Competitive Non-linear Pricing: Evidence from the French Automobile market

Isis Durrmeyer, Andras Niedermayer, Artyom Shneyerov*

October 30, 2016
Preliminary and Incomplete

Abstract

We develop a new empirical model of market equilibrium with second-degree price discrimination and oligopolistic competition based on an extension of Rochet and Stole (2002) non-linear pricing theory to multiproduct firms. The demand system is semi-parametrically identified. We estimate the model using French automobile data and take advantage of observing prices and market shares at the car model version level. We test the existence of second-degree price discrimination under imperfect competition. We extend the structural analysis of nonlinear pricing to an oligopolistic setting (see Luo, Perrigne, and Vuong, 2014). Our demand estimate is semi-parametric and does not rely on the exogeneity of characteristics assumptions and on instruments.

*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, EEA 2016 in Geneva, and participants of seminars at the University of Mannheim and the Toulouse School of Economics for helpful comments. [further acknowledgments to be added]

1 Introduction

Second-degree price discrimination plays an important role in many oligopolistic markets. As an example, more than 200 versions of the Renault Mégane are offered in France, with the price ranging from EUR 16,000 for the 72 horsepower version to over EUR 31,000 for the 127 horse power version. One of the many other examples is the telecommunications industry, with prices paid for mobile communication strongly varying depending on minutes talked, data downloaded and data speed, and the number of text messages sent.

Another important characteristic of these industries is that firms are very often multi-product firms. Renault does not only produce the Mégane, but also the Clio and several other car models: eight models under the brand *Renault* and one under the brand *Dacia*. Renault is no exception: over 70 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 a model of oligopolistic competition. 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. While we can estimate non-parametrically the distribution of consumers' type, we make a parametric assumption on the distribution of the heterogeneity of preferences over products. On the supply side, 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 and marginal costs. The estimation method is semi-parametric and relies on the assumption that sellers' pricing strategy follows second-degree price discrimination and induces consumers to self-select. The estimation method requires the observation of market shares and prices of products at the quality 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 72 horsepower version of the Renault Mégane). Price dispersion across versions of the same model is rather important in this market. 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 allows us to use data on different quality levels of a product (for our empirical application – different versions of a car model) rather than aggregate at the product level. We show that this model is semi-parametrically identified and develop an estimation technique to estimate the underlying distributions of consumers' price sensitivities without relying on instruments.

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 discrimination 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

[1] These conditions are that the firms are symmetric (i.e. they have the same costs to produce a certain level of quality), that the market is sufficiently competitive (i.e. the products are not too differentiated), and there is no outside good.

problem under these conditions. This lack of closed form solutions 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) results to multi-product firms.

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]

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 and non-linear pricing. So far, the literature has assumed that there is only one version of a product

[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 70 horse power baseline version of the Megane from Renault and the quality increment to 100 horse power from Ford.

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

being offered or aggregated the different version of a product to one version and in some papers to two versions (see e.g. Verboven (2002)). Our novelty lies in making use of the information on all the versions of a car model 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). Further, we estimate demand non-parametrically and without relying on characteristics assumptions and on BLP weak instruments.

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

The paper is structured as follows. Section 2 describes the theory of competitive non-linear pricing with multiproduct firms. Section 3 derives the empirical model and the identification strategy. In Section 4 we apply our methods 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 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 consumers that choose one product among the set $J = \sum_m J_m$ available. They also choose the quality level q_j. We follow the model developed by Rochet and Stole (2002) and assume that a consumer i's utility from product j at quality level q_j is

$$U_{ij}(q_j) = t_i q_j - p_j(q_j) - x_{ij} \tag{1}$$

where q_j is the quality chosen by the consumer, t_i is the willingness to pay for quality of consumer i and is drawn from the distribution F. $p_j(\cdot)$ is the price schedule of the firm, and x_{ij} is the outside option. The outside option is the utility derived from a competing product or from not buying any product.

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

$$q_j^*(t) = \arg\max_q tq - p_j(q) \tag{2}$$

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

$$t - p_j'(q_j^*(t)) = 0 \tag{3}$$

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 gets from choosing product j:

$$u_j(t) = tq_j^*(t) - p_j(q_j^*(t))$$

A consumer chooses product j if

$$u_j(t_i) - x_{ij} > \max\{u_0, \max_{j'} u_{j'}(t_i) - x_{ij'}\} \tag{4}$$

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 notation, which is used in the theoretical literature, and the standard approach in the empirical literature. 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

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

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

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

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

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

$$\hat{u}_j(\alpha_i) + \epsilon_{ij} > \max\{\hat{u}_0 + \epsilon_{i0}, \max_{j'} u_{j'}(\alpha_i) + \epsilon_{ij'}\} \tag{5}$$

Comparing (4) and (5) shows that the two notations are equivalent with $t_i = 1/\alpha_i$, $u_j(t_i) = \hat{u}_j(\alpha_i)/\alpha_i$, $u_0 = (\hat{u}_0 + \epsilon_0)/\alpha_i$ and

$$x_{ij} = \frac{\max\{\hat{u}_0 + \epsilon_0, \max_{j'} \hat{u}_{j'}(\alpha) + \epsilon_{ij'}\} - \epsilon_{ij}}{\alpha_i}$$

It is interesting to note that for a single product firm, 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

On the supply side we develop a model of competition with non-linear pricing. Firms compete in price schedules as opposed to the standard setting where firms set only one price. Our model follows Rochet and Stole (2002) and extends their approach to allow to consider 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

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

The profit associated with product j is

$$\pi_j(p) = \int_{\underline{t}}^{\bar{t}} M_j(\boldsymbol{u}(t), t)[p_j(q_j^*(t)) - C_j(q_j^*(t))]dF(t) \tag{6}$$

where the cost of producing quality q is $C_j(q)$, the total demand for product j is

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

The total demand is

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

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

The optimal price schedule can be derived using optimal control. 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

$$\max_{\{u_j,q_j\}_{j\in J_m}} \sum_{j\in J_m} \int_{\underline{t}}^{\bar{t}} M_j(\boldsymbol{u}(t),t)[tq_j(t) - u_j(t) - C_j(q_j(t))]dF(t) \tag{7}$$

subject to the constraints $u_j' = q_j$ for all $j \in J_m$.

Proposition 1. *A fully separating solution to the optimal control problem* (7) *satisfies the following necessary first-order condition (the* Euler *equation):*

$$\begin{aligned}&\left[\sum_{j'\in J_m} \frac{\partial G_{j'}(\boldsymbol{u}(t)/t)}{\partial u_j}(tq_{j'}(t) - C_{j'}(q_{j'}(t)) - u_{j'}(t)) - G_j(\boldsymbol{u}(t)/t)\right] f(t)\\ &= \frac{d}{dt}\left[G_j(\boldsymbol{u}(t)/t)f(t)\left(t - C_j'(q_j(t))\right)\right]\end{aligned} \tag{8}$$

and the boundary conditions

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

Proof. The proof follows the same logic as Rochet and Stole (2002) with the difference of taking into account multiproduct firms. Note that, as in Rochet and Stole (2002), there is no distortion both at the top ($\bar{t}$) and at the bottom ($\underline{t}$). □

For single product firms, this solution specializes to the solution provided by Rochet and Stole (2002). A much celebrated additional result from Rochet and Stole (2002) is that under some conditions, the solution of the single product firm problem becomes very simple: the price is simply costs 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

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

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. It is an empirical question, whether these conditions hold empirically (at least approximately). And if the conditions do not hold empirically, how close we are to cost plus fixed markup. In other words, to what extent are price differences between low-end and high-end versions explained by cost differences and to what extent are they explained by markup differences.

In the following, we will answer this question by first estimating demand and then the cost function, which will give us markups for different quality levels.

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 the heterogeneity is on the 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 parametrically identify the 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 parametric form on the cost function, we make a parametric assumption on the utility function and consider it is linear, as in the theoretical model.

3.1 Assumptions

We make some assumptions that 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:

$$G_j(\alpha) = \frac{\exp(u_j(\alpha))}{1 + \sum_{j'} \exp(u_{j'}(\alpha))} \tag{9}$$

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 types.

It is sometimes impossible to observe a reliable quality measure (see Luo et al. (2014)). Instead of relying on a unique variable to represent quality, we use several characteristics of the products to construct a quality index.

Assumption 3. *Separability of quality*

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

where the index jv stands for the version v of product j.

Note that while for the theoretical analysis we assume a continuum of quality levels $q_j(\alpha)$ for $\alpha \in \mathbb{R}$, for the empirical analysis we will use a discrete number of quality levels q_{jv}, $v = 1, ..., V_j$ – the quality levels that are observed empirically.

This assumption implies that the quality is a function of observable characteristics X_{jv} and unobservable characteristics, ξ_j and η_{jv}. 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. We assume that the error terms η_{jv} have a zero expected value over j, so that ξ_j captures the average per product unobserved quality.

3.2 Identification

Quality index

We first need to construct the quality index from the observed prices and characteristics of the versions of the products. For this, we rely on the theoretical model that implies that the price of a version p_{jv} depends on the quality q_{jv}:

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

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

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

β_j and p_j^{-1} are identified from the hedonic regression that uses the link between the characteristics X_{jv} of a version and its price p_{jv}. Note though that as long as one uses the hedonic regression alone (and not market share data), an affine transformation of p_j^{-1} is observationally equivalent to scaling β_j by a scalar and changing ξ_j. In other words, without market share data, we cannot identify ξ_j, we can 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 and the scale parameters

It is straightforward to back out the density of types, conditional on the product choice if one observes the true quality. Indeed, conditional on buying product j, the optimal choice of quality for a consumer is given by the first-order condition:

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

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 we recover the conditional density $\tilde{f}_j(\tilde{\alpha})$ (i.e. the density of $\tilde{\alpha}_j$ conditional on choosing j). We 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 the normal kernel:

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

To identify the scale parameters λ_j, we use some restrictions implied by the structural model. According to Bayes's rule for the joint probability of observing the type α and the product j

$$s_j f_j(\alpha) = G_j(\alpha) f(\alpha) \tag{10}$$

where $G_j(\alpha)$ is the probability of choosing model j, conditional on the type α (see Equation (9)). Equation (10) holds for all products and all consumer types. Expressing the ratio of Equation (10) for two different products, we get:

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

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 $\tilde{\alpha}_j$s and using the equality $u_j' = -p_j$ stemming from an envelope theorem argument, we get

$$\frac{\lambda_j \tilde{f}_j'(\lambda_j \tilde{\alpha}_j)}{\tilde{f}_j(\lambda_j \tilde{\alpha}_j)} - \frac{\lambda_{j'} \tilde{f}_{j'}'(\lambda_{j'} \tilde{\alpha}_{j'})}{\tilde{f}_{j'}(\lambda_{j'} \tilde{\alpha}_{j'})} = -\tilde{p}_j(\tilde{q}_j(\lambda_j \tilde{\alpha}_j)) + \tilde{p}_{j'}(\tilde{q}_{j'}(\lambda_{j'} \tilde{\alpha}_{j'})) \tag{11}$$

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 (11) 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 (11) for different values of α.

Identification of the density of types and the unobserved product qualities

The unobserved product specific quality is identified from market shares equations of products using a market share inversion that is in the spirit of Berry et al. (1995)'s contraction mapping. For this purpose, we define an estimator for f:

$$\hat{f}(\alpha) = \sum_{j=1}^{J} \frac{G_j(u(\alpha))}{1 - G_0(u(\alpha))} \left[\frac{s_j f_j(\alpha)}{G_j(u(\alpha))} \right], \tag{12}$$

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. [7]

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

$$s_j = \int G_j(u(\alpha))\hat{f}(\alpha)\alpha = \int \frac{\exp(u_j(\alpha))}{\sum_{j'=1}^{J}\exp(u_{j'}(\alpha))} \sum_{j'=1} s_{j'} f_{j'}(\alpha) d\alpha, \tag{13}$$

where the second equality can be obtained using (12) and some algebra. This equation differs from standard BLP, because the outside good does not show up 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 $u_j(\alpha) = \xi_j + \tilde{u}_j(\alpha)$.

Now, the above equation can be rewritten as

$$\tilde{s}_j = \int \frac{\exp(\tilde{\xi}_j + \tilde{u}_j(\alpha))}{\exp(\tilde{u}_1(\alpha)) + \sum_{j'=2}^{J}\exp(\tilde{\xi}_{j'} + \tilde{u}_{j'}(\alpha))} \sum_{j'=1} \tilde{s}_{j'} f_{j'}(\alpha) d\alpha, \tag{14}$$

for $j = 2, ..., J$. Note that ξ_1 cancels out and $\sum_{j'=1} \tilde{s}_{j'} f_{j'}(\alpha)$ integrates to 1. Then (14) is just the standard BLP market share equation, with the twist that product $j = 0$ is excluded for the moment and that model $j = 1$ should be considered 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:

$$s_0 = \int G_0(u(\alpha))\hat{f}(\alpha)d\alpha = e^{-\xi_1} \int \frac{1}{\sum_{j'=1}\exp(\tilde{\xi}_{j'} + \tilde{u}_{j'})} \sum_{j'=1} s_{j'} f_{j'}(\alpha) d\alpha. \tag{15}$$

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

[7] 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 version k, the predicted probability of choosing the model $G_j(u(\alpha))$ is close to zero, but we observe an outlier: a consumer with α chooses model j. Then the alternative estimator $\hat{\hat{f}}$ would explain this by a very large density $\hat{\hat{f}}(\alpha)$ at that point, whereas our estimator $\hat{f}$ defined by (12) would only be slightly affected.

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 Prop. 1. Rather than solving forward for u_j, q_j, we solve backward for C_j.

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 type of body) and the price. Prices come from manufacturers' catalogues and were merged with the car registrations dataset. We consider car models are of different levels of quality if they display different prices in the same year. We call this level of quality a *version* of a car model. We exclude the 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. These excluded models represent a small fraction of the sales, 26% of the market. We obtain a total of 73 models that belong to 22 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.

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

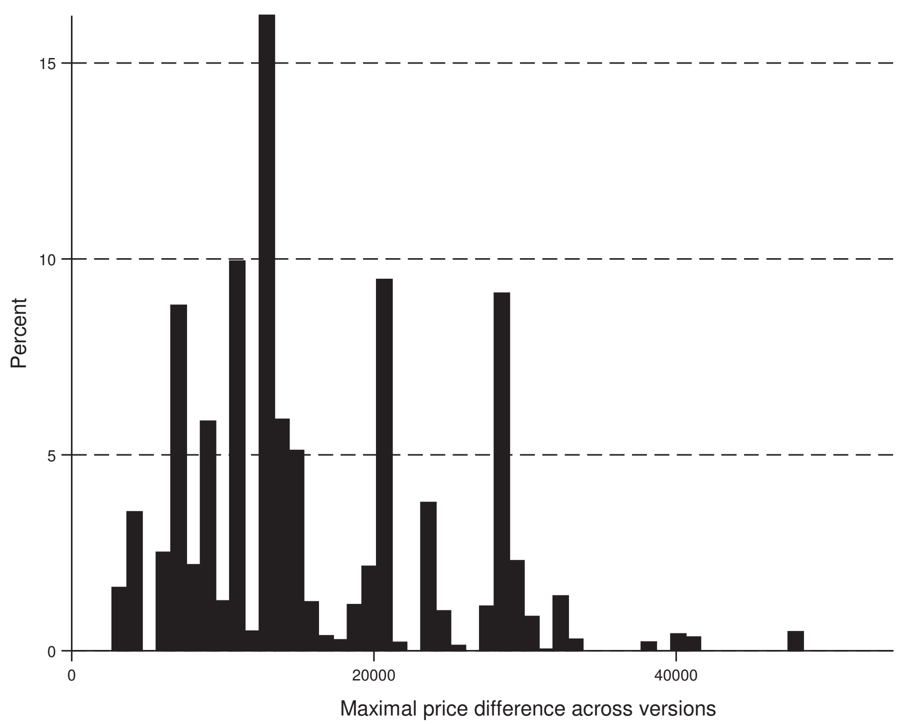
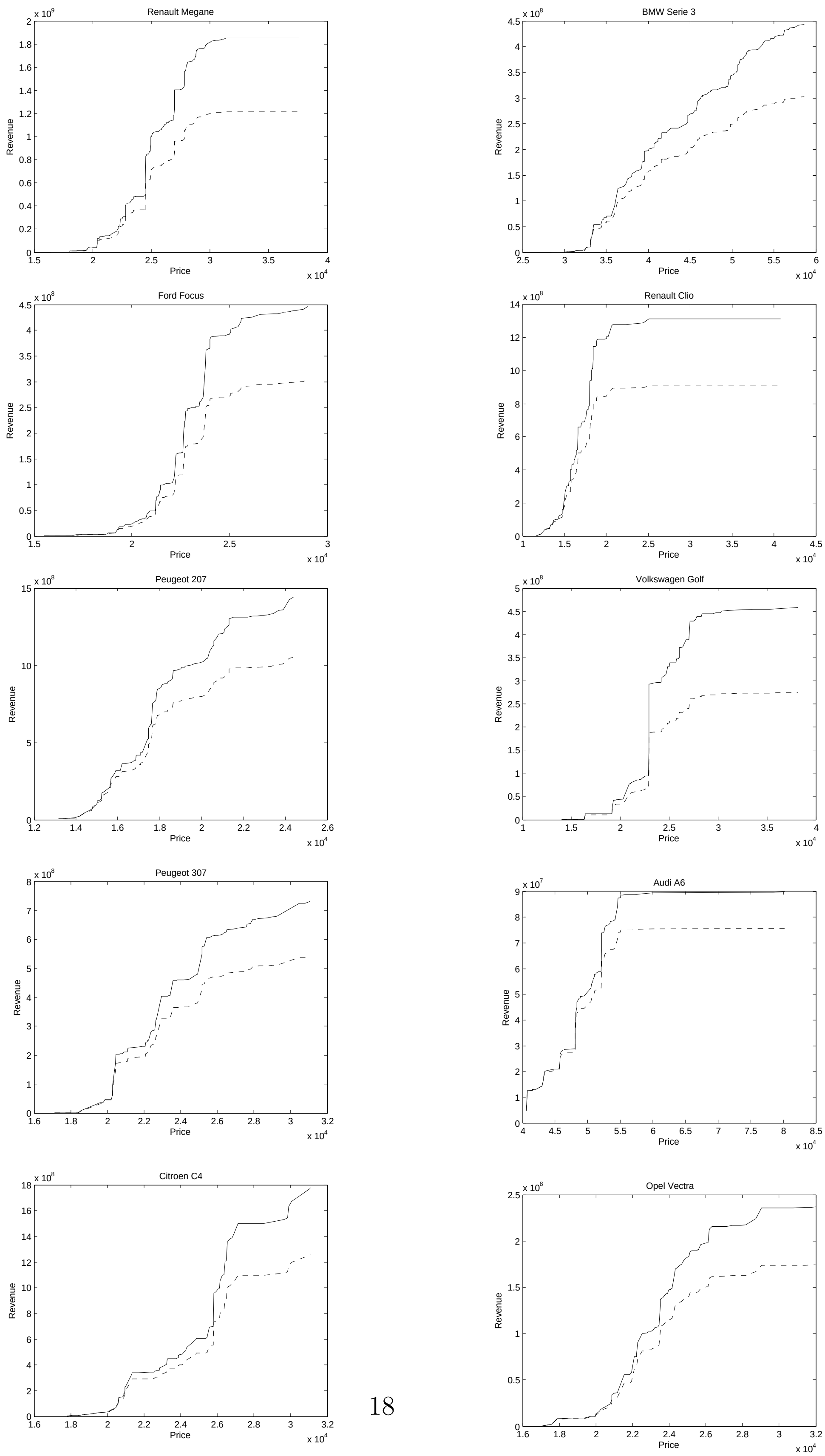


Figure 2 displays the revenue from selling the base version and the higher qualities for the top 10 models that offer the larger number of versions. The solid line represents the actual revenue while the dashed line represents the hypothetical revenue they would make if the price was uniformly set to the one of the baseline version. The revenue would be roughly two thirds of the actual revenue. Of course, this is an imperfect measure of the importance of price discrimination and the importance of price discrimination for firms' profits will be better understood using the structural model of market equilibrium.

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



4.2 Estimation

Estimation of the quality index

We first estimate the hedonic regression to construct the quality index. We assume that quality is a 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, uses diesel, is a minivan (with more than six seats). We also use squared variables and interactions between horsepower and dummies to allow for non-linear effects. These variables together explain 97.1% of the intra-model price variation. This is consistent with our model in which the quality of a version v of model 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 on the predictions from the theory. 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 that the β_j are identical for products

We estimate β 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 positive at the lower bound and 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 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. 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	Parameter
Horsepower	1
Horsepower^2	-0.003
HP$\times$coupe	0.087
HP$\times$stat. wagon	-0.003
HP$\times$MPV	-0.148
Fuel cost	-1.407
Fuel cost^2	0.09
Weight	35.299
Weight^2	-12.225
Cyl Cap	-45.696
Cyl Cap^2	9.037
Coupe	-9.846
Three doors	-0.207
Stat. Wagon	0.887
Diesel	3.557
MPV	12.572

The scale parameter λ is estimated to be 0.512. 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.

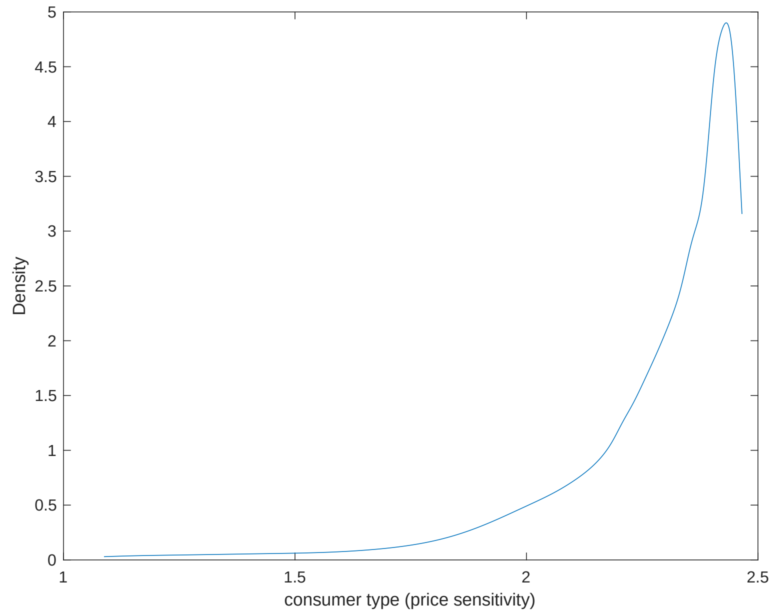


Figure 3: Estimated density of types α

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 Chebyshev polynomial

$$\hat{C}_j(q) = \sum_{l=0}^{N} \gamma_{j,l} T_l(2\frac{q}{\bar{q}} - 1)$$

where $2\frac{q}{\bar{q}} - 1$ makes sure that q is linearly transformed into the domain of the Chebyshev polynomial, $[-1, 1]$.

We evaluate the equation (8) for 20 different values of t_i. Denote the difference between the LHS and the RHS of (8) by ν_i. Further, define the violation of the boundary value condition by $\nu_{21} = \hat{C}_j'(q_j(\bar{t})) - \bar{t}$.

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

$$\min_{\{\gamma_{j,l}\}} \sum_{i=1}^{21} \nu_i^2$$

Since the coefficients show up 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.

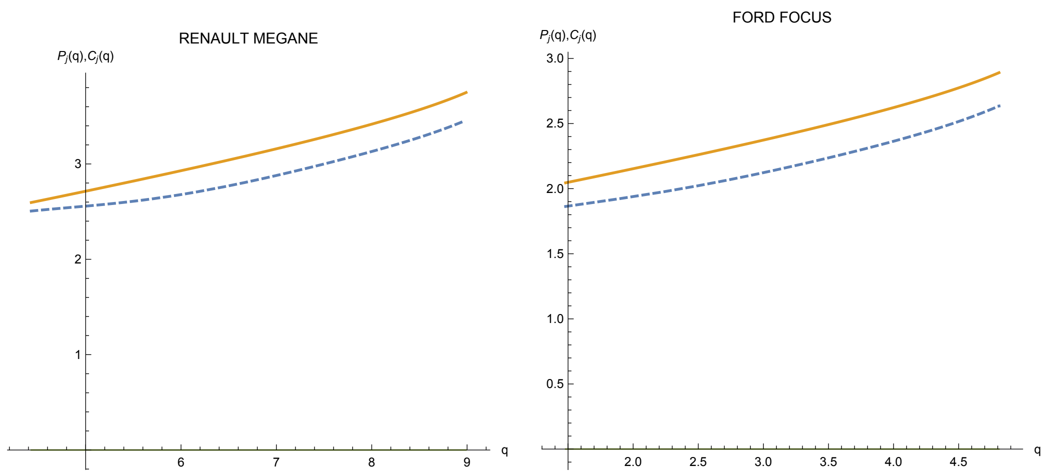


Figure 4: Prices and costs as a function of the quality level q for the Renault Megane and the Ford Focus. Prices and costs are in €10,000.

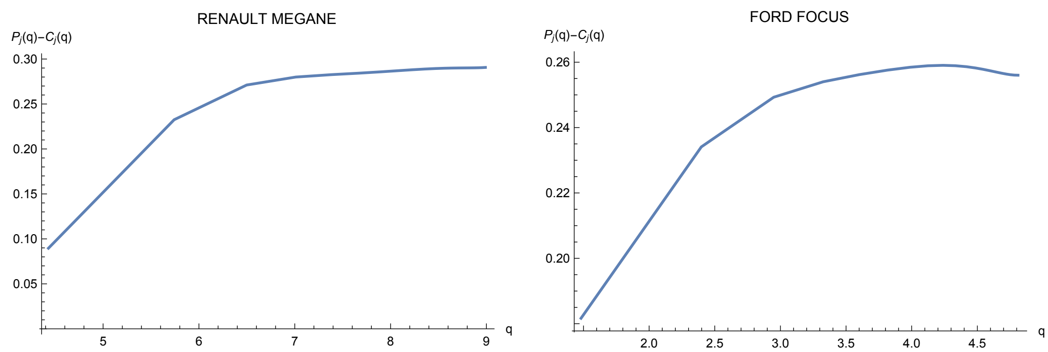


Figure 5: Absolute markups (in €10,000) as a function of the quality level q for the Renault Megane and the Ford Focus.

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 (boostrap estimation)
- Test how robust the results are to the use of horsepower only 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 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. Perringe, 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'Haultfœuille, X. and P. Février (2015): "The Provision of Wage Incentives: A Structural Estimation Using Contracts Variation," .

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," .

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

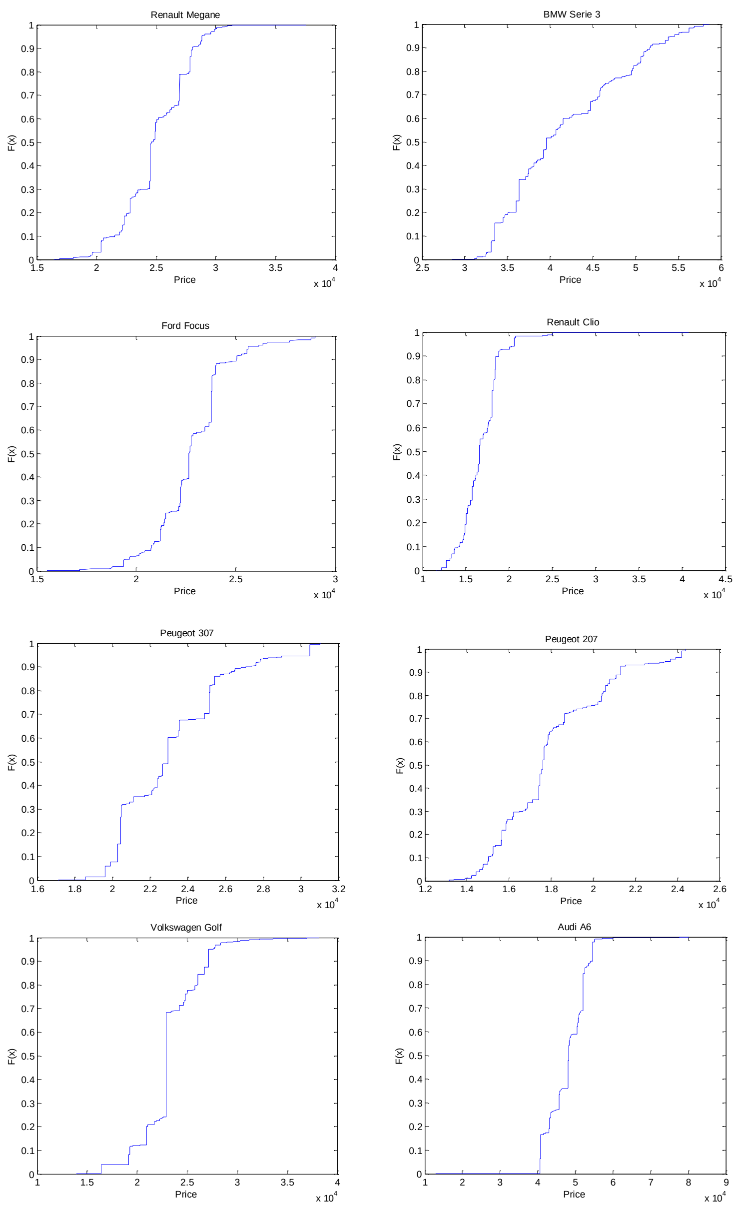


Figure 6: Empirical cumulative distribution of prices for the 8 car models with the largest number of different versions