Structural Model of Product Line Pricing in Differentiated Good Markets

Isis Durrmeyer, Andras Niedermayer, Artyom Shneyerov*

May 15, 2017
Preliminary and Incomplete

Abstract

We develop a new empirical model of market equilibrium with horizontal and vertical product differentiation. Our model is an extension of Rochet and Stole (2002)'s non-linear pricing theory to multiproduct firms. In this model, product line pricing strategies follow second-degree price discrimination. We propose a new method to identify demand and supply primitives. This methodology leverages the rich theoretical framework and provides identification of the parameters of interest under very mild assumptions on functional forms. We apply our new methodology to the French automobile market and measure the importance of second-degree price discrimination.

*We thank participants of the MaCCI IO Day 2015 in Mannheim, IIOC 2016 in Philadelphia, MaCCI Summer Institute in Competiton Policy 2016 at Schloss Burgellern, SED 2016 in Toulouse, EEA 2016 in Geneva, the CSIO/IDEI conference in Evanston and participants of seminars at the University of Mannheim, the University of Melbourne, the Bank of Canada and the Toulouse School of Economics for helpful comments. [further acknowledgments to be added]

1 Introduction

In many oligopolistic markets products are both horizontally and vertically differentiated. Modeling and understanding product line pricing in such industries presents several challenges. First of all, the pricing of different levels of quality of the same product is likely to reflect price discrimination by producers to take advantage of heterogeneous tastes for quality in the consumer population. As an example, more than 200 versions of the Renault Mégane were offered in France in 2007, with the price ranging from EUR 16,000 for the version with 72 horsepower to over EUR 31,000 for the version with 127 horsepower. One other example is the telecommunications industry, with prices paid for mobile communication strongly varying depending on the number of minutes talked, data downloaded, data speed, and the number of text messages sent.

Another challenge for the analysis pricing strategies in such industries is that firms very often offer multiple products. Renault does not only produce the Mégane, but also the Clio and several other car models: ten models under the brand *Renault* and one under the brand *Dacia*. Renault is no exception: over 267 car models sold in France are produced by 22 manufacturers.

This also holds for other industries, such as the telecommunications industry: e.g. Telefónica Germany is behind the brands o2, E-Plus, BASE, simyo, and Blau.

We develop a new empirical model of market equilibrium with horizontal and vertical product differentiation. Our model allows for second-degree price discrimination in an oligopolistic market. The model is based on Rochet and Stole (2002) and extends their approach to multiproduct firms. On the demand side, we assume that consumers are heterogeneous in terms of preference for quality (hereafter *type*) and preference over products. Consumers choose which product to buy and the quality level of the good. On the supply side, we assume that firms can offer their products at any level of quality. Firms play a game where they simultaneously set the optimal pricing schedules (i.e. the price for any quality level) for all the products.

We propose a new methodology to estimate the primitives of the model, i.e. the parameters of preferences of consumers and the marginal costs of firms. The identification takes advantage of the rich theoretical model and relies on the assumption that product line pricing reflects an optimal screening strategy of firms. We show that we can non-parametrically identify the distribution of consumer types under standard functional form assumptions on the utility function. By leveraging the theory, identification does not rely on the exogeneity of product characteristics as typically done in the literature (see Berry et al. (1995)). We propose a five-step estimation algorithm that uses data on prices and sales at the product-version level.

We apply our methodology to the new automobile market in France. We take advantage of observing prices, characteristics, and market shares not just at the car model level (such as the Renault Mégane), but also the version level (such as the version with 72 horsepower of the Renault Mégane). Price dispersion across versions of the same model is rather important in this market as the price difference between the cheapest and the most expensive version of a car model is EUR 16,000 on average and can be as high as EUR 60,000.

We estimate the primitives of demand and supply and quantitatively investigate the importance of price discrimination in the market. Preliminary results suggest that price discrimination plays an important role in the French automobile industry: the profits related to price discrimination are between one third and two thirds of the total profits, depending on manufacturers.

This paper relates to three strands of literature: the theory of competitive non-linear pricing (Armstrong and Vickers (2001), Rochet and Stole (2002) and Jullien (2000)), the empirical literature on non-linear pricing, and the empirical literature on oligopolistic pricing in differentiated products markets.

To the best of our knowledge, this is the first paper to bring a competitive non-linear pricing theory a la Armstrong and Vickers (2001) and Rochet and Stole (2002) to the data. These two papers are mostly cited for the result that price discrimination disappears if there is enough competition, and quality increments above the baseline version are priced at marginal costs under certain conditions.[1] However, under different conditions there can be price dis-

[1]These conditions are that the firms are symmetric (i.e. they have the same costs to produce a certain

crimination and quality increments are priced above marginal costs. These two papers do not obtain closed form solutions to characterize the solution to firms' profit maximization problem under these conditions. It is perhaps this lack of a closed form solutions which has lead to more emphasis on the case with closed form solutions; to the extent that sometimes these articles are informally referred to a "showing that price discrimination does not matter if there is enough competition".[2] Bringing the theory to the data allows us to solve the model numerically even absent closed form solutions. Our estimates suggest that – at least for the French automobile industry – the case in which there is price discrimination is the relevant one.

An auxiliary theoretical contribution of our article is that we extend the Rochet and Stole (2002) model to a multi-product firm setting.

The empirical literature on non-linear pricing has almost exclusively focused on monopolists or local monopolists (Byrne (2015), Luo, Perrigne, and Vuong (2014), Crawford, Shcherbakov, and Shum (2011)). While the assumption of a local monopoly is reasonable for the setups considered in these papers, it would not be appropriate for the French automobile industry with 22 manufacturers and a large number of car models. Notable exceptions are Ivaldi and Martimort (1994), Miravete and Röller (2003), and Aryal (2013) who consider a duopoly with non-linear pricing. Our main difference to these articles is that they consider markets in which consumers can combine quantities purchases from different suppliers (which corresponds to screening with non-exclusive contracts). While this assumption is appropriate for the markets considered in these articles, it would not fit the automobile industry in which the baseline version of a car and the quality increments have to be bought from the same manufacturer.[3] Another difference is that these articles restrict firms' endogenous price schedules to be quadratic, whereas we estimate price schedules non-parametrically.[4] A recent

level of quality), that the market is sufficiently competitive (i.e. the products are not too differentiated), and there is no outside good.

[2]It should be emphasized that Armstrong and Vickers (2001) and Rochet and Stole (2002) make it clear that competition does not necessarily lead to price discrimination disappearing, but only under the conditions stated in their papers.

[3]For example, you cannot buy the baseline version with 70 horsepower of the Mégane from Renault and the quality increment to 100 horsepower from Ford.

[4]This is also important for our application: one would not necessarily believe that the price of a car is a

paper by Fan and Yang (2016) estimates a model of demand and supply for smartphones in which firms choose both prices and quality. Our paper differs as we consider a continuous level of quality are offered while in their model, there is a fixed cost of offering a different level quality.

We extend the differentiated products empirical literature in the tradition of Berry (1994) and Berry, Levinsohn, and Pakes (1995) (BLP) to allow for quality differentiation in addition to the horizontal product differentiation. Very ofter, the literature has assumed that there is only one version of a product being offered or aggregated the different version of a product to one version and in some papers to a restricted number versions (see e.g. Verboven (2002) who considers two versions). Our novelty lies in making use of the information on prices and sales of the versions of a product rather than aggregating. On the supply-side, we depart from the literature since we consider firms competing with price schedules a la Rochet and Stole (2002) rather than with single prices per product (or two prices for two versions of a product). Furthermore, we estimate demand non-parametrically and without relying on exogenous characteristics assumptions and do not use instrumental variables.

In a wider sense, this article also relates to the recent literature on the identification of monopolistic multidimensional screening (see Aryal (2016) and Aryal, Perrigne, and Vuong (2016)).

The paper is structured as follows. Section 2 describes the theoretical model of product line pricing with multiproduct firms. Section 3 derives the empirical model, the identification strategy and the estimation method. In Section 4 we apply our methodology to the French automobile industry.

2 Theory

We consider an industry with M firms indexed by $m = 1, ..., M$. Each firm produces J_m different products, indexed by $j_m = 1, ..., J_m$. A product j is offered at any quality level q_j over the interval $[\underline{q}_j, \overline{q}_j] \subset \mathbb{R}+$. In this settings firms compete in price schedules $p_j(q)$ as

quadratic function of horsepower.

opposed to the standard setting in which firms set one price for each product.

Following the non-linear pricing literature q_j can either be the quality the quantity, the analysis remains the same. To simplify the exposition, we will always refer to q_j as quality level in the following, even though it can also be interpreted as quantity.

2.1 Demand

We consider N consumers that choose one product among the set $J = \sum_m J_m$ available or decide not to buy the product. If they decide to buy the product, they also choose the quality level q_j of this product. We follow the model developed by Rochet and Stole (2002) and the standard literature on demand for differentiated products (e.g. Berry et al. (1995)) and assume that a consumer i's utility from product j at quality level q_j is

$$U_{ij}(q_j) = t_i q_j - p_j(q_j) - x_{ij}$$

where q_j is the quality chosen by the consumer, t_i is the willingness to pay for quality of consumer i (or its type) and is drawn from the distribution F. $p_j(\cdot)$ is the price schedule of the firm, and x_{ij} is the outside opportunity cost. The outside opportunity cost is the utility the consumer would derive from a competing product or from not buying any product. The utility from choosing not to buy the product is denoted by u_0.

A consumer's optimal choice of quality conditional on choosing product j is

$$q_j^*(t_i) = \arg\max_q t_i q - p_j(q)$$

The function $q_j^*(t_i)$ is implicitly given by the first-order condition[5]

$$t_i - p_j'(q_j^*(t_i)) = 0$$

The interpretation is standard, the optimal quality level chosen is such that the marginal utility of quality is equal to the marginal price increase. A useful concept for the further analysis is the indirect utility a consumer of type t_i gets from choosing product j:

$$u_j(t_i) = t_i q_j^*(t_i) - p_j(q_j^*(t_i))$$

[5]We assume that there is no corner solution, i.e. all the quality levels are offered by firms.

A consumer chooses product j if

$$u_j(t_i) - x_{ij} > \max\{u_0, \max_{j'} u_{j'}(t_i) - x_{ij'}\} \tag{1}$$

It is useful to define the vector of indirect utility functions $\boldsymbol{u}(t) = (u_1(t), u_2(t), ..., u_J(t))$.

It will also be helpful for the further analysis to see the equivalence of the above approach with consumers' types, which is used in the theoretical literature, and the standard approach in the empirical literature which relies on consumers' price sensitivities. The empirical literature typically considers that consumers have heterogeneous price sensitivities α_i and an i.i.d. preference shock for each product ϵ_{ij}. For this notation, the consumer's utility is written as

$$\hat{U}_{ij}(q_j) = q_j - \alpha_i p_j(q_j) + \epsilon_{ij} \tag{2}$$

With this notation, the quality chosen conditional on choosing product j is given by the first-order condition

$$1 - \alpha_i p_j'(q_j^*(\alpha_i)) = 0$$

and the corresponding indirect utility $\hat{u}_j(\alpha_i) = q_j^*(\alpha_i) - p_j(q_j^*(\alpha_i))$. A consumer chooses product j if

$$\hat{u}_j(\alpha_i) + \epsilon_{ij} > \max\{\hat{u}_0 + \epsilon_{i0}, \max_{j'} \hat{u}_{j'}(\alpha_i) + \epsilon_{ij'}\} \tag{3}$$

Comparing (1) and (3) shows that the two notations are equivalent with $t_i = 1/\alpha_i$, $u_j(t_i) = \hat{u}_j(\alpha_i)/\alpha_i$, $u_0 = (\hat{u}_0 + \epsilon_{i0})/\alpha_i$ and

$$x_{ij} = \frac{\max\{\hat{u}_0 + \epsilon_{i0}, \max_{j'} \hat{u}_{j'}(\alpha_i) + \epsilon_{ij'}\} - \epsilon_{ij}}{\alpha_i}$$

It is interesting to note that for a single product firm, the outside opportunity cost x_{ij} is exogenous from firm j's perspective. On the other hand, a multiproduct firm takes into account that offering a different price schedule for a product j' also changes the outside option x_{ij} for its product j.

2.2 Supply

We develop a model of competition with second degree price discrimination. Firms compete in price schedules as opposed to the standard setting where firms set only one price. Our model extends Rochet and Stole (2002) to a setting with multiproduct firms. A firm strategically sets price schedules $p_j(q)$ for all the products it offers. The program of firm m that sells the set of products J_m is

$$\max_{\{p_j\}_{j\in J_m}} \sum_{j\in J_m} \pi_j(p)$$

where π_j is the profit associated with product j and is

$$\pi_j(p) = \int_{\underline{t}}^{\bar{t}} M_j(\boldsymbol{u}(t), t)[p_j(q_j^*(t)) - C_j(q_j^*(t))]dF(t)$$

where $C_j(q)$ is the cost of producing quality q and $M_j(\boldsymbol{u}(t), t)$ is the total demand for product j:

$$M_j(\boldsymbol{u}, t) = \text{Prob}((x, t)|x \leq u_j)$$

Let G_j be the probability that a consumer chooses product j given utility vector $\boldsymbol{u}$, that is $G_j(\boldsymbol{u}(t)/t) = \text{Prob}\left(u_j(t_i) - x_{ij} > \max\{u_0, \max_{j'} u_{j'}(t_i) - x_{ij'}\}\right)$ Then, the total demand is

$$M_j(\boldsymbol{u}(t), t) = \int G_j(\boldsymbol{u}(t)/t)dF(t)$$

The usual approach to simplify this optimal control problem is to solve of $u_j(t)$ rather than $p_j(q_j)$. As a first step, consider that by a standard envelope theorem argument, the indirect utility function has to satisfy $u_j'(t) = q_j(t)$.[6] Also, observe that $u_j(t) = tq_j(t) - p_j(q_j(t))$, so that given u_j and q_j, the price schedule can be obtained by $p_j(q_j(t)) = tq_j(t) - u_j(t)$. So a firm can use u_j and q_j as control functions to maximize profits and then simply recover the price schedule from this. A firm's maximization problem can be rewritten as

$$\begin{aligned} \max_{\{u_j, q_j\}_{j\in J_m}} \quad & \sum_{j\in J_m} \int_{\underline{t}}^{\bar{t}} M_j(\boldsymbol{u}(t), t)[tq_j(t) - u_j(t) - C_j(q_j(t))]dF(t) \\ \text{s.t.} \quad & u_j' = q_j \text{ for all } j \in J_m \end{aligned} \tag{4}$$

The optimal utilities solution to this problem can be derived using optimal control. The following Proposition characterizes the solution.

[6]This stems from consumers' first-order condition $t - p_j'(q_j^*(t)) = 0$, which implies $u_j'(t) = q_j^*(t) + t - p_j'(q_j^*(t)) = q_j^*(t)$.

Proposition 1. *A fully separating solution to the optimal control problem* (4) *satisfies the following necessary first-order condition (the* Euler *equation):*

$$\left[\sum_{j'\in J_m} \frac{\partial G_{j'}(\boldsymbol{u}(t)/t)}{\partial u_j}(tq_{j'}(t) - C_{j'}(q_{j'}(t)) - u_{j'}(t)) - G_j(\boldsymbol{u}(t)/t)\right] f(t)$$

$$= \frac{d}{dt}\left[G_j(\boldsymbol{u}(t)/t)f(t)\left(t - C_j'(q_j(t))\right)\right] \tag{5}$$

and the boundary conditions

$$C_j'(q_j(\bar{t})) = \bar{t}, \quad C_j'(q_j(\underline{t})) = \underline{t}.$$

The proof follows the same logic as Rochet and Stole (2002) with the difference of taking into account multiproduct firms. For single product firms, this solution specializes to the solution provided in their paper. Note that, as in Rochet and Stole (2002), there is no distortion both at the top ($\bar{t}$) and at the bottom ($\underline{t}$).

A much celebrated additional result in Rochet and Stole (2002) is the *no-quality distortion* result (see also Armstrong and Vickers (2001) for a similar result): under some conditions, the solution of the single product firm problem becomes very simple: the price is simply equal to the cost of producing a quality level q plus a fixed markup, formally $p_j(q) = C_j(q) + \kappa$, where the markup κ is determined by the competitiveness of the market. The condition for this result to hold is that (i) there is no outside good, (ii) the firms are symmetric (i.e. $c_j(q) = c_{j'}(q)$ for all j, j'), and (iii) the market is sufficiently competitive (the last condition is stated more formally in Rochet and Stole (2002)).

This result means that under these conditions, firms effectively choose not to price discriminate based on the chosen quality, so that consumers buy the quality level they receive under a first-best allocation. It is an empirical question to investigate by how much the provision of quality is distorted in a given market with differentiated products. In other words, we can measure to what extent price differences between low-end and high-end versions of products are explained by cost differences versus markup differences.

We propose a methodology to estimate demand parameters and the cost function which allows one to recover the markups for different quality levels.

2.3 Discussion

Before moving on to the empirical model, it is worthwhile to discuss the similarities and differences of our theoretical model to existing models. The similarity is the largest to Rochet and Stole (2002), since our model is an extension of their model to multiproduct firms. The Euler equation in Rochet and Stole (2002) is

$$\left[\frac{\partial G_j(\boldsymbol{u}(t)/t)}{\partial u_j}(tq_j(t) - C_j(q_j(t)) - u_{j'}(t)) - G_j(\boldsymbol{u}(t)/t)\right] f(t)$$
$$= \frac{d}{dt}\left[G_j(\boldsymbol{u}(t)/t)f(t)\left(t - C_j'(q_j(t))\right)\right].$$

The key difference is that (5) contains the sum over the products sold by the firm, which accounts for cannibalization, i.e., a multiproduct firm is concerned that lowering the price of one product not only increases demand for that product, but also decreases demand for other products sold by the same firm.

Compared to the standard Mussa and Rosen (1978) monopolistic non-linear pricing, one of the key differences is that the outside opportunity cost x_{ij} has a non-degenerate distribution in our setting. Our models reduces to the standard Mussa and Rosen (1978) when there is only one firm offering a single product and the outside opportunity cost x_{ij} is constant (which can be normalized to zero), so that a consumer's utility becomes $U_i(q) = t_i q - p(q)$. The consumer's first-order condition implies the same choice of quality $q^*(t)$ based on his type as in our model, given implicitly by $t - p'(q^*(t)) = 0$. The monopolist problem then becomes the maximization $\pi(p) = \int_{t^*(p)}^{\bar{t}} [p(q^*(t)) - C(q^*(t))]\, dF(t)$, where $t^*(p)$ is the lowest consumer type buying the good. Consumers with $t < t^*(p)$ are excluded. In our model, no consumer is excluded solely based on his type t: if the outside opportunity cost x_{ij} is sufficiently low (which is always possible if x_{ij} has unbounded support), then a consumer chooses to buy the product. The reverse also holds in our setting: in contrast to Mussa and Rosen (1978) no type t buys with probability 1, since there is always a possibility of a very high outside opportunity cost x_{ij}.

When comparing our model to the discrete choice model underlying the BLP approach, one has to consider two interpretations of how BLP could be applied to product line pricing in a differentiated product setting, which we refer informally to *product BLP* and *version-product BLP*.

In *product BLP*, different versions of a product are aggregated to one single product. For example, the Mégane version with 72 horsepower and the Mǵane version with 127 horsepower are considered to be the same product. For the estimation, one simply uses the sum of sales of all the different versions of the product. Regarding the choice of which price and characteristics to consider, different methods of aggregation have been used in the literature. Berry et al. (1995) uses the prices and characteristics of the baseline version of each car model. Another possibility is to take the version located at a specific quantile of the distribution of versions or the mode (e.g. Nurski and Verboven (2016) define the baseline version as the the 25% percentile version, D'Haultfoeuille et al. (2014) use the prices and characteristics of the most frequently purchased version).

If the no-quality distortion result in Rochet and Stole (2002) holds, then this aggregation does not constitute a problem and the aggregated model implies the correct welfare predictions, provided that one uses the baseline version for the aggregated product. The reason is that under this specific case of product line pricing, firms choose optimally the markup for their baseline version while the quality increments do not directly enter the profit function since they are sold at marginal cost. The optimal markups then satisfy the same first order conditions as in the standard BLP model. Quality increments of course affect the demand for the product as the demand for the product is higher when the quality increments are attractive. But this effect is captured by the product specific unobserved characteristics in the standard BLP model. The welfare implications are also the same in the BLP framework and under this particular case of the Rochet and Stole (2002) model. Indeed, second degree price discrimination disappears and the markups are constant across different versions of the same model. Therefore, the *product BLP* model can be considered to be the special case of our model in which the no-quality distortion result holds.

The other application of the standard BLP model to our setting is what we call *version-product BLP*: each version is considered to be a different product (see e.g. Grigolon et al. (2014)). For this, consider the continuous quality increments $q_j \in \mathbb{R}$ of a product j to be a continuous approximation of a discrete distribution of quality increments q_{jv} with $v = 1, ..., V_j$, where V_j is the number of versions of product j offered. In this setting, a consumer's utility of choosing product (j, v) is

$$\hat{U}_{ijv} = q_{jv} - \alpha_i p_{jv} + \epsilon_{ijv}.$$

The utility function we specify in our model in equation (2) can be viewed a restriction of the *product-version BLP* that imposes $\epsilon_{ijv} = \epsilon_{ij}$, i.e., the horizontal differentiation shocks are perfectly correlated for all versions v of a model j.

If the true model is the *version-product BLP*, aggregating the versions of a product leads to an incorrect specification of the model and has different welfare implications, especially since the number of products is reduced.

Our model nests the *product-BLP* model, but is more restrictive than the *version-product BLP* approach. However, the latter poses serious challenges for estimation as we discuss in the next section.

3 Empirical model and identification

We want to identify primitives of the theoretical model from aggregate level data on equilibrium prices and sales at the version level. We use the specification in which consumers are heterogeneous in terms of price sensitivity to be close to the standard literature on demand for differentiated products. We show that we can non-parametrically identify the distribution of types $F(\alpha)$ and identify very flexible cost functions $C_j(q)$. As D'Haultfœuille and Février (2015) and Luo et al. (2014) point out, it is impossible to non-parametrically identify simultaneously the distribution of types, the cost function and the utility function without an exogenous source of variation.As opposed to Luo et al. (2014) who impose a simple parametric form on the cost function, we specify a very flexible form for the cost functions and make a simpler parametric assumption on the utility function and consider it is linear, as in the theoretical model.

3.1 Assumptions

We present the main distributional assumptions we need to identify the primitives of the model. These assumptions are standard in the differentiated products demand estimation literature.

Assumption 1. *Distribution of the outside option*

ϵ_{ij} *are i.i.d. over products and consumers and do not depend on the level of quality chosen.*

As usually done in the differentiated products demand estimation literature (see Berry (1994)) we assume that ϵ_{ij} are distributed according to a Type 1 extreme value distribution.

Assumption 2. *Normalization of the outside good utility*

$$u_0(\alpha) = 0, \text{ for all } \alpha$$

This normalization is necessary as only the differences in utilities matter for the choice of products. This is also a standard assumption in the literature on demand for differentiated products.

Given these assumptions, the probability that a consumer of type α_i chooses product j is:

$$G_j(\alpha) = \frac{\exp(u_j(\alpha))}{1 + \sum_{j'} \exp(u_{j'}(\alpha))} \tag{6}$$

The market share of product j is:

$$s_j = \int \frac{\exp(q_j(\alpha) - \alpha p_j(q_j(\alpha)))}{1 + \sum_{j'} \exp(q_{j'}(\alpha) - \alpha p_{j'}(q_{j'}(\alpha)))} dF(\alpha)$$

where $F(\alpha)$ is the cumulative distribution function of consumer price sensitivities.

It is sometimes impossible to observe a reliable quality measure (see Luo et al. (2014)).[7] Instead of relying on a unique variable to represent quality, we use several characteristics of the versions to construct a quality index. For our theoretical analysis we used a continuous quality level q_j. However, empirically, we observe a discrete empirical distribution of qualities,

[7] Note that when we consider quantity instead of quality the approach simplifies, since one does not have to construct a quality index.

which we will denote as q_{jv}, where v is the index of a specific version of a product. Denote the price of version v of product j as p_{jv}.

Assumption 3. *Separability of quality*

$$q_{jv} = X_{jv}\beta_j + \xi_j + \eta_{jv}$$

where X_{jv} are observable characteristics of version v of product j, ξ_j and η_{jv} are the unobserved characteristics.

ξ_j represents the unobserved component of the quality that is common across versions of the same product. η_{jv} are assumed to be i.i.d. over products and versions and have a zero expected value over j so that ξ_j captures the average per product unobserved quality. The parameter β_j represents the contribution of each characteristic to the quality and is allowed to vary across products. For instance, fuel efficiency is allowed to be more important for quality of the Renault Mégane than it is for a Mercedes Class C.

3.2 Identification

We face the standard challenge for identification that the price p_{jv} of a version of a model is affected by six factors that are non-trivial to disentangle: (i) the observable characteristics of the version $X_{jv}\beta_j$, (ii) the unobservable characteristics η_{jv} of the version, (iii) the distribution f of α which determines the elasticity, (iv) a scale parameter that scales β_j, (v) unobserved characteristics ξ_j of the model, (vi) costs $C(q_{jv})$ of producing the good.

We will deal with this by considering variations in the data which shut down some of the factors but not others.

Price schedule and quality

We first need to construct the quality indexes and the price schedules from the observed prices and characteristics of the versions of the different products. While the price schedule $p_j(\cdot)$ is an endogenous object, our theoretical model implies that $p_j(\cdot)$ is a twice differentiable, monotone, convex function. We specifically make use of the monotonicity of $p_j(\cdot)$ in the following:

$$p_{jv} = p_j(q_{jv})$$

Taking the inverse of the price function and combining with Assumption 3, we get:

$$\begin{array}{l} q_{jv} = p_j^{-1}(p_{jv}), \text{ or,} \\ X_{jv}\beta_j + \xi_j + \eta_{jv} = p_j^{-1}(p_{jv}) \end{array}$$

β_j and p_j^{-1} are identified from within-product variations of characteristics X_{jv} and prices p_{jv}. Note that an affine transformation of p_j^{-1} is observationally equivalent to scaling β_j by a scalar and changing ξ_j. In other words,at this stage we cannot identify ξ_j and can only identify β_j up to a scale parameter, i.e. we can identify $\tilde{\beta}_j = \beta_j/\lambda_j$, where λ_j is a scale parameter, and we can identify p_j^{-1} up to an affine transformation, i.e. we can identify

$$\tilde{p_j}^{-1}(p_{jv}) \quad = \tilde{q}_{jv} = \tfrac{q_{jv}-\xi_j}{\lambda_j}$$

We now turn to the identification of the distribution of types and postpone the discussion on the identification of the unobserved quality and the scale parameters as it uses the conditional distribution of types.

Identification of the conditional distribution of types

Next, we turn to within model variations of market shares. Within model variations have the advantage that ξ_j is constant within a model. The scale parameter λ_j does not matter either, as long as we are (temporarily) content with identifying the distribution of the scaled parameter $\tilde{\alpha}_j$ rather than the distribution of the true α. We will return to the latter later.

It is straightforward to back out the density of types, conditional on the product choice if one observes the price schedules and the quality levels chosen by consumers. Indeed, conditional on buying product j, the optimal choice of quality for a consumer of type α is given by the first-order condition:

$$p_j'(q_j) = \frac{1}{\alpha}$$

Observing the chosen quality q_{jv}, we know that a consumer's type is

$$\alpha(jv) = \frac{1}{p_j'(q_{jv})}, \text{ or equivalently,}$$
$$\alpha(jv) = \left(p_j^{-1}\right)'(p_{jv})$$

As mentioned above, the quality is identified from the prices and characteristics up to an affine transformation. So we can only recover $\tilde{p}_j^{-1}(p_{jv})$. Given that

$$p_j^{-1}(p_{jv}) = \lambda_j \tilde{p}_j^{-1} + \xi_j$$

we can recover the distribution of $\tilde{\alpha}_j = \alpha/\lambda_j$ using

$$\tilde{\alpha}_{jv} = \left(\tilde{p}_j^{-1}\right)'(p_{jv})$$

Using the relative market shares s_{jv}/s_j as weight, we recover the conditional density $\tilde{f}_j(\tilde{\alpha})$ (i.e. the density of $\tilde{\alpha}_j$ conditional on choosing j). We can use a kernel density estimator to estimate f_j:

$$\hat{f}_j(\tilde{\alpha}_j) = \sum_{v=1}^{V_j} \frac{1}{h} K\left(\frac{\tilde{\alpha}_j(jv) - \tilde{\alpha}_j}{h}\right) \frac{s_{jv}}{s_j}$$

where h is the bandwidth and $K(.)$ is a Kernel, e.g. the normal kernel:

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-u^2}{2}\right).$$

Identification of the scale parameters

To identify the scale parameters λ_j, we compare the effect of within model price variations on market shares for different versions. (This similar to a standard diff-in-diff approach, except that we need our structural model to deal with endogeneity.) The basic idea is best glanced by looking at an extreme case: assume that λ_j were close to zero, so that consumers would hardly care at all about quality increments. In this extreme case, market share differences within a model would be almost purely driven by price differences across versions. If, however, λ_j were large, market shares would be less affected by price differences and rather by quality differences.

To be more specific, we use the following restrictions implied by the structural model. According to Bayes' rule for the joint probability of observing the type α and the product j

$$s_j f_j(\alpha) = G_j(\alpha) f(\alpha) \tag{7}$$

where $G_j(\alpha)$ is the probability of choosing model j, conditional on the type α presented in equation (6)). Since equation (7) holds for all products and all consumer types, the ratio of this equation for two different products is

$$\frac{s_j f_j(\alpha)}{s_{j'} f_{j'}(\alpha)} = \frac{\exp(\tilde{u}_j(\alpha) + \xi_j)}{\exp(\tilde{u}_{j'}(\alpha) + \xi_{j'})},$$

where $\tilde{u}_j(\alpha) = u_j(\alpha) - \xi_j$.

To get rid of the unobserved product quality ξ that are, at this stage unknown, we take the derivative of the logarithm of the previous equation with respect to α:

$$\frac{d \log f_j(\alpha)}{d\alpha} - \frac{d \log f_{j'}(\alpha)}{d\alpha} = \frac{d\tilde{u}_j(\alpha)}{d\alpha} - \frac{d\tilde{u}_{j'}(\alpha)}{d\alpha}$$

Rewriting this in terms of $\tilde{\alpha}_j$ and using the equality $u'_j = -p_j$ stemming from an envelope theorem argument, we get

$$\frac{\lambda_j \tilde{f}'_j(\lambda_j \tilde{\alpha}_j)}{\tilde{f}_j(\lambda_j \tilde{\alpha}_j)} - \frac{\lambda_{j'} \tilde{f}'_{j'}(\lambda_{j'} \tilde{\alpha}_{j'})}{\tilde{f}_{j'}(\lambda_{j'} \tilde{\alpha}_{j'})} = -\tilde{p}_j(\tilde{q}_j(\lambda_j \tilde{\alpha}_j)) + \tilde{p}_{j'}(\tilde{q}_{j'}(\lambda_{j'} \tilde{\alpha}_{j'})) \tag{8}$$

In the current version of this paper, we make the additional identifying assumption that for all car models in the same category (such as small vehicles) the values of β_j and hence also λ_j are the same. (Note that it is usual to assume that β_j is the same for all car models j, so this is still a relatively weak assumption.) Using this assumption and taking (8) for a j and j' in the same category, we get an equation for λ_j that allows identification of λ_j.

We hope that in a future version of this paper we can show identification without this assumption. This could possibly be achieved by taking (8) for different values of α.

Identification of the unconditional density of types

With λ_j at hand we can recover the conditional distribution of α for a model j from the conditional distribution $\tilde{f}_j$ of $\tilde{\alpha}_j$:

$$f_j(\alpha) = \lambda_j \tilde{f}_j(\tilde{\alpha})$$

The unconditional density is related to the conditional density by Bayes' rule:

$$f(\alpha) = \frac{s_j f_j(\alpha)}{G_j(u(\alpha))}.$$

This equality has to hold for all products j. For an estimator, it is better to take a weighted average of the above equation. Take for example the following estimator for f:

$$\hat{f}(\alpha) = \sum_{j=1}^{J} \frac{G_j(u(\alpha))}{1 - G_0(u(\alpha))} \left[\frac{s_j f_j(\alpha)}{G_j(u(\alpha))} \right], \tag{9}$$

where the expression in square brackets is equal to f for all j and $\sum_{j=1}^{J} G_j/(1 - G_0) = 1$, so that we have a weighted average.[8]

Identification of the unobserved product qualities

We have identified all components of demand except for the unobserved product specific quality ξ_j. The unobserved product specific quality ξ_j is identified from product level (i.e. aggregating over versions) market share equations. We show that the market share equations can be inverted to recover a unique vector of ξ as in Berry et al. (1995).

The market share equation for $j = 1, ..., J$ can be written as

$$s_j = \int G_j(u(\alpha))\hat{f}(\alpha)\alpha = \int \frac{\exp(u_j(\alpha))}{\sum_{j'=1}^{J} \exp(u_{j'}(\alpha))} \sum_{j'=1} s_{j'} f_{j'}(\alpha) d\alpha, \tag{10}$$

where the second equality can be obtained using (9) and some algebra. This equation differs from standard BLP, because the outside good does not show up in the denominator and $\sum_{j'=1} s_{j'} f_{j'}(\alpha)$ does not integrate to 1. To take care of this define $\tilde{s}_j = s_j/(1 - s_0)$ the market share conditional on not choosing the outside good. Also define the normalized random coefficients $\tilde{\xi}_j = \xi_j - \xi_1$ for $j = 1, ..., J$. Recall that we defined $u_j(\alpha) = \xi_j + \tilde{u}_j(\alpha)$.

[8] Note that we could consider an alternative estimator

$$\hat{\hat{f}}(\alpha) = \sum_{j'=1} s_j^2 f_j(\alpha)/[(1 - s_0)G_j(u(\alpha))]$$

which is also a weighted average but uses the true outside good market share rather than the theoretical one. Our estimator is preferable as it is more robust to outliers. To clarify, consider the following situation. For a particular type the predicted probability of choosing the a product $G_j(u(\alpha))$ is close to zero, but we observe an outlier: a consumer with α chooses product j. Then the alternative estimator $\hat{\hat{f}}$ would explain this by a very large density $\hat{\hat{f}}(\alpha)$, whereas our estimator $\hat{f}$ defined by (9) would only be slightly affected.

Now, the above equation can be rewritten as

$$\tilde{s}_j = \int \frac{\exp(\tilde{\xi}_j + \tilde{u}_j(\alpha))}{\exp(\tilde{u}_1(\alpha)) + \sum_{j'=2}^{J} \exp(\tilde{\xi}_{j'} + \tilde{u}_{j'}(\alpha))} \sum_{j'=1} \tilde{s}_{j'} f_{j'}(\alpha) d\alpha, \tag{11}$$

for $j = 2, ..., J$. Note that ξ_1 cancels out and $\sum_{j'=1} \tilde{s}_{j'} f_{j'}(\alpha)$ integrates to 1. Then (11) is just the standard BLP market share equation, with the twist that product $j = 0$ is excluded for the moment and that product $j = 1$ plays the role of the outside good for the moment. Therefore, we can simply use standard BLP argument for identifications of $\tilde{\xi}_2, ..., \tilde{\xi}_J$. We can also use their contraction mapping to back them out.

To get ξ_1, we have to use the market share equation for the outside good:

$$s_0 = \int G_0(u(\alpha)) \hat{f}(\alpha) d\alpha = e^{-\xi_1} \int \frac{1}{\sum_{j'=1} \exp(\tilde{\xi}_{j'} + \tilde{u}_{j'})} \sum_{j'=1} s_{j'} f_{j'}(\alpha) d\alpha. \tag{12}$$

Observe that in the above equation, everything except ξ_1 is known, so one can trivially solve for ξ_1.

Identification of the costs

The cost function $C_j(\cdot)$ for all j can be identified using the Euler equations of the profit maximization problem described in Proposition 1. Rather than solving forward for u_j, q_j, we solve backward for C_j.

3.3 Comparison with standard models

We first compare our identification results with Luo et al. (2014) who show identification of the standard model of non-linear pricing under monopoly described in Mussa and Rosen (1978). We face the additional challenges that consumer's types are multidimensional rather than one-dimensional and that we do not have observations about the price sensitivities of consumers who choose the outside good $j = 0$. However, we have additional data: we have prices and market shares at the version level for multiple products and not just one. Further, while we have partial exclusion for all types t, we do not have full exclusion for any type t, which contrasts with Luo et al. (2014) result that the distribution of types cannot be recovered for $t < t^*$.

Luo et al. (2014) assume a linear cost function, but identify non-parametrically consumers' utility and the density of types. In contrast, we assume linear utility (a standard assumption in the BLP literature), but are non-parametric on the density of types and impose a flexible form on the cost function.[9] Another difference is that Luo et al. (2014) use the supply side structure for demand identification, since the supply side in their model implies simple first-order conditions; whereas the Euler equation in our setting cannot be used for demand side identification due to its complexity. Instead, we use a combination of within and between product variations for identification.

The identification of the standard BLP model for differentiated products relies on a different type of variation in the data compared to our model. Indeed, here we rely on within model variation of characteristics and market shares of versions which is absent, at least in the *product BLP* setting. On the contrary, BLP model is identified using cross-products and cross-market variations. Cross-market variation is actually crucial to precisely estimate the heterogeneity parameters.In our approach, we do not use cross-markets variation for the identification of the distribution of consumers' types. Instead we rely only on within-product variation together with restrictions implied by the theoretical model. Of course data on multiple markets is useful to improve the precision of the estimates. One other key difference between our model and the BLP standard approach is that we do not rely on the assumption that product characteristics are exogeneous or instruments. In the standard BLP setting, product characteristics are considered to be exogenous but the price is allowed to be endogeneous. Typical instruments are cost shifters and function of other products characteristics, which are valid under the assumption of exogeneous product characteristics. It is often difficult to get data on cost shifters that vary across products (typical cost shifters such as wages or input prices display little variation across products). This is one reason why endogenizing the quality choice in differentiated products market is in general challenging and the papers that relax the assumption of exogeneous product characteristics rely instead on

[9] We do not identify non-parametrically the cost function because the cost functions are given by a boundary value differential equation system, for which there is no general result for uniqueness of the solution. However, for any parametric form of the cost function, there is a unique solution for the parameters of the cost function.

alternative instruments that are valid because of the specificity of the market they analyze.[10] By assuming that any level of quality is offered for every product, we do not model the choice of product characteristics or quality and our approach does not rely on instruments.

Next, in contrast to BLP, we identify non-parametrically the heterogeneity of price sensitivities while it is standard to impose a parametric form without individual-level data.On the other hand, our model contains a restriction compared to the the standard BLP approach: we do not allow for heterogeneity on product characteristics (β). What we can allow for is heterogeneity of preferences for characteristics that are invariant across versions of a product. For instance, we can assume that the utility depends on the characteristics of the baseline version of the product, regardless of the level of quality chosen. The preferences for characteristics of the baseline version can be seen as a proxy for the preferences the product, and we can allow heterogeneity of these preferences. Denote by $\bar{X}_j$ the characteristics of the baseline version of product j. We consider the utility of consumer i choosing the quality level q of product j is:

$$U_{ij}(q_j) = q_j + \bar{X}_j\tilde{\beta}_i - \alpha_i p_j(q_j) + \epsilon_{ij},$$

where $\tilde{\beta}_i$ represents consumer specific preference for the baseline version characteristics. We cannot allow for heterogeneity on characteristics that are version specific because our supply side model would become a multidimensional screening problem and could not be leveraged as we do here for the identification of the demand and supply primitives. Note that if we allow for this specific type of heterogeneity, our model is equivalent to the *product BLP* model when there is no price discrimination and quality increments are priced at marginal cost. However this is different from the *product-version BLP* model in which one can introduce heterogeneity of preferences over version characteristics.

Finally, the substitution patterns our model implies are less restrictive than the *product BLP* but are more restrictive than the *version-product BLP*. Our model implies that each

[10] Crawford et al. (2011) take advantage of the feature that cable TV is offered by local providers but they belong to the same system operator. They use the average prices and qualities of providers from the same operator in other local markets as instruments. Fan (2013) relies on the partial overlap of newspapers circulation's areas so the demand shifters in non-overlapping markets of a newspaper's competitors are valid instruments for newspapers' qualities.

version of a product substitutes to only one version of other products, which is the same as in *product BLP*, the baseline version competes with other baseline versions of the other products. However in the *version-product BLP*, all the versions of other products are substitutes to a given version of a product. Because of the logit error assumption, even the most flexible specification demand implies strictly positive cross price elasticities.

4 Application to the automobile market

4.1 Data and descriptive evidence

To estimate the model, we use a dataset that contains all the registrations of new cars for the year 2007. We observe the main characteristics of the vehicles (brand, model, horsepower, trim, weight, number of seats, number of doors and body type) and the price. Prices come from manufacturers' catalogues and were merged with the car registrations dataset. We consider each car model name to be a different product. We consider two versions to be different if they display different prices and/or characteristics in the same year. We exclude the car models for which we observe less than 20 versions sold because we need a sufficient number of versions of a model to have a precise estimate of the price schedule.[11] These excluded car models represent a small fraction of the sales, 26% of the market. We obtain a total of 73 car models that belong to 13 different manufacturers. The total number of versions per model observed is between 20 and 214. The average is 50 versions per model. Table 1 displays statistics for the 10 models offering the largest number of versions.

[11]Thanks to some additional restrictions, we could define a lower threshold on the number of versions.

Table 1: Top 10 models in terms of number of versions

Brand	Model	No. versions	Share
Renault	Megane	214	8.6%
BMW	Serie 3	187	1.2%
Ford	Focus	160	2.3%
Renault	Clio	114	9.1%
Peugeot	207	111	9.3%
Volkswagen	Golf	99	2.3%
Peugeot	307	97	3.7%
Audi	A6	86	0.2%
Citroen	C4	86	8.3%
Opel	Astra	86	1.2%

We observe considerable price dispersion across car versions as Table 2 suggests, sharing the price difference between the most expensive and the cheapest version of a car model. The sales weighted average gross price difference is 15,850 euros. The maximum of 60,100 euros corresponds to the Mercedes CLK-class. Figure 1 displays the distribution of maximal price difference across versions of the car models. Price dispersion concerns the majority of car models.

Table 2: Maximum price difference across versions of a model weighted by model market shares

Average	Minimum	Maximum
15,849	2,700	60,100

Figure 1: Distribution of maximal price difference over models

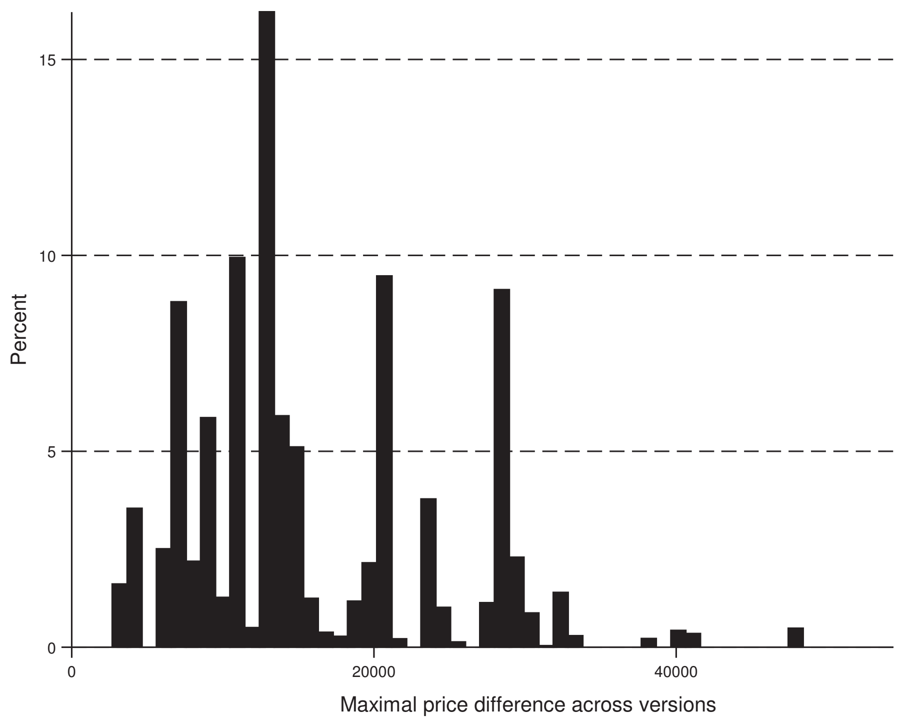
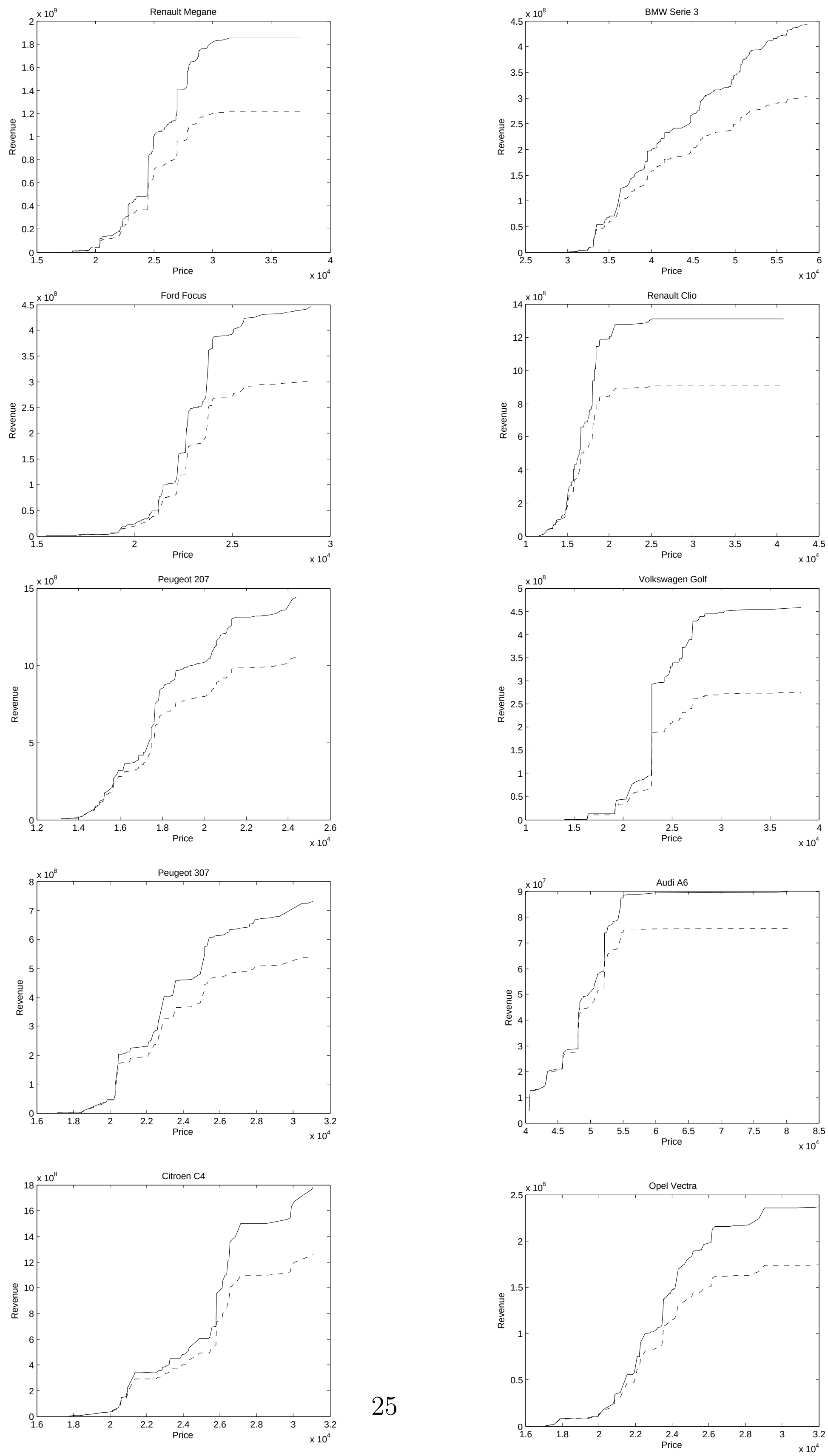


Figure 2 displays the revenue from selling the baseline version and the higher quality levels for the top 10 models that offer the largest number of versions. The solid line represents the actual revenue while the dashed line represents the hypothetical revenue they would make if the price was uniformly set to the one of the baseline version. The revenue would be roughly two thirds of the actual revenue. Of course, this is not a measure of the importance of price discrimination as the cost of quality matters. The importance of price discrimination for firms' profits will be assessed using the structural model of market equilibrium that disentangles markups from the costs.

Figure 2: Revenue from the baseline version and from the add-ons for the top 10 car models



4.2 Estimation

Estimation of price schedules and the quality index

We first estimate the price schedule of each car model and the quality index of each car version assuming that quality is a linear combination of the following observable car characteristics: horsepower, fuel cost, weight, cylinder capacity, and dummies if the car is a convertible, has three doors, has a station-wagon body and uses diesel. We also use squared variables to allow for non-linear effects. These variables together explain 97.1% of the intra-product price variation. This is consistent with our model in which the quality of a version v of product j depends on some unobservables η_{jv}.

We approximate p_j^{-1} by a twice integrated negative fifth order Bernstein polynomial. This parametrization has the advantage that we can easily impose restrictions on the slope and curvature of the function to take advantage of the theoretical restrictions. Recall that a nth order Bernstein polynomial has the following expression:

$$B_n(f)(x) = \sum_{i=0}^{n} a_i b_{n,i}(x)$$

With

$$b_{n,i}(x) = \binom{n}{i} x^i (1-x)^{n-i}$$

Formally, the inverse pricing function p_j^{-1} is

$$p_j^{-1}(x) = \sum_{i=0}^{n} a_i \binom{n}{i} \int_0^x \int_0^y z^i (1-z)^{n-i} dz dz + a_0^* + a_1^* x$$

To gain precision, we assume here that the β_j are identical for products that belong to the same segment. We define three car segments: compact cars, family cars, and luxury cars.

We estimate β_s and the coefficients of the Bernstein polynomial using non-linear least squares to minimize $\sum_{jv} \eta_{jv}^2$. The coefficients a_i, $i = 0, ..., n$ are constrained to be negative to ensure that p_j^{-1} is concave. A concave p_j^{-1} implies a convex p_j, which in turn implies that the second-order condition of the consumer's quality choice problem $[q - \alpha p_j(q)]'' < 0$ is satisfied. We constrain a_0^* and a_1^* such that p_j^{-1} is increasing at the upper bound, respectively. This ensures that the function is positive and increasing everywhere in the support. We also force the slopes of all the products within a segment to be identical at the minimum and maximum prices. Formally we impose:

$$(p_j^{-1})'(\max p_{jv}) = \bar{q} \quad \forall j$$
$$(p_j^{-1})'(\min p_{jv}) = \underline{q} \quad \forall j$$

This makes sure that the conditional distributions of types have the same support for all products. This restriction is made mostly for convenience, it should be possible to do the estimation without this assumption, albeit at the cost of more complexity. The restriction is not as strict as it may appear: it states that a consumer of type $\underline{\alpha}$ compares the highest quality $\overline{q}_j$ version of all products j within a segment s. But it does not impose any constraint for the second highest and intermediate versions of a model. Typically, only a very small fraction of consumers buy the highest quality version of a product. The same applies for the lower bound of support with $\overline{\alpha}$ and $\underline{q}_j$. The results are displayed in Table 3.

Table 3: Estimation results for quality index parameters

Variable	Compact	Family	Luxury
Weight	1.53	0.74	2.71
Cyl Cap	0.09	-0.03	-0.22
Fuel cost	-1.51	-0.37	-1.67
Horsepower	-0.03	-0.02	0.01
Cyl Cap^2	-3.54	0.13	4.56
Fuel cost^2	7.13	1.16	4.78
Horsepower^2	5.88	5.42	-0.71
Coupe	-1.64	0.09	-0.78
Three doors	0.14	0.29	0.95
Stat. Wagon	-0.71	-0.49	-0.61
Diesel	-2.56	-1.07	-4.65

The density of types is displayed in Figure 3. Recall that a low α represents a low price sensitivity or equivalently a high preference for quality. As we expect, consumers are rather price sensitive but a non-negilgible fraction have very low price sensitivities.

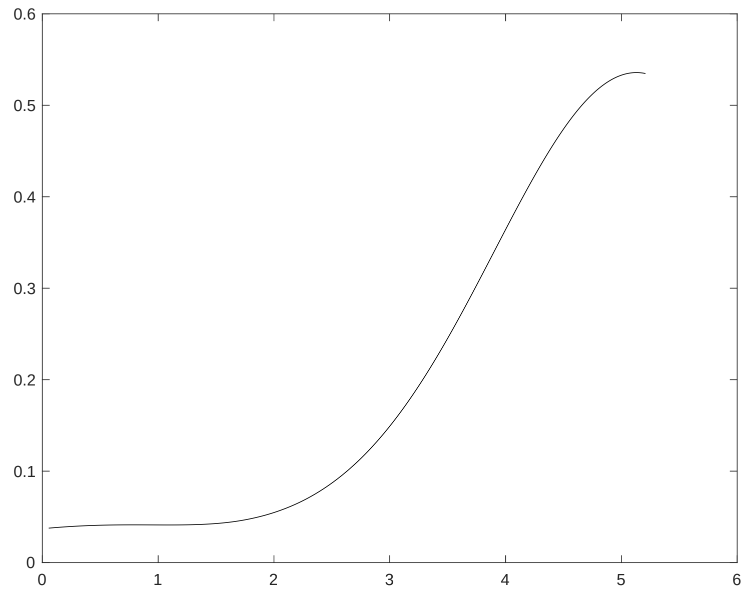


Figure 3: Estimated density of types α

Estimation of the marginal costs

We use the first-order condition provided in Proposition 1 to back out costs. We solve the equations backwards for the cost functions C_j, $j \in J_m$.

To solve for the functions C_j numerically, we parametrize C_j with a Bernstein

$$\hat{C}_j(q) = \sum_{l=0}^{N} \gamma_{j,l} T_l(\frac{q}{\bar{q}})$$

where $\frac{q}{\bar{q}}$ makes sure that q is linearly transformed into the domain of the Bernstein polynomial, $[0, 1]$.

We evaluate the equation (5) for different values of t_i. Denote the difference between the LHS and the RHS of (5) by ν_i. Further, define the violation of the boundary value condition by $\nu_{\bar{t}} = \hat{C}_j'(q_j(\bar{t})) - \bar{t}$.

We solve the equation system by minimizing the violations of the conditions ν_i:

$$\min_{\{\gamma_{j,l}\}} \sum_{t=\underline{t}}^{\bar{t}} \nu_i^2$$

Since the coefficients enter in the objective function quadratically, the solution is given by the first-order condition which is a linear equation system. For our estimates, the linear equation system is invertible and we get a unique solution.

In the following we report the estimated marginal costs. Figure 4 shows prices and marginal costs for the Renault Megane and the Ford Focus. Figure 5 shows the markup as a function of quality. It can be seen that the markup is increasing in quality, i.e. firms price discriminate despite the presence of competition. Estimates for other car models are work in progress.

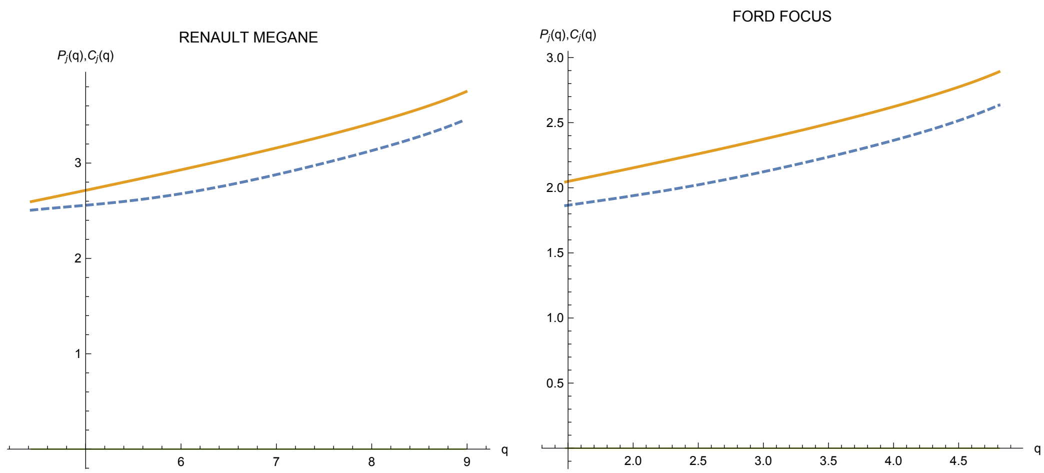


Figure 4: Prices and costs as a function of the quality level q for the Renault Megane and the Ford Focus. Prices and costs are in €10,000.

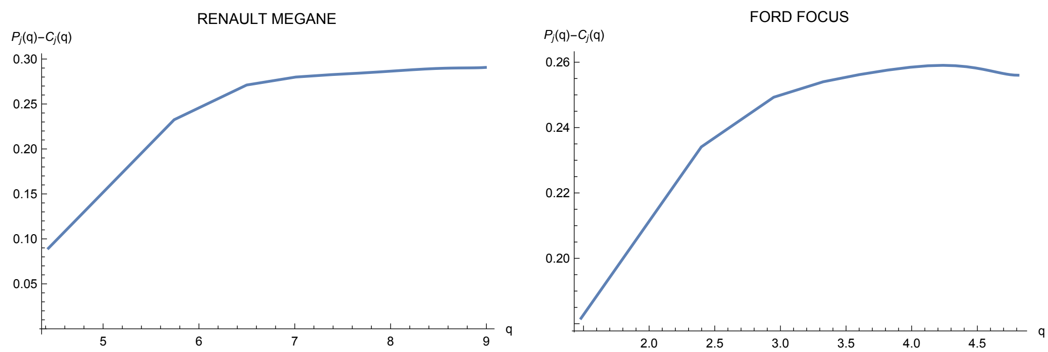


Figure 5: Absolute markups (in €10,000) as a function of the quality level q for the Renault Megane and the Ford Focus.

The markup of a version of a model $p_j(q) - c_j(q)$ can be decomposed into two parts: the baseline markup $p_j(\underline{q}_j) - c_j(\underline{q}_j)$ and the the add-on markup $[p_j(q) - c_j(q)] - [p_j(\underline{q}_j) - c_j(\underline{q}_j)]$. One can also decompose the profits a manufacturer makes into baseline profits (baseline markup times sold quantity) and add-on profits (add-on markup weighted by sold quantity of an individual version). Formally, the total profit is

$$\pi_j = \int [p_j(q_j(t)) - c_j(q_j(t))]\, G_j(\boldsymbol{u}(t)/t) dF(t)$$

The base profit is

$$\pi_j^{\text{baseline}} = \int \left[p_j(\underline{q}_j) - c_j(\underline{q}_j)\right] G_j(\boldsymbol{u}(t)/t) dF(t)$$

and the add-on profit is simply $\pi_j^{\text{addon}} = \pi_j - \pi_j^{\text{baseline}}$.

For the Renault Megane we get that the ratio of add-on profits to total profits is $\pi_j^{\text{addon}}/\pi_j = 0.69$. For the Ford Focus we get $\pi_j^{\text{addon}}/\pi_j = 0.29$.

4.3 Next Steps, Robustness Checks, Extensions

- confidence intervals (using boostrap)
- Test how robust the results are to the use of only one characteristic as quality measure
- Use a nonlinear utility function $U(q) = (1 - \gamma)q^{1-\gamma}$
- Compare estimates to what we would get from a standard BLP. We can directly compare ξ and β_s and α_i

4.4 Counter-factual simulations

Simulation exercises we could do to assess the importance of non-linear pricing:

- Simulate the "competitive" equilibrium: cost + fixed mark-up
- Simulate the monopoly equilibrium (Mussa-Rosen model)
- Predictions from a merger simulation and how different it is from BLP prediction

References

Armstrong, M. and J. Vickers (2001): "Competitive price discrimination," *RAND Journal of Economics*, 579–605.

Aryal, G. (2013): "An Empirical Analysis of Competitive Nonlinear Pricing," Tech. rep.

——— (2016): "Identifying Multidimensional Adverse Selection Models," Tech. rep.

Aryal, G., I. Perrigne, and Q. Vuong (2016): "Identification of Insurance Models with Multidimensional Screening," Tech. rep.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, S. T. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 242–262.

Byrne, D. P. (2015): "Testing Models of Differentiated Products Markets: Consolidation in the Cable TV Industry," *International Economic Review*, 56, 805–850.

Crawford, G. S., O. Shcherbakov, and M. Shum (2011): "The welfare effects of monopoly quality choice: Evidence from cable television markets," *Unpublished paper*.

D'Haultfoeuille, X., I. Durrmeyer, and P. Février (2014): "Automobile Prices in Market Equilibrium with Unobserved Price Discrimination," Tech. rep., SFB/TR 15 Discussion Paper.

D'Haultfœuille, X. and P. Février (2015): "The Provision of Wage Incentives: A Structural Estimation Using Contracts Variation," .

Fan, Y. (2013): "Ownership consolidation and product characteristics: A study of the US daily newspaper market," *The American Economic Review*, 103, 1598–1628.

Fan, Y. and C. Yang (2016): "Competition, product proliferation and welfare: A study of the us smartphone market," .

GRIGOLON, L., M. REYNAERT, AND F. VERBOVEN (2014): "Consumer valuation of fuel costs and the effectiveness of tax policy: Evidence from the European car market," .

IVALDI, M. AND D. MARTIMORT (1994): "Competition under nonlinear pricing," *Annales d'Economie et de Statistique*, 71–114.

JULLIEN, B. (2000): "Participation constraints in adverse selection models," *Journal of Economic Theory*, 93, 1–47.

LUO, Y., I. PERRIGNE, AND Q. VUONG (2014): "Structural analysis of nonlinear pricing," Tech. rep.

MIRAVETE, E. J. AND L.-H. RÖLLER (2003): "Competitive non-linear pricing in duopoly equilibrium: the early US cellular telephone industry," .

MUSSA, M. AND S. ROSEN (1978): "Monopoly and product quality," *Journal of Economic theory*, 18, 301–317.

NURSKI, L. AND F. VERBOVEN (2016): "Exclusive dealing as a barrier to entry? Evidence from automobiles," *The Review of Economic Studies*, 83, 1156–1188.

ROCHET, J.-C. AND L. A. STOLE (2002): "Nonlinear pricing with random participation," *Review of economic Studies*, 277–311.

VERBOVEN, F. (2002): "Quality-based price discrimination and tax incidence: evidence from gasoline and diesel cars," *RAND Journal of Economics*, 275–297.

Appendix

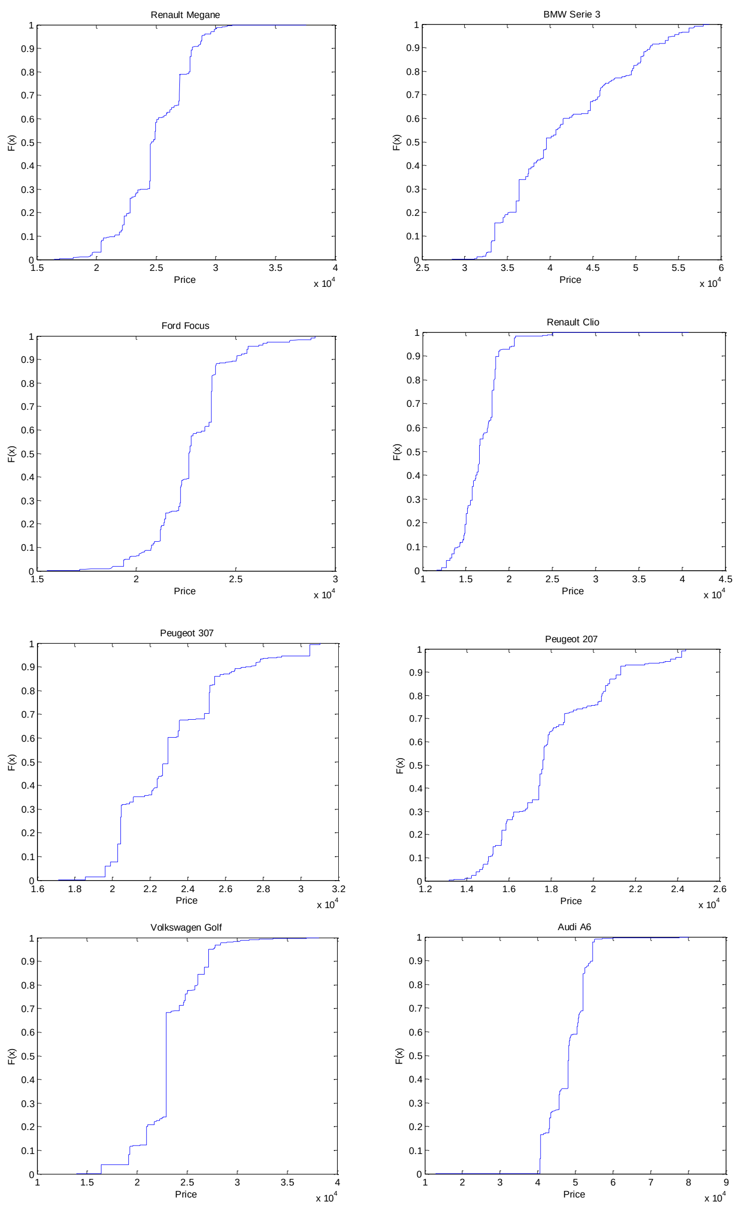


Figure 6: Empirical cumulative distribution of prices for the 8 car models with the largest number of different versions