line decisions. Given that the profitability of a product line depends on the expected price competition each firm will face, this section will also describe the game theoretic assumptions about pricing decisions and then the strategic assumptions about product line introduction choices. For ease of exposition, it is assumed that the number of firms is equal to two (i.e., $F = 2$). It is straightforward to extend this model to allow for any arbitrary number of firms, although evidently the computational cost of the estimation method will evidently grow with the number of firms.

#### 2.2.1 Variable and Fixed Profits

The profits obtained by each firm have two main components: variable and fixed profits. The first component (variable profits) depends on the prices, market shares and marginal costs of each of the alternatives in a firm’s product line. In particular, the variable profits obtained by firm $f$ in market $m$ when it introduces product line $a_{fm} \in A_f$ and its competitor
introduces product line \( a_{f'm} \in A_f' \) are given by:

\[
\pi_{\text{Var}}(a_{f'm}|a_{f'm}) = \sum_{j \in a_{f'm}} (p_{fjm} - c_{fjm})s_{fjm},
\]  

(3)

where \( p_{fjm} \) and \( c_{fjm} \) are the price and (constant) marginal cost for product \( j \) offered by firm \( f \) in market \( m \). Furthermore, marginal costs \( c_{fjm} \) are modeled as a function of a vector of cost shifters \( Z_{fjm} \) (e.g., input prices) that explain variation across markets in marginal costs: \( c_{fjm} = \gamma_f Z_{fjm} \), where \( \gamma_f \) is a vector of coefficients. The second component of the profit function of each firm includes the fixed benefits and costs of introducing each of the alternatives in the product line (fixed profits), which by definition do not depend on the sales of each of the products. Items contributing to these fixed profits may include, for example, the necessary investments in manufacturing and distribution to make a particular product line \( a_{f'm} \) available in market \( m \). Denote this second component by \( \pi_{\text{Fixed}}(a_{f'm}) \).

It is important to note that there might be common factors that determine the profits of two different product lines of the same firm. In terms of the example described in the previous subsection, the following two product lines: \{small\} and \{small, large\} have one variety in common. First, the variable profits for both product lines will depend on the unobserved product characteristics (\( \xi \)) for the small product, inducing a correlation between the profits from these two product lines that include the small product. Second, it might be expected that the fixed costs associated with the introduction of the small version to contribute to the profits of both product lines. In order to capture the effect of these common factors, \( \pi_{\text{Fixed}}(a_{f'm}) \) can be formulated as follows:

\[
\pi_{\text{Fixed}}(a_{f'm}) = \sum_{j \in a_{f'm}} \kappa_{fj} + \varepsilon_{a_{f'm}},
\]  

(4)

where the summation takes into account that if an alternative \( j \) appears in two different product lines, then the associated term \( \kappa_{fj} \) will be used to compute the fixed profits of both product lines. In addition, \( \varepsilon_{a_{f'm}} \) captures all other factors contributing to the fixed profits when firm \( f \) introduces product line \( a_{f'm} \) in market \( m \) and it is assumed to be private to firm \( f \). Assuming that this information is private to each firm, competing firms will have prior beliefs about \( \varepsilon_{a_{f'm}} \) which for mathematical tractability are assumed to be characterized by an

---

1Please note that a slightly different formulation of these fixed profits will be used in the empirical section.
extreme value distribution $\text{EV}(0, \frac{1}{\mu})$. Note that if $\mu$ approaches infinity, then the magnitude of each $\varepsilon_{a_{fm}}$ tends to zero and, in the limit, all the elements of the profit function become common knowledge to all firms. At the same time, the full support of the private shocks implies that every possible product line $(a_{fm})$ has a non-zero probability of being introduced in a given market, even though this probability could be very small depending on the values of $\pi_{\text{Var}}(a_{fm})$ and $\pi_{\text{Fixed}}(a_{fm})$.

Based on these definitions, the total profits obtained by firm $f$ when it introduces product line $a_{fm}$ and its competitor $f'$ introduces $a_{f'm}$ are given by:

$$\Pi(a_{fm}|a_{f'm}) = \pi_{\text{Var}}(a_{fm}|a_{f'm}) + \pi_{\text{Fixed}}(a_{fm}).$$

Finally, denote by $\overline{\Pi}(a_{fm}|a_{f'm})$ the elements of $\Pi(a_{fm}|a_{f'm})$ that are common knowledge to all players:

$$\overline{\Pi}(a_{fm}|a_{f'm}) = \sum_{j \in a_{fm}} (p_{fjm} - c_{fjm}) s_{fjm} + \sum_{j \in a_{fm}} \kappa_{fj},$$

which implies that total profits can be expressed as the sum between the common-knowledge profits $\overline{\Pi}(a_{fm}|a_{f'm})$ and the private-information component $\varepsilon_{a_{fm}}$:

$$\Pi(a_{fm}|a_{f'm}) = \overline{\Pi}(a_{fm}|a_{f'm}) + \varepsilon_{a_{fm}}.$$  

The next subsection characterizes the price equilibrium, which requires us to focus on the variable profits obtained by each firm $\pi_{\text{Var}}(a_{fm}|a_{f'm})$.

### 2.2.2 Pricing Decisions

A Bertrand-Nash formulation to characterize pricing decisions among the competing firms is proposed. Accordingly, firms simultaneously set prices for their products and, therefore, when each firm $f$ introduces product line $a_{fm}$ in market $m$, equilibrium prices must satisfy the following first order conditions which are obtained by differentiating the variable profits $\pi_{\text{Var}}(a_{fm}|a_{f'm})$ with respect to each price $p_{fjm}$:

$$s_{fjm} + \sum_{l \in a_{fm}} (p_{flm} - c_{flm}) \frac{\partial s_{flm}}{\partial p_{fjm}} = 0, \quad j \in a_{fm}, \ f = 1, \ldots, F, \ m = 1, \ldots, M.$$  

10
Denote by $p^*_fjm(a_m)$ and $s^*_fjm(a_m)$ the equilibrium price and market share, respectively, of alternative $j$ offered by firm $f$ in market $m$ when firms introduce products in $a_m$ (note that market shares depend on unobserved product characteristics $\xi$). The equilibrium prices and shares will then satisfy equation (8) and they can be used to determine common knowledge equilibrium profits: $\Pi^*(a_f|a_{f_m})$.

In the case of the introduced products ($a^o_m$), following Su and Judd (2012), assume that the data on equilibrium prices contain measurement error, such that the researcher does not directly observe $p^*_fjm(a^o_m)$, but instead:

$$
p^o_{fjm}(a^o_m) = p^*_fjm(a^o_m) + \eta_{fjm},
$$

where $\eta_{fjm}$ is Gaussian with zero mean and variance $\varphi^2_{\eta}$. Allowing for measurement error is consistent with the findings of Einav et al. (2010), who checked the consistency between store and panel data, and more importantly makes the estimation problem more tractable. In addition, note that if the measurement error were negligible, $\varphi^2_{\eta}$ would be estimated to be very small.

### 2.2.3 Product Line Decisions

As described in equation (7), the profits of each firm have a private information component ($\varepsilon_{a_{f_m}}$). Therefore, the analysis of the strategic interactions between the firms in terms of their product line decisions can be characterized using a Bayesian-Nash equilibrium concept. Under this incomplete information setup, each firm forms beliefs about the probability that its competitor will introduce a particular set of products in a given market. In particular, denote by $\sigma_{a_{f_m}}$ the belief held by firm $f$’s competitor about the likelihood of firm $f$ introducing the set of alternatives $a_{f_m}$. These beliefs can then be used by each firm to compute the expected profits it would obtain if it introduces a particular set of products in the market. For example, firm 1 can estimate the expected profits it would obtain in market $m$ from product line $a_{1m} \in \mathcal{A}_1$ as follows:

$$
E[\Pi^*_{1m}(a_{1m})] = \sum_{a_{2m} \in \mathcal{A}_2} \sigma_{a_{2m}} \Pi^*_1(a_{1m}|a_{2m}) + \varepsilon_{a_{1m}},
$$
where the summation considers the profits that firm 1 would obtain under all possible product lines that firm 2 could introduce (i.e., $a_{2m} \in A_2$) and each of these terms is weighted by firm 1’s belief about firm 2’s product line choice ($\sigma_{a_{2m}}$).

In equilibrium, each firm will introduce the product line that maximizes its expected profits $E[\Pi^*_{fm}(a_{fm})]$. Therefore, given that $\varepsilon_{a_{fm}} \sim \text{EV}(0, \frac{1}{\mu})$, the probability that firm $f$ will introduce product line $a_{fm}$ corresponds to:

$$
\sigma_{a_{fm}} = \frac{\exp\{\mu(\sum_{a'_{fm} \in A_f} \sigma_{a'_{fm}} \Pi^*(a_{fm} | a'_{fm}))\}}{\sum_{a'_{fm} \in A_f} \exp\{\mu(\sum_{a'_{fm} \in A_f} \sigma_{a'_{fm}} \Pi^*(a'_{fm} | a'_{fm}))\}},
$$

(10)

This condition needs to be satisfied for every possible product line $a_{fm} \in A_f$, firm $f \in \{1, 2\}$ and market $m = 1, \ldots, M$, which defines a fixed point relationship in the space of firm product introduction beliefs.

### 3 Model Estimation

This section discusses estimation issues related to the model presented in Section 2. First, it is explained why a method that ignores the endogeneity of product line decisions may lead to biased demand estimates. Moreover, this section provides evidence from a Monte Carlo simulation study which illustrates the magnitude and prevalence of this problem. Finally, an iterative MPEC approach (Su and Judd (2012)) is proposed to estimate the primitives of the structural model accounting for the endogeneity of both price and product line decisions.

#### 3.1 Product line endogeneity and sample selection bias

One of the most basic assumptions in demand estimation of random-coefficient demand models from aggregate data is that the mean of unobserved product characteristics ($\xi_{fjm}$) is zero. This assumption is used in simulated GMM approaches (e.g., Berry et al. (1995)) and also in likelihood-based methods (e.g., Park and Gupta (2009)). If it is acknowledged, however, that firms are more likely to introduce the alternatives that lead to the highest profits or that generate a sufficient level of demand, this assumption may not hold. In particular, the mean of the unobserved product characteristics of the introduced products...
(i.e., the products for which a researcher typically observes price and sales data), $E[\xi_{jm} | j \in a_{fm}]$, may not be equal to zero. This introduces an estimation problem that it is conceptually similar to the one described in Heckman (1979). In this subsection, it is explained why this problem may lead a researcher to overestimate total demand in a given market.

Considering a very simple example will be sufficient to illustrate the connection between the sample selection problem described in Heckman (1979) and the estimation problem discussed in this paper. Suppose a monopolist can produce $J$ varieties of a product and offers different sets of products in different markets (e.g., products 1 and 2 in market 1 and products 1 and 3 in market 2). For simplicity, demand is characterized by a homogenous logit model where the utility of alternative $j$ for a consumer $i$ in market $m$ is given by: $U_{ijm} = \alpha + \xi_{jm} + \epsilon_{ijm}$. As before, the utility of the outside good is given by: $U_{iom} = \epsilon_{iom}$. Note that $\alpha$ is related to total category demand given that larger values of this parameter increase the share of all the purchase options. Also assume that $\epsilon_{ijm} \sim \text{EV}(0,1)$ and each $\xi_{jm}$ is i.i.d. Normal(0,1). Suppose the firm uses a very simple rule to select its product line such that only products that generate enough sales compared to the outside good option are introduced in a given market. Defining $\delta$ as a minimum demand threshold, the firm could choose to introduce product $j$ in market $m$ only if $s_{jm} > \delta$, which is equivalent to the following condition: $\xi_{jm} > \ln(\delta) - \alpha$.

Consequently, the unobserved characteristics of introduced products will no longer be i.i.d. Normal(0,1), but instead their distribution corresponds to a truncated Normal(0,1), where the truncation is such that all unobserved characteristics of introduced products ($\xi_{jm}$) smaller than $\ln(\delta) - \alpha$ have zero probability. Therefore, in this case, the mean of the introduced products will no longer be equal to zero, but instead it will be equal to: $\phi(\ln(\delta) - \alpha) \Phi(\ln(\delta) - \alpha)$, where this expression corresponds to the inverse Mills ratio (Heckman (1979)) and $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and cumulative distribution function, respectively, of a standard normal distribution. Thus, if this form of sample selection is ignored, a researcher might estimate $\alpha$ by maximizing the likelihood of the market shares for the introduced products:

$$L(s) = \prod_{m=1}^{M} \prod_{j \in a_{m}} \phi(\ln(s_{jm} / s_{om}) - \alpha),$$

where $L(s)$ denotes the likelihood of the market share $(s)$ data for the introduced products.

It is easy to see that this optimization would lead $\alpha$ to be overestimated on average by a
magnitude equal to the mean of the unobserved characteristics of the introduced products: 
\[ E[\xi_{jm}|j \in a_m] = \frac{\phi(\ln(\delta) - \alpha)}{1 - \Phi(\ln(\delta) - \alpha)} > 0. \] Given that \( \alpha \) is directly related to the demand of the purchase options, this would lead to overoptimistic estimates of demand.

One approach to address this problem is to maximize a likelihood function that not only considers the likelihood of the observed market shares but also the the probability of the observed product lines: \(^2\)

\[
L(s, a) = \prod_{m=1}^{M} \left( \prod_{j \in a_m} \phi(\ln(s_{jm}/s_{om}) - \alpha) \right) \left( \prod_{j \in a_m} 1_{\{\xi_{jm} > \ln(\delta) - \alpha\}} \right) \left( \prod_{j \notin a_m} \Phi(\ln(\delta) - \alpha) \right), \tag{12}
\]

where \( L(s, a) \) denotes the likelihood of the market share \( s \) and product introduction \( a \) data. It is important to note that in this case, where demand is characterized by a homogeneous logit model and the firm uses a very simple rule to introduce products in each markets, it is straightforward to address this sample selection bias, by maximizing the likelihood in equation (12) as opposed to the likelihood in equation (11). In more general cases, such as the one described in Section 2 where i) consumers have heterogenous preferences, ii) firms choose their prices and product lines to maximize their profits and iii) firms have imperfect information about each other, computing the likelihood of the data (demand, prices and product introduction) is evidently not as trivial as in this simple example. In fact, the complexity of the general model precludes us from deriving closed form expression for the estimation bias. Nevertheless, the same basic argument is applicable to the more general case if \( E[\xi_{jm}|j \in a_{fm}] \neq 0. \)

The next subsection describes a simulation study that shows the consequences of ignoring the endogeneity of product line decisions in the context of the structural model described in Section 2, while Subsections 3.3 and 3.4 propose an estimation strategy that addresses this problem.

### 3.2 Monte Carlo Evidence: A Simulation Study

A simulation study is conducted based on the structural model described in Section 2. I consider \( F = 2 \) firms making price and product introduction decisions in \( M = 30 \) markets.

\(^2\)This approach is conceptually equivalent to maximum likelihood methods for censored regression models (see Pagan and Ullah (1999), p. 321).
Each firm is able to produce and sell $J_f = 3$ varieties for any given market. Five variables $(x_{fjm})$ are included in the utility function of each consumer: i) a dummy variable $(x_{fjm,1})$ equal to 1 for all purchase alternatives, ii) two variables $(x_{fjm,2}$ and $x_{fjm,3})$ generated from a normal distribution with zero mean and variance 0.1, iii) a fourth variable generated from a uniform distribution in the $(-0.5, 0.5)$ interval and iv) the price of each alternative $p_{fjm}$.

The mean of consumer coefficients is assumed to be equal across markets, i.e. $\theta_m = \bar{\theta}$, $m = 1, \ldots, M$, where $\bar{\theta} = (3, 0.5, -0.25, -1, -2)$. In this study some coefficients are heterogenous (i.e., random) across consumers while others are fixed. In particular, in this simulation the standard deviation of the random-coefficients is zero for the first four variables and it is equal to 0.5 for the price coefficients (i.e., $\sqrt{\Lambda_{5,5}} = 0.5$).

Unobserved product characteristics $(\xi_{fjm})$ and price measurement errors $(\eta_{fjm})$ are generated from zero-mean normal distributions with standard deviations equal to $\varphi_\xi = 1$ and $\varphi_\eta = 0.1$. As described in Section 2, marginal costs are given by: $c_{fjm} = \gamma_f Z_{fjm}$. Cost shifters $(Z_{fjm})$ include a dummy variable for each product of each firm and a second variable generated from a uniform distribution in the $(0.5, 1.5)$ interval. The coefficients associated with the first component of $(Z_{fjm})$ correspond to $\gamma_{1,1} = 0.4$ for firm 1 and $\gamma_{2,1} = 0.1$ for firm 2, while the coefficients for the second component are equal to $\gamma_{1,2} = 0.2$ for firm 1 and $\gamma_{2,2} = 0.3$ for firm 2. In terms of the fixed profits, the parameters $(\kappa_{fj})$ are equal across alternatives (i.e., $\kappa_{fj} = \kappa$ for all $f$ and $j$), where $\kappa = -0.2$. Finally, the scale $(\mu)$ of the private information term of the profit function is equal to 15.

Given these parameter values, market share, price and product introduction data for every market $m$ are generated as follows:

1. Bertrand-Nash price equilibrium: For every possible pair of product lines $(a_{fm}, a_{f'm})$, a vector of prices that satisfy the first order condition in equation (8) is found (a second order check is also implemented to ensure the solution found is indeed an equilibrium) using the fixed point method proposed by Morrow and Skerlos (2011). Note that given that $J_f = 3$ for both firms, the number of possible pairs of product lines is equal to $2^3 \times 2^3 = 64$. Using these equilibrium prices, equilibrium market shares and equilibrium variable profits are computed for each of these 64 cases, where these market shares were obtained by using 500 simulated consumers.

2. Bayesian-Nash equilibrium: Given the equilibrium variable profits obtained for each pair of product lines in each market, a set of product introduction beliefs for each firm
that satisfy the equilibrium conditions in equation (10) is found.

3. Sampling equilibrium product lines: a product line \((a_{fm}^o)\) is sampled for each firm from a multinomial distribution where the probability of each product line \((a_{fm})\) is equal to the equilibrium product introduction belief obtained in the previous step.

4. Market share data and sampling observed prices: The market share data \((s_{fjm}^o)\) for a given market corresponds to the equilibrium market shares obtained in the first step for the pair of sampled product lines \((a_{fm}^o, a_{f'm}^o)\). The observed prices \((p_{fjm}^o)\) are equal to the equilibrium prices under \((a_{fm}^o, a_{f'm}^o)\) plus the corresponding measurement error \((\eta_{fjm})\).

Simulated data are generated for 30 replications. Table 1 displays summary statistics for unobserved product characteristics across all replications. As expected, it is observed that the mean of these unobserved characteristics for the introduced products is higher than the one corresponding to the non-introduced products. In fact, the fraction of positive unobserved characteristics for the introduced products (0.79) is more than twice as big as the one for the non-introduced products (0.35). Therefore, introduced products are more likely to have higher and positive values of the unobserved characteristics. Consequently, if the endogeneity of product lines is ignored, one should observe a positive bias in the estimation of the total demand, given that the mean of the unobserved characteristics of the introduced products is greater than zero. In particular, one would expect \(\theta_1\), the parameter that affects the utility of all purchase alternatives, to be upwardly biased (estimation results are discussed in the next paragraphs). Moreover, it is also noted that the standard deviations of the unobserved product characteristics in both subgroups (introduced and non-introduced products) are smaller than 1, the value of the standard deviation of these unobservables when considering all products. This indicates some degree of homogeneity within each of these groups and suggests that the standard deviation of these unobserved product characteristics \((\varphi \xi)\) might be underestimated if product line endogeneity is ignored.

Accordingly, parameter estimates for each of the simulated data sets were first obtained ignoring the endogeneity of product line decisions.\(^3\) This is accomplished by maximizing the

\(^3\)Results for the estimation method that accounts for product-line endogeneity are presented later in this section.
Table 1: Summary statistics of the unobserved product characteristics ($\xi$).

<table>
<thead>
<tr>
<th></th>
<th>Introduced Products</th>
<th>Non-Introduced Products</th>
<th>All Products</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.67</td>
<td>-0.35</td>
<td>-0.01</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>0.88</td>
<td>0.86</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>Fraction positive</strong></td>
<td>0.79</td>
<td>0.35</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>Number of cases</strong></td>
<td>1,915</td>
<td>3,485</td>
<td>5,400</td>
</tr>
</tbody>
</table>

The likelihood of the market shares and prices subject to a set of equilibrium constraints:

$$
\max_{\bar{\theta}_m, \Lambda, \gamma, \xi, \eta^2, \varphi, \varphi_\eta} \prod_{m=1}^{M} |J_m|^{-1} \prod_{f=1}^{F} \prod_{j \in a_{f,m}} \phi(\xi_{fjm}; 0, \varphi_\xi^2) \phi(\eta_{fjm}; 0, \varphi_\eta^2)
$$

subject to:

$$
0 = s_{fjm}^* - \frac{1}{R_\theta} \sum_{i=1}^{R_\theta} \exp(\alpha_{im} X_{fjm} + \beta_{im} p_{fjm}^* + \xi_{fjm})
$$

$$
\sum_{f=1}^{F} \sum_{k \in a_{f,m}} \exp(\alpha_{im} X_{fkm} + \beta_{im} p_{fkm}^* + \xi_{fkm})
$$

$$
\forall f, \forall m, j \in a_{f,m},
$$

$$
0 = s_{fjm}^* + \sum_{l \in a_{f,m}} (p_{flm}^* - c_{flm}) \frac{\partial s_{fjm}^*}{\partial p_{fjm}^*}
$$

$$
\forall f, \forall m, j \in a_{f,m},
$$

where $\theta_{im} = (\alpha_{im}, \beta_{im})'$ is a draw from a multivariate normal distribution with mean $\bar{\theta}_m$ and variance $\Lambda$; $R_\theta$ is the number of draws of $\theta_{im}$ used to approximate the market share of the random-coefficients logit model (see equation 1); and $|J_m|$ is the determinant of the Jacobian of the transformation of unobserved product characteristics ($\xi$) into equilibrium market shares ($s$) and depends on all utility coefficients and the unobserved characteristics of the introduced products.

In terms of the two sets of constraints specified in equations (14) and (15), the first set requires expected market shares to match observed market shares as in Berry et al. (1995), while the second set requires equilibrium prices, which are decision variables in this constrained optimization problem, to satisfy a profit-maximization first-order condition.
Table 2: Simulation study ignoring product line endogeneity. Summary statistics of the estimated demand parameters.

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\theta}_1 )</th>
<th>( \hat{\theta}_2 )</th>
<th>( \hat{\theta}_3 )</th>
<th>( \hat{\theta}_4 )</th>
<th>( \hat{\theta}_5 )</th>
<th>( \sqrt{ \Lambda_{5,5} } )</th>
<th>( \varphi_\xi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>3.00</td>
<td>0.50</td>
<td>-0.25</td>
<td>-1.00</td>
<td>-2.00</td>
<td>0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>Mean</td>
<td>3.71</td>
<td>0.74</td>
<td>0.03</td>
<td>-0.77</td>
<td>-2.01</td>
<td>0.51</td>
<td>0.84</td>
</tr>
<tr>
<td>SD</td>
<td>0.34</td>
<td>2.02</td>
<td>2.58</td>
<td>0.64</td>
<td>0.21</td>
<td>0.15</td>
<td>0.10</td>
</tr>
<tr>
<td>( P_{\text{above}} )</td>
<td>1.00</td>
<td>0.47</td>
<td>0.53</td>
<td>0.67</td>
<td>0.50</td>
<td>0.47</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Note: \( P_{\text{above}} \) corresponds to the fraction of replications in which a parameter is overestimated.

Note that the likelihood of the market share data is determined using a change of variables approach. Specifically, the probability of a combination of market shares is obtained as a function of the unobserved product characteristics (\( \xi \)). This approach is valid since after conditioning on \( \bar{\theta}_m \) and \( \Lambda \) there is a one-to-one relationship between unobserved product characteristics and market shares (Berry et al. (1995)). The likelihood of prices is directly derived by evaluating the density of the measurement errors (\( \eta_{fjm} \)), which are in turn defined by equation (9). Note that unobserved product characteristics (\( \xi \)) will affect market shares and equilibrium prices and, hence, price measurement errors. Therefore the values of these unobserved characteristics are not only chosen to match expected and observed market shares, but also to make the pricing data more likely.

Table 2 shows a summary of the estimation results obtained after solving this constrained optimization problem for each replication.\(^4\) From these results it is first noticed that the average bias of \( \hat{\theta}_1 \) is equal to 0.71. This implies that the results overestimate the total demand for the products and, in fact, it is verified that the estimate of this parameter is greater than its true value for all replications. As expected, this bias is comparable to the mean of the unobservables of the introduced products reported in Table 1 (0.63). Similarly, the mean of the estimated standard deviation of these unobservables (\( \varphi_\xi \)) is equal to 0.84, which is close to the standard deviation of these unobservables for the introduced products (0.88) reported in Table 1. Moreover, the estimated standard deviation is smaller than the true value for 90% of the replications.

\(^4\)In this simulation study \( R_\theta = 100 \) draws are used to compute expected market shares, and a different random seed was used for each simulated data set.
Most of the remaining parameters do not exhibit large deviations from their true values. In particular, note that the mean and variance of the price coefficient are estimated with a small bias. This is not surprising given that the estimation method accounts for the endogeneity of prices. Nevertheless, this does not imply that price elasticities of demand or sales projections will be unbiased, since these quantities also depend, for example, on the intercept of the utility function and the unobserved product characteristics. For example, if a researcher is interested in estimating the demand for an alternative that was not introduced, the results from a method that ignores product-line endogeneity will be overoptimistic.

Specifically, even though unobserved product characteristics are estimated for introduced products (by matching observed and predicted market shares), estimates for the unobserved characteristics of non-introduced alternatives are not obtained from the model estimation. These unobservables are needed, for example, to perform counterfactual analysis since firms make introduction decisions conditioning on these unobservables. Therefore, any inferences about the non-introduced products would have to be made integrating over the distribution of their unobservables. Consequently, if product line endogeneity is ignored, the researcher would not be using the correct distribution for these unobservables. This can be further understood by noticing that the unconditional distribution of unobserved product characteristics of the non-introduced alternatives (i.e., Normal$(0, \varphi^2_\xi)$ as in our model) is different from the distribution of these unobservables conditional on the observed product lines:

\[
f(\xi_m) = \phi(\xi_m; 0, \varphi^2_\xi) \neq f(\xi_m|a^o_{f,m}, a'_{f,m}) = \frac{\sigma_{a^o_{f,m}}(\xi_m)\sigma_{a'_{f,m}}(\xi_m)\phi(\xi_m; 0, \varphi^2_\xi)}{\int \sigma_{a^o_{f,m}}(\xi)\sigma_{a'_{f,m}}(\xi)\phi(\xi; 0, \varphi^2_\xi)d\xi},
\]

where the last equality follows from Bayes’ theorem; $\xi_m$ is a vector of unobserved characteristics for the introduced and non-introduced alternatives of all firms in market $m$; and $\sigma_{a^o_{f,m}}(\xi_m, \eta_m)$ is the probability of firm $f$ introducing the observed product line $a^o_{f,m}$ when unobserved product characteristics are equal to $\xi_m$. Therefore, if a researcher relies on the unconditional distribution of unobserved product characteristics $f(\xi_{f,m})$, wrong conclusions will be drawn about demand projections, equilibrium prices and profits.

In order to illustrate this issue, the following policy experiment is performed. For each of the 30 replications and for every market the Bertrand-Nash equilibrium prices, market shares and profits that would be obtained if one additional product was introduced are computed. In each case, the true value of the equilibrium quantities and their estimated values using the parameters obtained when the endogeneity of product line decisions is ignored are computed.
Figure 1: Histogram of the differences between true and estimated equilibrium prices, market shares and profits of the new products ignoring (top) and accounting (bottom) for the endogeneity of product line decisions.

The comparison between these two sets of quantities provides an opportunity to assess the consequences of not accounting for the endogeneity of product introduction decisions. The top panel of Figure 1 shows a histogram of the differences between the true and estimated equilibrium prices, market shares and profits for each of the products added to the product line (new products) when product line endogeneity is ignored. Similarly, the top panel of Figure 2 shows the corresponding differences for products belonging to the introduced product lines (existing products). In addition, the first two columns in Table 3 show summary statistics for the deviations between the estimated and true values of each of the quantities (prices, market share and profits) for both new and existing products.

In the case of the new products (Figure 1, top panel), on average the model that ignores product line endogeneity overestimates prices, market shares and profits yielding overoptimistic estimates about the performance of these products. In terms of the existing products (see the top panel in Figure 2 and the first two columns of results in Table 3), even though there is a small positive bias for equilibrium prices, market shares and consequently profits
Figure 2: Histogram of the differences between true and estimated equilibrium prices, market shares and profits of the existing products ignoring (top) and accounting (bottom) for the endogeneity of product line decisions.
Table 3: Summary statistics of the deviations between estimated and true equilibrium prices, shares and profits, when ignoring and accounting for product line (PL) endogeneity.

<table>
<thead>
<tr>
<th></th>
<th>Ignoring PL Endogeneity</th>
<th>Accounting for PL Endogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>New Products</td>
<td>Existing Products</td>
</tr>
<tr>
<td>Prices</td>
<td>Mean</td>
<td>0.0932</td>
</tr>
<tr>
<td></td>
<td>S.E. of Mean</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>Fraction positive</td>
<td>0.8884</td>
</tr>
<tr>
<td></td>
<td>Number of cases</td>
<td>3,485</td>
</tr>
<tr>
<td>Shares</td>
<td>Mean</td>
<td>0.1126</td>
</tr>
<tr>
<td></td>
<td>S.E. of Mean</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>Fraction positive</td>
<td>0.9105</td>
</tr>
<tr>
<td></td>
<td>Number of cases</td>
<td>3,485</td>
</tr>
<tr>
<td>Profits</td>
<td>Mean</td>
<td>0.1391</td>
</tr>
<tr>
<td></td>
<td>S.E. of Mean</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>Fraction positive</td>
<td>0.9156</td>
</tr>
<tr>
<td></td>
<td>Number of cases</td>
<td>3,485</td>
</tr>
</tbody>
</table>

are on average underestimated. This is explained by the overoptimistic predictions about the utility of the new products which lead this model to forecast too many consumers switching from the existing to the new products.

The next section proposes an approach to account for the endogeneity of product line and pricing decisions.

### 3.3 A Solution to the Product-Line Endogeneity Problem

As suggested in Subsection 3.1, an approach to account for the endogeneity of product line decisions is to formulate a likelihood function that considers not only the probability of the observed sales and pricing data, but also the likelihood of the observed product lines. This likelihood can be computed as a function of the equilibrium product introduction beliefs $\sigma_{a_{fm}}$ for each firm and market. Evaluating this likelihood, however, is not a trivial task, as explained below.

Suppose one would like to compute a set of equilibrium product introduction beliefs given values of the parameters that determine the demand for each of the introduced alternatives ($\theta_m, \Lambda, \xi_{fjm}$ for $j \in a_{fm}$), the marginal costs ($\gamma$) and profits ($\kappa_j$ and $\mu$). Let $\xi_{m,t}$ denote
the vector of unobserved characteristics of the introduced alternatives in market \( m \), while \( \xi_{m,N} \) denote the corresponding vector of observables for the non-introduced alternatives. Equilibrium product introduction beliefs are also a function of the observables of the non-introduced products \( (\xi_{m,N}) \), which are observed by the firms but not by the researcher. Accordingly, these observables need to be integrated out when computing the likelihood of a product line \( a_{fm} \) being introduced in the market:

\[
\sigma_{a_{fm}}(\xi_{m,I}) = \int_{\xi_{m,0}}^{\xi_{m,I}} \sigma_{a_{fm}}(\xi_{m,I},\xi_{m,N}) f(\xi_{m,N}) d\xi_{m,N}. \tag{17}
\]

Note that this result follows directly from the law of total probability. As a consequence, when computing the likelihood function it is necessary to use the unconditional distribution of the unobserved product characteristics \( f(\xi_{m,N}) \) in this integration (i.e., not the distribution of these observables conditional on the product line \( a_{fm} \)), while the term \( \sigma_{a_{fm}}(\xi_{m,I},\xi_{m,N}) \) takes into account the likelihood of the product line \( a_{fm} \) under each value of the observables for the non-introduced products \( (\xi_{m,N}) \).

The distinction between the conditional and the unconditional distribution of unobserved product characteristics is crucial. In particular, when estimating the model parameters it is necessary to specify a likelihood function which considers the probability of a set of product lines given all model parameters. Since unobserved product characteristics are not known, they can be integrated out using their unconditional distribution, as is common in a maximum simulated likelihood approach. If instead of computing the likelihood of the product line data one wished to estimate the most likely value of an unobserved product characteristic given the observed product lines, then one would need the conditional probability of that unobservable given the observed assortments. This distinction will become relevant again when performing policy simulations as those described later in this paper (see subsection 3.5).

In the context of the estimation of the model, the integration described in equation (17) can be performed, for example, using simulation:

\[
\hat{\sigma}_{a_{fm}}(\xi_{m,I}) = \frac{1}{R_{\xi}} \sum_{r=1}^{R_{\xi}} \sigma_{a_{fm}}(\tilde{\xi}_{m}^{(r)}), \tag{18}
\]

where \( R_{\xi} \) is the number of simulation draws used to estimate \( \sigma_{a_{fm}}; \sigma_{a_{fm}}(\tilde{\xi}_{m}^{(r)}) \) is the equil-
rium product introduction belief about the likelihood of firm $f$ introducing product line $a_{fm}$ when unobservable characteristics for all alternatives are equal to $\tilde{\xi}_m^{(r)}$; and $\tilde{\xi}_m^{(r)}$ is a vector of unobserved product characteristics for all alternatives in market $m$ such that the components associated with introduced products are equal to the estimated value of the corresponding unobservables (i.e., $\tilde{\xi}_m^{(r)} = \xi_{jm}$, $j \in a_{fm}$), while the values of the unobserved characteristics of the non-introduced products are generated from a zero-mean normal distribution with standard deviation equal to $\varphi_2^\xi$, respectively.

Using this approach to estimate $\sigma_{a_{fm}}$, the parameters of the structural model can be recovered by maximizing the likelihood of the market share, price and assortment data subject to a set of equilibrium constraints:

$$
\max_{\theta, \gamma, \mu, \alpha, \beta, \sigma, \kappa} \prod_{m=1}^M |J_m|^{-1} \prod_{f=1}^F \tilde{\sigma}_{a_{fm}}(\xi_{m,I}) \prod_{j \in a_{fm}} \phi(\xi_{fm,j}; 0, \varphi_2^\xi) \phi(\eta_{jm}; 0, \varphi_2^\eta) \tag{19}
$$

subject to:

$$
R_a \sum_{i=1}^R \frac{\exp(\alpha_{im} X_{fm,j} + \beta_{im} p_{fm,j} + \xi_{fm,j})}{1 + \sum_{f'=1}^F \sum_{k \in a_{f'm}} \exp(\alpha_{im} X_{f'km} + \beta_{im} p_{f'km} + \xi_{f'km})}, \quad \forall f, \forall m, j \in a_{fm}, 
$$

$$
0 = s_{fm,j} + \sum_{l \in a_{fm}} (p_{flm}(a_m) - c_{flm}) \frac{\partial s_{fm,j}}{\partial p_{fm,j}}, \quad \forall f, \forall m, \forall a_m, j \in a_{fm}, 
$$

$$
\sigma_{a_{fm}}(\tilde{\xi}_m^{(r)}) = \frac{\exp\{\mu(\sum_{a_{f'm} \in A_f} \sigma_{a_{f'm}}(\tilde{\xi}_m^{(r)}))\Pi^*(a_{fm}|a_{f'm}, \tilde{\xi}_m^{(r)})\}}{\sum_{a_{f'm} \in A_f} \exp\{\mu(\sum_{a_{f'm} \in A_f} \sigma_{a_{f'm}}(\tilde{\xi}_m^{(r)}))\Pi^*(a_{fm}|a_{f'm}, \tilde{\xi}_m^{(r)})\}}, \quad \forall f, \forall m, \forall a_{fm} \in A_f, \forall r,
$$

where $\tilde{\sigma}_{a_{fm}}(\xi_{m,I})$ is defined in equation (18); and $\Pi^*(a_{fm}|a_{f'm}, \tilde{\xi}_m^{(r)})$ denotes the common-knowledge equilibrium profits obtained by firm $f$ in market $m$ when it introduces product line $a_{fm}$, its rival introduces $a_{f'm}$ and unobserved product characteristics are given by $\tilde{\xi}_m^{(r)}$.

A common problem in the estimation of entry models and other multi-agent discrete choice games is the possible existence of multiple equilibria. In particular, there could be
more than one set of equilibrium product line introduction beliefs that meet constraint (22). The use of a maximum likelihood approach implies that among all sets of product introduction beliefs satisfying the Bayesian Nash equilibrium condition, the one that maximizes the likelihood of the data will be chosen (Su and Judd (2012)). From a practical standpoint, this implies that numerical optimization strategies aimed at achieving global (as opposed to local) optimality should be employed, such as the use of multiple initial values.

Under this approach, solving the constrained optimization problem, even for low-dimensionality instances (i.e., small values of $J_f, M, R_\theta$, and $R_\xi$), requires a large number of computations. In particular, note that in order to evaluate the last set of constraints of this problem it is necessary to compute the common-knowledge equilibrium profits $\Pi(a_{fm}|a_{f'm},\tilde{\xi}^{(r)}_m)$ for every possible pair of product lines that can be introduced by the firms in every market and for each of the $R_\xi$ draws of $\xi_m$. This in turn requires us to find a Bertrand-Nash price equilibrium for $M \cdot R_\xi \cdot 2^{J_1+J_2}$ instances just to make a single evaluation of this set of constraints. Interestingly, this problem is conceptually similar to the challenges faced when estimation consumers’ consideration sets, where computing the likelihood of consumer choices requires integrating over all possible consideration sets (e.g., van Nierop et al. (2010)).

In the next subsection, an iterative approach that substantially reduces the computational burden is proposed to estimate the model parameters.

### 3.4 Iterative estimation

A central aspect of the likelihood function in equation (19) is that unobserved product characteristics not only affect the likelihood of the market share and pricing data, but also the likelihood of the observed product lines given that the equilibrium product introduction beliefs $\sigma_{afm}(\xi_{m,l})$ are a function of the unobserved product characteristics (see equation (18)). This important feature captures the correlation between introduced product lines and unobservables for a given market and, therefore, formulating the likelihood of the observed product lines as a function of the unobservables enables us to explicitly account for the endogeneity of product line decisions.

Many of the extant two-step and iterative approaches for estimating incomplete information games (e.g., Bajari et al. (2007)) approximate the probability of a firm choosing a certain action using the data available to the researcher. For example, a frequency estimator could be employed where the fraction of times that a firm introduces a certain product line
across markets or periods is used to approximate these choice probabilities. These approximated probabilities are then used in successive steps to estimate the model parameters. These methods, however, can not be used in the context of the model discussed in this paper, because as described in the previous paragraph, the probability of a firm introducing a product line depends on unobserved product characteristics. If this dependency is ignored, the researcher would fail to take into account that unobserved product characteristics for introduced products may, for example, be more likely to be positive for introduced products than for non-introduced products. This in turn would lead to incorrect parameter inferences.

In contrast, the approach described in this subsection is based on iteratively approximating the relationship between product introduction and unobserved product characteristics (as opposed to observables). Specifically, equilibrium product introduction beliefs \( \sigma_{a_{fm}}(\tilde{\xi}_m^{(r)}) \) can be approximated by a flexible function of these unobservables denoted by \( g_{afm}(\tilde{\xi}_m^{(r)}; \rho) \). In this approximation, \( \rho \) is a vector of parameters to be estimated that determines the shape of \( g_{afm} \) as a function of \( \tilde{\xi}_m^{(r)} \). One simple example of such a function, is given by a choice (e.g., logit) model, where the probability of firm \( f \) choosing an assortment \( a_{fm} \) in market \( m \) is modeled as a function of all the unobserved product characteristics for market \( m \):

\[
g_{afm}(\tilde{\xi}_m^{(r)}; \rho) = \frac{\exp(W_{afmr})}{1 + \exp(W_{afmr})},
\]

(23)

where \( W_{afmr} \) is an \( n \)-degree polynomial of the vector of unobservables \( \tilde{\xi}_m^{(r)} \). For example, a quadratic polynomial, as employed in the subsequent analysis is given by:

\[
W_{afmr} = \upsilon_{fa} + \sum_{f'j'} \delta_{fa'j'}\tilde{\xi}_{f'j'm} + \sum_{f'j'n} \omega_{fa'j'n'j'n} \tilde{\xi}_{f'j'm} \tilde{\xi}_{f'n'j'n'm} \quad (24)
\]

and \( \rho \equiv (\upsilon, \delta, \omega) \). Given these definitions, an iterative estimation approach is presented below:

1. **Initialize** \( \rho \): Start by setting initial values for each of the components of \( \rho \).

2. **Price-Demand Parameters**: Then estimate the demand and marginal cost parameters \( (\bar{\theta}, \Lambda, \gamma, \xi_I, \xi, \zeta_n) \) and the equilibrium prices \( (p^*) \) for the observed assortment
replacing $\sigma_{a_f m}(\tilde{\xi}_m(r))$ by $g_{a_f m}(\tilde{\xi}_m(r); \rho)$ in (19), therefore, solving:

$$\max_{\theta, \gamma, \psi, \xi} \prod_{m=1}^{M} |J_m|^{-1} \prod_{f=1}^{F} \prod_{j \in a_f m} \phi(\xi_{fjm, \lambda}; 0, \varphi_f^2) \phi(\eta_{fjm}; 0, \varphi_n^2)$$

subject to constraint (20), which matches the estimated to the observed market shares, and constraint (21), which imposes best response conditions on prices. Note that this estimation problem is much simpler than the one described in equation (19) given that in this step it is not necessary to impose the set of constraints (22), which are related to the Bayesian-Nash equilibrium of product line decisions. This avoids the need of computing Bertrand-Nash profits for every assortment combination in every optimization iteration within this step. These profits are actually obtained in the next step.

3. **Bertrand-Nash Profits for every assortment combination:** Given demand and marginal cost parameters, Bertrand-Nash profits are computed. More specifically, for every combination of firm assortments ($a_m$) in every market $m$, and for every draw $r$ of the unobserved product characteristics $\tilde{\xi}_m(r)$, prices $p_{fjr \lambda m}(a_m)$ satisfying the first order conditions in (8) are found. This can be a very computationally intensive step, but it is also a task that lends itself very nicely to parallelization. For example, the price equilibrium problems for different ($a_m, m, r$) combinations can be assigned to different processor cores or threads and, hence, they can be solved simultaneously (i.e., in parallel). The output of this step is the set of Bertrand-Nash profits for each firm for every ($a_m, m, r$) combination: $\pi_{Y \lambda m}^*(a_{f m}|a_{f m}, \tilde{\xi}_m(r))$, which will be used in the next step to compute product line introduction equilibrium beliefs.

4. **Bayesian-Nash Product Line Introduction Beliefs:** Given the Bertrand Nash

---

5. This approach relies either on price equilibrium uniqueness or, if multiple price equilibria exist, on an assumed or known by the researcher equilibrium selection rule (e.g., Pareto efficiency). In the context of a pricing game with demand characterized by a random coefficients logit model, Pierson et al. (2013) provide very simple conditions to verify uniqueness when consumers might be heterogenous in terms of their valuation of all product attributes, but price. Therefore, a researcher may choose to impose those conditions to ensure that there is a unique pricing equilibrium. In this simulation, I do not necessarily restrict the analysis to situations under which consumers are homogeneous in their price coefficients. In particular, I have performed simulation experiments with 5,760,000 instances of the random-coefficients logit model used in the previously-described Monte Carlo experiment and have not found any evidence of price equilibrium multiplicity.
profits estimated in the previous step \( \pi^*_{\text{Var}}(a_{fm}|a_{f'm},\tilde{\xi}_m^{(r)}) \), equilibrium product line introduction beliefs \( \sigma_{a_{fm}}(\tilde{\xi}_m^{(r)}) \) are estimated for each draw of \( \tilde{\xi}_m^{(r)} \) and for every possible firm assortment. In this step the remaining parameters of the profit function are estimated: the scale of the profit shocks (\( \mu \)) and the fixed components (\( \kappa \)). This is accomplished by solving the following problem:

\[
\max_{\sigma,\mu,\kappa} \prod_{m=1}^{M} \prod_{f=1}^{F} \tilde{\sigma}_{a_{fm}}(\xi_{m,I})
\]

subject to:

\[
\sigma_{a_{fm}}(\tilde{\xi}_m^{(r)}) = \frac{\exp\{\mu\left(\sum_{a_{f'm} \in A_f'} \sigma_{a_{f'm}}(\xi_{m,I}^{(r)}) \hat{\pi}_{\text{Var}}(a_{fm}|a_{f'm},\tilde{\xi}_m^{(r)}) + \sum_{j \in a_{fm}} \kappa_j\right)\} \sum_{a'_{fm} \in A_f} \exp\{\mu\left(\sum_{a_{f'm} \in A_f'} \sigma_{a_{f'm}}(\xi_{m,I}^{(r)}) \hat{\pi}_{\text{Var}}(a'_{fm}|a_{f'm},\tilde{\xi}_m^{(r)}) + \sum_{j \in a'_{fm}} \kappa_j\right)\}}{R_{\xi} \sum_{r=1}^{R_{\xi}} \sigma_{a_{fm}}(\tilde{\xi}_m^{(r)})},
\]

\( \forall f, \forall m, \forall a_{fm} \in A_f, \forall r. \)

where the last constraint ensures that the estimated product introduction beliefs satisfy the conditions of a Bayesian Nash equilibrium and \( \tilde{\sigma}_{a_{fm}}(\xi_{m,I}) \equiv \frac{1}{R_{\xi}} \sum_{r=1}^{R_{\xi}} \sigma_{a_{fm}}(\tilde{\xi}_m^{(r)}) \). Note that this second step is very similar to the estimation method in Vitorino (2012), although here a structural formulation for firm profits is used, which is derived from consumer utility maximization and Bertrand-Nash pricing assumptions.

5. **Approximate Product Line Introduction Beliefs and Estimate \( \rho \):** Using the estimated product line introduction beliefs \( \sigma_{a_{fm}}(\tilde{\xi}_m^{(r)}) \), one can construct an approximation \( g_{afm}(\tilde{\xi}_m^{(r)}; \rho) \) of these beliefs. This is accomplished by finding the value of \( \rho \) that minimizes the sum of squared residuals:

\[
\min_\rho \sum_{m=1}^{M} \sum_{r=1}^{R_{\xi}} (g_{afm}(\tilde{\xi}_m^{(r)}; \rho) - \sigma_{a_{fm}}(\tilde{\xi}_m^{(r)}))^2
\]

where \( g_{afm}(\tilde{\xi}_m^{(r)}; \rho) \) follows the functional form specified in equations (23) and (24).

6. **Assess convergence and repeat:** Assess model parameter convergence and return to step 2 if the parameters have not converged.
It is important to note that as the approximation function $g_{afm}$ approaches $\sigma_{afm}$ (for example, by using an increasingly flexible approximation), the output of the iterative problem converges to a solution to the Kuhn Tucker conditions of the original problem (19). This is because as the approximation becomes increasingly precise, the first derivatives of the objective function considered in step 2 of the iterative method converge to the true objective function (19) and the constraints of that problem are satisfied by the output of the iterative method. The first issue is easily verified, since the objective function in step 2 of the iterative method depends on $g_{afm}(\xi_m; \rho_{afm})$ and as it approaches $\sigma_{afm}(\xi_m)$ the first derivatives of this objective function converge to the ones of the objective function presented in equation (19). The second issue is also easily verified since all constraints are considered in the iterative procedure: observed market shares match expected market shares, prices are best responses and product introduction beliefs about competitors are rational. In fact, this is verified not only when $g_{afm}(\xi_m; \rho_{afm})$ exactly matches $\sigma_{afm}(\xi_m)$, but for any approximation error (i.e., regardless of the quality of the approximation).

The next subsection demonstrates the performance of this method by means of a numerical study that uses the same data described in Subsection 3.2 to facilitate the comparison with the method that ignores product line endogeneity.

3.5 Simulation Study: Performance of the iterative estimator

The iterative estimation method described in previous subsection is applied to each of the 30 data sets and using $R_\xi = 100$ draws to integrate out the distribution of the unobserved characteristics of the non-introduced products. A summary of the estimation results is presented in Table 4.

In contrast with the results in Table 2, it is observed that the iterative method produces good results for all parameters. In particular, in the case of $\theta_1$ which measures the overall demand for the purchase alternatives, notice that the mean of the estimates (3.10) is much closer to the truth than the one obtained under the method that ignores the endogeneity of product lines (3.71). Moreover, recall that the method that ignores product line endogeneity overestimated $\theta_1$ for all the data sets, while the two-step method proposed in 3.4 exhibits a more balanced performance given that the fraction of estimates above the truth is equal to 60%.

In addition, recall that the model that ignores product line endogeneity underestimates
the standard deviation of the unobserved product characteristics. This is explained by the fact that this model only considers the products that were introduced, which are more likely to be similar to each other in terms of the values of the unobservables. In contrast, the iterative method generates estimates that are on average closer to the truth (1.01 versus 0.84). Also note that as in Table 2, most of the remaining parameters do not exhibit large deviations from their true values.

The last two columns of results in Table 4 report summary statistics for the estimated values of $\kappa$ and $\mu$, which are related to the total profits of introducing a product line. These results show that the estimated values of $\kappa$ and $\mu$ are on average close to their true values. It is worth mentioning that the $g_{a fm}(\xi_{m}^{(r)}; \rho)$ function approximates equilibrium product introduction beliefs $\sigma_{a fm}(\xi_{m}^{(r)})$ very accurately, explaining on average 98.49% of the variance of the latter, with a standard deviation of this $R^2$ of only 0.41% across all data sets.

Finally, as in Subsection 3.2, a policy experiment is conducted computing equilibrium prices, market shares and profits for the case in which a product is added to set of introduced products. As before, to compute these quantities, one needs the unobserved characteristics of all products, including those that are not introduced. The unobservables for the non-introduced products, however, are not part of the output of the estimation methodology (as opposed to those of the introduced product, where a value of $\xi_{fjm}$ is estimated for each introduced alternative). Therefore, inferences about the non-introduced products require integrating over the distribution of their unobservables. This distribution is defined in equation (16). Numerically, this can be done via simulation by relying on $R_\xi$ draws of $\xi_m$ to compute each quantity of interest (e.g., an equilibrium price) and then weighting each of the $R_\xi$ values of the quantity of interest by a weight consistent with equation (16). More specifi-
cally, $R_\xi$ values of $\xi_m$ for the non-introduced products are drawn from a Normal distribution with zero mean and variance $\varphi_\xi^2$. Then this value is used along with the estimated demand and marginal cost parameters to estimate a quantity of interest. Subsequently, the $R_\xi$ values of the quantity of interest are used to compute a weighted average, with the following expression for the weight given to the value of the quantity of interest under the $r^{th}$ draw:

$$w(\tilde{\xi}_m^{(r)}) = f(\tilde{\xi}_m^{(r)} | a_{f_m}^{o}, a_{f_m}'^{o}) = \frac{\sigma_{a_{f_m}^{o}}(\tilde{\xi}_m^{(r)}) \sigma_{a_{f_m}'^{o}}(\tilde{\xi}_m^{(r)})}{\sum_{r'=1}^{R_\xi} \sigma_{a_{f_m}^{o}}(\tilde{\xi}_m^{(r')}) \sigma_{a_{f_m}'^{o}}(\tilde{\xi}_m^{(r')})}.$$  \hfill (25)

It is evident from the last expression that the weight considers how likely the observed assortment is given a value $\tilde{\xi}_m^{(r)}$ of the unobserved product characteristics: $\sigma_{a_{f_m}^{o}}(\tilde{\xi}_m^{(r)}) \sigma_{a_{f_m}'^{o}}(\tilde{\xi}_m^{(r)})$. Note that to compute these weights, one needs the product introduction equilibrium beliefs about the probability of the observed assortment. This is not a trivial task since it requires in turn knowing the Bertrand Nash equilibrium profits under all possible assortments and then finding a solution to equation (22).

This procedure was implemented using $R_\xi = 100$ and the results of this policy experiment are presented in the bottom panels of Figures 1 and 2. In addition, the last two columns in Table 3 show summary statistics for the deviations between the estimated and true values of each of the quantities (prices, market share, profits) for both new and existing products. In contrast with the results obtained when the endogeneity of product line decisions is ignored (see the top panel of both figures), it is observed that the iterative method produces estimates which are much more evenly distributed around the true values reducing the overoptimistic expectations about new products and correcting the downward bias in the equilibrium shares and profits of the extant products.

In summary, the simulation study presented in this section demonstrates the efficacy of the method at recovering the true demand and profit parameters. Moreover, it is evident that methods that ignore the endogeneity of product line decisions, as is common in the extant empirical literature, may substantially overestimate total demand. The next section complements these findings by applying this methodology to a real data set about firm and consumer decisions across several markets.
Table 5: Product characteristics and summary statistics for the products in the liquid laundry detergent category.

<table>
<thead>
<tr>
<th>Size Category</th>
<th>Vendor</th>
<th>Introd. Mean</th>
<th>Price Mean</th>
<th>Share Mean (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>P&amp;G</td>
<td>1.00</td>
<td>7.69</td>
<td>8.15</td>
</tr>
<tr>
<td>Medium</td>
<td>P&amp;G</td>
<td>0.73</td>
<td>10.54</td>
<td>0.85</td>
</tr>
<tr>
<td>Large</td>
<td>P&amp;G</td>
<td>0.90</td>
<td>16.55</td>
<td>1.55</td>
</tr>
<tr>
<td>Small</td>
<td>Unilever</td>
<td>1.00</td>
<td>6.68</td>
<td>3.65</td>
</tr>
<tr>
<td>Medium</td>
<td>Unilever</td>
<td>0.57</td>
<td>8.14</td>
<td>0.22</td>
</tr>
<tr>
<td>Large</td>
<td>Unilever</td>
<td>0.73</td>
<td>11.35</td>
<td>1.11</td>
</tr>
</tbody>
</table>

4 Data and Estimation

The proposed methodology is illustrated using a cross-sectional data set that contains information about purchases in the liquid laundry detergent product category for 30 different Zip codes in the state of Washington during the third quarter of 2000.\(^6\) To limit the computational burden, the analysis focuses on the two largest vendors, Procter & Gamble (P&G) and Unilever, and groups each manufacturer’s products into three size categories: small (less than 110 fluid ounces), medium (between 110-175 fluid ounces) and large (more than 175 fluid ounces). Market size is determined by the number of households in each Zip code. Summary statistics for each of these size categories are displayed in Table 5. In particular, note that some firm-size combinations are introduced in all markets (e.g., P&G’s small and medium sizes). This decision is primarily modeled from the perspective of the manufacturer, who chooses whether to offer these sizes in each market. It is acknowledged that there is a bargaining process between the manufacturer and the retailer than underlies these decisions, but at the same time it is reasonable to assume that the retailer should be willing to carry a product if the manufacturer is willing to offer sufficiently attractive terms (e.g., slotting fees). This implies that the estimated marginal cost also includes distribution (retailing) costs. An important improvement over this approach would be to model the vertical game between manufacturers and retailers (e.g., Berto Villas-Boas (2007)), which is a computationally non-trivial extension left for future research.

\(^6\)I thank IRI for making this data set available. The IRI Academic Dataset is described in Bronnenberg et al. (2008).
In addition, to explain variations in marginal costs across markets, cost shifters that vary across Zip codes are needed. In this application, the following variables are used as proxies for cost shifters: number of households, population density, median housing price and median rent, since these variables might be related to distribution costs (e.g., it might be more expensive to reach markets with a high population density) and the real state cost incurred by retailers (e.g., shelf space might be more expensive in markets with higher real state prices). Summary statistics for these market characteristics are presented in Table 6.

Denoting these market characteristics by $Z_m$, the marginal cost for firm $f$ and size category $j$ in market $m$ is given by $c_{fjm} = \text{vol}_{fj}(\gamma_{f0} + \sum_{k=1}^{4} \gamma_{fk}Z_m)$, where $\text{vol}_{fj}$ is the average volume of products from firm $f$ and size category $j$. This formulation is chosen to allow marginal costs to increase with the size of the product.\footnote{A constraint is added to the estimation to ensure that marginal costs are always positive, otherwise numerical issues could arise when computing equilibrium prices and profits.}

Furthermore, since markets have different sizes, we amplify variable profits (see equation 3) by the number of households in each market ($N_m$) before entering them into the Bayesian Nash equilibrium conditions (see equation 10) to account for these differences when modeling product introduction decisions (i.e., the larger the variable profits for a particular product line, the more likely that it will be chosen for a given market). Finally, the fixed component of profits is modeled as follows: $\pi_{\text{Fixed}}(a_{fm}) = (\kappa_1 - \kappa_2(|a_{fm}| - 1))N_m$, for all $a_{fm} \neq \emptyset$, where $|\cdot|$ denotes the cardinality operator for a given set. Under this formulation, $\kappa_1$ reflects the fixed baseline benefits of being present in the market with at least one product compared to the option of being absent (i.e., compared to the case of $a_{fm} = \emptyset$). Furthermore, $\kappa_2$ reflects the fixed incremental costs or savings from adding a variety to a non-empty product line in a given market. Both terms are amplified by the market size to take into account that the magnitude of these fixed profits might depend on the size of the corresponding market.

The following covariates for the consumer utility function are used: an intercept for all purchase alternatives, a size variable (6.0, 11.7 and 17.9 for the small, medium and large size categories of P&G and 6.0, 12.3 and 17.9 for the same size categories for Unilever, respectively), a dummy variable for Unilever products and the average price of each size category for each firm. The model parameters are estimated ignoring and accounting for the endogeneity of product introductions and using different $R_\theta = R_\xi = 100$ random draws for each market to approximate the distribution of random coefficients and unobserved product characteristics. Estimation results are presented in Table 7 with standard errors obtained.
Table 6: Summary Statistics of Market Characteristics.

<table>
<thead>
<tr>
<th>Product Size</th>
<th>Units</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>population</td>
<td>10,000 households</td>
<td>1.14</td>
<td>0.41</td>
<td>0.47</td>
<td>2.31</td>
</tr>
<tr>
<td>population density</td>
<td>1,000 habitants per sq mile</td>
<td>3.75</td>
<td>2.26</td>
<td>0.29</td>
<td>9.50</td>
</tr>
<tr>
<td>median housing price</td>
<td>$100,000’s</td>
<td>1.96</td>
<td>0.96</td>
<td>0.43</td>
<td>4.33</td>
</tr>
<tr>
<td>median rent</td>
<td>$1,000’s</td>
<td>0.67</td>
<td>0.19</td>
<td>0.35</td>
<td>1.06</td>
</tr>
</tbody>
</table>

using 30 balanced bootstrap samples (Chernik (2008)). A simpler version of the approximated product introduction function is used by setting $\omega$ to be zero in equation (24), since this more parsimonious version yields a very good fit ($R^2 = 95.2\%$), while it also speeds up and facilitates convergence.

From these results, it is first verified that the mean of the unobserved characteristics of the introduced products, is significantly higher when accounting for this form of endogeneity than when ignoring it (0.156 versus 0.000). The mean of these unobservables is greater than zero for all bootstrap samples in the case of the full model, while only for 40% of the bootstrap samples in the case of the model that ignores product introduction endogeneity. Moreover, for each of the introduced product, the estimated unobserved characteristics are higher when accounting for this form of endogeneity than when it is ignored (considering the uncertainty of these estimates using the bootstrap samples, on average for 95.7% of the cases the estimated unobserved product characteristics are higher when accounting as opposed to when ignoring this endogeneity). Therefore, as expected, these results provide evidence that after controlling for differences in observed characteristics, the introduced products generate on average more demand.

Given that the price-demand estimation underestimates the mean of the unobserved characteristics of the introduced products, in order to match expected and observed demand, the estimates of some of the remaining parameters of the utility function are likely to be affected. In particular, note that even though all other parameters across both estimations are very close, the intercept that affects the utility of all purchase alternatives is greater under the model that ignores product line endogeneity than under the full model (-0.675 vs. -0.942, respectively). This overestimation is verified for all bootstrap samples. Consequently, in the case of the price-demand estimation, the smaller mean of the unobserved characteristics of the
Table 7: Empirical Results: Estimated model parameters and standard errors accounting for and ignoring product line endogeneity.

<table>
<thead>
<tr>
<th></th>
<th>Full Model</th>
<th>Ign. Endog.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>S.E.</td>
</tr>
<tr>
<td>Demand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>-0.942</td>
<td>0.147</td>
</tr>
<tr>
<td>M. of random</td>
<td>-0.942</td>
<td>0.147</td>
</tr>
<tr>
<td>coefficients</td>
<td>0.033</td>
<td>0.017</td>
</tr>
<tr>
<td>Unilever price</td>
<td>-2.461</td>
<td>0.572</td>
</tr>
<tr>
<td>price</td>
<td>-0.322</td>
<td>0.017</td>
</tr>
<tr>
<td>( \Lambda )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD. of random</td>
<td></td>
<td></td>
</tr>
<tr>
<td>coefficients</td>
<td>1.644</td>
<td>0.553</td>
</tr>
<tr>
<td>Unilever price</td>
<td>0.089</td>
<td>0.006</td>
</tr>
<tr>
<td>( \varphi_\xi )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD. of unobs. charact.</td>
<td>1.285</td>
<td>0.078</td>
</tr>
<tr>
<td>( \xi_I )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. of unobs. charact.</td>
<td>0.156</td>
<td>0.017</td>
</tr>
<tr>
<td>introduced</td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-introd.</td>
<td>-0.376</td>
<td>0.065</td>
</tr>
<tr>
<td>Pricing and Marginal</td>
<td>( \gamma )</td>
<td>intercept</td>
</tr>
<tr>
<td>Costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>population</td>
<td>0.552</td>
<td>0.074</td>
</tr>
<tr>
<td>population density</td>
<td>0.034</td>
<td>0.022</td>
</tr>
<tr>
<td>median housing price</td>
<td>0.004</td>
<td>0.005</td>
</tr>
<tr>
<td>median rent</td>
<td>0.006</td>
<td>0.026</td>
</tr>
<tr>
<td>intercept</td>
<td>0.003</td>
<td>0.150</td>
</tr>
<tr>
<td>Unilever</td>
<td></td>
<td></td>
</tr>
<tr>
<td>population</td>
<td>0.346</td>
<td>0.125</td>
</tr>
<tr>
<td>population density</td>
<td>0.104</td>
<td>0.042</td>
</tr>
<tr>
<td>median housing price</td>
<td>-0.001</td>
<td>0.007</td>
</tr>
<tr>
<td>median rent</td>
<td>0.042</td>
<td>0.056</td>
</tr>
<tr>
<td>( \varphi_\eta )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD. of meas. error</td>
<td>1.309</td>
<td>0.108</td>
</tr>
<tr>
<td>Product</td>
<td>( \mu )</td>
<td>private shock scale</td>
</tr>
<tr>
<td>Introduction</td>
<td>( \kappa )</td>
<td>baseline benefit</td>
</tr>
<tr>
<td>( R^2_\sigma )</td>
<td>belief approx. fit</td>
<td>0.952</td>
</tr>
</tbody>
</table>
introduced products is being compensated by a higher estimate of this intercept. As described in the simulation study, these biases have important implications for policy recommendations given that they may lead to overoptimistic predictions about the performance of products that have not been introduced in a given market. For those products, a more positive utility intercept would be used (i.e., \(-0.675\) instead \(-0.942\), respectively) and the mean of these unobservables would be assumed to be zero even though a negative value is more likely. In fact, based on the estimated parameters and using equation (25), this mean is estimated at \(-0.376\). Both aspects (higher utility intercept and higher mean of unobserved characteristics of non-introduced products) would lead a researcher to derive over-optimistic estimates about the demand and willingness to pay for non-introduced products, as also shown in the Monte Carlo study described in the previous sections. For example, considering the average consumer in this market, this difference in utility expressed in terms of its corresponding monetary value is equivalent to: \((-0.675 - (-0.942) + 0.376)/0.089 = \$7.22\), i.e. a substantial overestimation of the willingness to pay of the average consumer for one unit of any of the non-introduced products. In terms of the results for the fixed profit parameters, both \(\kappa_1\) and \(\kappa_2\) are significant.\(^8\) These estimates reveal a preference from firms to have a presence in the market and at the same time imply (fixed) profit gains from adding varieties to a non-empty assortment.

Finally, there are important similarities across the two estimation approaches, we see that under both of them it is concluded that consumers prefer products from P&G (given the negative mean brand intercept for Unilever) although there is substantial heterogeneity in brand preferences, as evidenced by the standard deviation of the Unilever utility coefficients. Regarding marginal costs, differences across markets are significantly explained by variation in market size (i.e., marginal costs, which include also distribution costs, are higher in more populated markets). Furthermore, the marginal cost function achieves a good fit with equilibrium prices being reasonably close to observed prices (i.e., the standard deviation of the price measurement error is estimated at \$1.309 for the full model).

\(^8\)Even though the standard error for \(\kappa_2\) is relatively large compared to its estimate, the latter is significant since a positive estimate is obtained across all bootstrap samples.
5 Conclusions and Future Research

This paper proposes a structural model and an estimation method to account for the endogeneity of the mix of products made available by firms across different markets. Furthermore, it is shown that ignoring this form of endogeneity leads to substantial biases in terms of parameter estimation and also about demand and profit projections which are often important to managers.

Several extensions are possible, such as the consideration of different assumptions about the information available to the firms. For example, instead of relying on a Bayesian-Nash equilibrium framework, this could be replaced by complete information assumptions. This would be potentially interesting in industries with a small number of players where it is likely for one firm to have a reasonable knowledge about the cost structure and technology used by its rivals. Other extensions include allowing for economies of density and scope. For example, the introduction of a sufficiently large number of products in multiple markets may lead to manufacturing or distribution efficiencies (Holmes (2011); Gimeno and Woo (1999)). In addition, it might be interesting yet potentially challenging to consider strategic coordination issues. In particular, the multimarket contact literature suggests that having a presence in multiple markets can provide stronger incentives to sustain tacit collusion given that a firm’s deviation from a collusive agreement can be punished by its competitors in multiple markets (Bernheim and Whinston (1990)). Another interesting extension corresponds to the consideration of the strategic interactions between manufacturers and retailers (e.g., Berto Villas-Boas (2007)). Finally, product line decisions could be also be modeled using a dynamic setup (Aguirregabiria et al. (2007)). This could be relevant in cases where the costs or benefits of making a product available in a given period depend on previous actions (e.g., due to the existence of slotting fees paid only the first time a product is introduced or due to state dependence in consumer preferences).

In sum, I hope the model and method presented in this paper might enable researchers to account for the endogeneity of the mix of products made available by firms across different markets. Furthermore, I hope that this work may stimulate further research aimed at generalizing and extending the model and methodology introduced in this paper along the dimensions highlighted in this concluding section.
References

*Econometrica* **75**(1) 1–53.

Aguirregabiria, Victor, Pedro Mira, Hernan Roman. 2007. An estimable dynamic model 
449–454.

Aguirregabiria, Victor, Junichi Suzuki. 2015. Empirical games of market entry and spatial 
competition in retail industries. Working Paper, University of Toronto.


Berto Villas-Boas, Sofia. 2007. Vertical relationships between manufacturers and retailers: 

Besbes, O., D. Saure. 2014. Product assortment and price competition under multinomial 

Science* **27**(4) 745–748.

ed. Wiley Inter-Science.

Draganska, Michaela, Michael Mazzeo, Kajta Seim. 2009. Beyond plain vanilla: Modeling 
joint pricing and product assortment choices. *Quantitative Marketing and Economics* **7**(2) 
105–146.


