GOVERNANCE AND THE EFFICIENCY
OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 477

Automobile Prices in Market Equilibrium with Unobserved Price Discrimination

Xavier D'Haultfoeuille*
Isis Durrmeyer**
Philippe Février***

* CREST
** University of Mannheim
*** CREST

September 2014

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.sfbtr15.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Dr. Klaus M. Schmidt · Department of Economics · University of Munich · D-80539 Munich,
Phone: +49(89)2180 2250 · Fax: +49(89)2180 3510

Automobile Prices in Market Equilibrium with Unobserved Price Discrimination.*

Xavier D'Haultfœuille[†] Isis Durrmeyer[‡] Philippe Février[§]

September 4, 2014

Abstract

This paper deals with the estimation of structural models of demand and supply with incomplete information on prices. When the seller is able to price discriminate, or the buyer to bargain, individuals pay different prices that are usually not collected in the data. This paper explores a method to estimate the supply and demand models jointly when only posted prices are observed. We consider that heterogenous transaction prices occur due to price discrimination by firms on observable characteristics of consumers. Within this framework, the identification is secured by (i) supposing that at least one group of individuals does pay the posted prices and (ii) assuming that the marginal costs of producing and selling the goods does not depend on the characteristics of the buyers. This methodology is applied to estimate the demand in the new automobile market in France. Results suggest that discounting arising from price discrimination is important. The average discount is estimated to be 5.2%, with large variation according to the buyers' characteristics. Our results are in line with discounts generally observed in European and American automobile markets.

*We would like to thank Steve Berry, Juan Esteban Carranza, Rob Clark, Pierre Dubois, Philippe Gagnepain, Gautam Gowrisankaran, John Morrow, Kathleen Nosal, Mario Pagliero, Philipp Schmidt-Dengler, Michelle Sovinsky, Frank Verboven and the participants of various seminars and conferences for their valuable comments. We also thank Pierre-Louis Debar and Julien Mollet from the CCFA for providing us with the data. Isis Durrmeyer acknowledges financial support from the Deutsche Forschungsgemeinschaft through SFB-TR 15.

[†]CREST. E-mail: xavier.dhaultfoeuille@ensae.fr

[‡]University of Mannheim. Corresponding author. E-mail: isis.durrmeyer@gmail.com

[§]CREST. E-mail: philippe.fevrier@ensae.fr

1 Introduction

The standard aggregate-level estimation of demand and supply models of differentiated products relies on the observation of the market shares and the characteristics of the products, in particular prices (see Berry, 1994). However, because of price discrimination and price negotiation, precise data on prices may be hard to obtain. Transaction prices for an identical product may indeed differ from one individual to another. Sellers can practice third degree price discrimination according to observable demographic characteristics such as age, gender or town of residence. There may also be room for individual negotiation, the sellers willing to offer discounts to consumers who argue about the prices and use competition to obtain better deals. Automobiles, furniture, kitchens or mobile phone contracts are examples for which there is either documented or anecdotal evidence that consumers receive some discounts (on automobiles, see, e.g. Harless & Hoffer (2002), Morton et al. (2003), Chandra et al. (2013) or Langer (2012)). Loans and insured mortgages have also proved to be negotiable (see Charles et al., 2008 and Allen et al. (2014)). Such phenomena also exist in vertical relationships between producers and retailers. Producers are required to edit general terms and conditions of sale. These conditions are then the starting point for individual negotiation with each retailer.

In all these cases, one typically observes either transaction prices on a small sample issued from a survey, or only posted prices on a large sample. If, in the first case, price discrimination can be studied, policy exercises cannot be performed. With large data, on the other hand, policy simulations are usually done without taking the issue of limited observation of prices into account. Because the instrumental variables approach used in Berry et al. (1995) to control for price endogeneity does not solve this nonclassical measurement error problem (namely, observing posted prices instead of transaction prices), ignoring it generally results in an inconsistent estimation of the structural parameters and biases in policy exercises.

This paper proposes a method to estimate a structural demand and supply model with unobserved transaction prices. Our rationale for the existence of transaction prices below posted prices is that they allow firms to price discriminate between heterogenous consumers and thus extract more surplus than they would with an uniform price.[1] Sellers have to edit only one price, namely the posted price, since it is usually legally forbidden to price

[1] In some cases, all profits could be higher if *all firms* did not price discriminate (see, e.g. Holmes (1989), Corts (1998)). But if one cannot commit not to price discriminate, price discrimination occurs for each firm at equilibrium.

discriminate between consumers and difficult to implement, but in practice the transaction prices differ from one individual to another. We suppose that sellers use observable characteristics of the buyers to price discriminate. Importantly, the market shares of all groups of buyers defined by these characteristics should be observed by the econometrician. This assumption may be problematic in settings where there are few buyers, such as the producers-retailers example. But it seems plausible that price discrimination is based on only a few easily observable characteristics, such as sex, age and city of residence, in markets where the sellers usually do not know the buyers before the transaction.

We then consider on the demand side the random coefficients discrete choice model popularized by Berry et al. (1995) with unobserved price discrimination. In this case, the contraction mapping that they use is no longer valid here, and we obtain biased estimates of mean utilities of products from market shares. Second, to recover parameters from the mean utilities, Berry et al. (1995) exploit the exogeneity of observed product characteristics (apart from price) and the sum of product characteristics of others to yield moment conditions involving these parameters. Here, however, transaction prices are unobserved and enter into the error term, so that all other product characteristics are endogenous.

Instead, we rely on the supply side, and replace the unobserved prices by their expression stemming from the first-order condition of profit maximization. These conditions have identifying power under two assumptions. First, the marginal cost of a product is supposed to be identical for all buyers. This amounts to neglecting differences in selling costs to different consumers in the total cost of a product. This assumption is likely to be satisfied in many markets, such as the automobile market, where the major part of the marginal cost is production, not sale, and the cost of selling is probably not very different from one consumer to another. The second condition concerns posted prices. We suppose that the sellers post the higher discriminatory price and then offer some discounts according to demographic characteristics in order to reach discriminatory optimal prices. This assumption is necessary since otherwise, we could shift all discounts by an arbitrary constant. It is also consistent with empirical evidence reported for instance by the UK Competition Commission on the automobile market, showing that some people actually pay the posted prices (see *New cars: A report on the supply of new motor cars within the UK* (2000)).

We then apply our method to the new automobile industry. Following Berry et al. (1995), the demand for automobile has always been estimated using posted prices when transaction prices are unobserved. Yet, many descriptive studies focusing on price discrimination (Ayres & Siegelman, 1995, Goldberg, 1996, Harless & Hoffer, 2002, Morton et al., 2003) have argued that important dispersion in transaction prices was at stake in the new car market. In France, it is also commonly admitted that negotiation is possible when purchasing a new car. An article published in *Le Figaro* (October 2011) suggests that discounts up to 26% can be obtained. We rely on an exhaustive dataset recording all the registrations of new cars bought by households in France between 2003 and 2008. Apart from detailed cars attributes, some characteristics of the owner are provided. We observe in particular age, expected income (namely, the median income of people in the same age class living in the same municipality). As these characteristics are easily observable by sellers and presumably strong determinants of purchases, we suppose that they are used to price discriminate.

Our results suggest that price discrimination is substantial in France. The average discount is estimated to be 5.2% of the posted price. The major part of estimated discounts is between 0 and 30% depending on the car purchased and demographic characteristics. Income appears to be the most important determinant of price discrimination and clearly negatively correlated to the value of discount. Purchasers under 40 with low income get the highest discount in average (9%). The average discount for middle-aged purchasers ais around 6%. The purchasers over 60 with a high income are usually those who pay the posted price while the ones with a low income get in average 4.9% discount. These results are in line with Harless & Hoffer (2002). Using observed dealer gross margins, they find evidence of price discrimination by age group, older buyers paying significant higher prices. The magnitude of discounts we obtain are also comparable to anecdotal evidences found in specialized magazines or on internet.

Finally, because unobserved discounts appear to be important here, we show that using posted prices instead of our method, as is usually done, lead to a substantial underestimation of price sensitivity parameters and an overestimation of mark-ups. We also study the effect of price discrimination on manufacturers profits and consumers surplus. We show that if firms did not price discriminate, most of them would be worse off. The gains appear to be larger for luxury brands and French brands that have important market shares. On the consumer side, as the theory suggests, they are winners and losers but price discrimination is moderately welfare enhancing.

We propose a methodology to overcome the problem of unobserved transaction prices in the standard framework of the widely used Berry et al. (1995) demand and supply model. However, we believe that this methodology can have can be easily extended to other demand and supply models. This method has a lot of applications when the data collected by the econometrician is either unreliable or restricted. Our methodology is useful when the econometrician observes the demand for different markets but the prices for only a subset of these market (for example, cars in different European countries or supermarket chains is different municipalities).

The paper is organized as follows. The first section presents the model and explains how to deal with the unobserved transaction prices in the estimation. The application on the French new cars market is developed in the second part of the paper. Section 3 analyzes the importance of discounts and the impact of price discrimination. We conclude in Section 4.

2 Theoretical model

We first present our theoretical model. The approach is identical to the BLP model except that the demand arises from a finite number of heterogenous groups of consumers. The firms are supposed to observe the group of each consumer, as well as their corresponding characteristics, such as their average price elasticity. They then price discriminate among these groups, in order to take advantage of the heterogeneity in preferences.

Specifically, heterogenous consumers are supposed to be segmented in n_D groups of consumers, and we denote by d the group of consumer i. Each consumer chooses either to purchase one of the J products or not to buy any, which is the outside option denoted by 0. As usually, each product is assimilated to the bundle of its characteristics. Consumers are utility maximizers, and the utility of choosing j is supposed to be a linear function of characteristics:

$$U_{ij}^d = X_j \beta_i^d + \xi_j^d + \alpha_i^d p_j^d + \epsilon_{ij}^d,$$

where X_j corresponds to the vector of observed characteristics and ξ_j^d represents the valuation of unobserved characteristics. p_j^d is the price set by the seller for the category d and is not observed by the econometrician. Consumers with characteristics d are supposed to face the same transaction price p_j^d. This is crucial, but not more restrictive than the assumption that ξ_j^d is common to all individuals with characteristics d, shown by Berry & Haile (2010) to be necessary for identifying demand models nonparametrically from aggregated data. In other words, we need to abstract from individual negotiation in this model.

We assume, as usually, that individual parameters can be decomposed linearly into a mean, an individual deviation from the mean and a deviation related to individual characteristics:

$$\begin{cases} \beta_i = \overline{\beta}_0^d + \pi_0^{X,d} E_i + \Sigma_0^{X,d} \zeta_i^X \\ \alpha_i = \overline{\alpha}_0^d + \pi_0^{p,d} E_i + \Sigma_0^{p,d} \zeta_i^p, \end{cases}$$

where E_i denotes demographic characteristics that are unobserved by the firm for each purchaser but whose distribution is common knowledge. $\zeta_i = (\zeta_i^X, \zeta_i^p)$ is a random vector with a specified distribution such as the standard multivariate normal distribution. We suppose, as usually, that the idiosyncratic error terms ϵ_{ij}^d are extreme-value distributed.

The utility function can be expressed as a mean utility and an individual deviation from this mean:

$$U_{ij}^d = \delta_j^d(p_j^d) + \mu_j^d(E_i, \zeta_i, p_j^d) + \epsilon_{ij}^d,$$

with

$$\delta_j^d(p_j^d) = X_j \overline{\beta}_0^d + \overline{\alpha}_0^d p_j^d + \xi_j^d$$

and

$$\mu_j^d(E_i, \zeta_i, p_j^d) = X_j \left(\pi_0^{X,d} E_i + \sigma_0^{X,d} \zeta_i^X \right) + p_j^d \left(\pi_0^{p,d} E_i + \sigma_0^{p,d} \zeta_i^p \right).$$

We let the dependence in p_j^d explicit for reasons that will become clear below. Because of the logistic assumption on the ϵ_{ij}^d, the aggregate market share $s_j^d(p^d)$ of good j for demographic group d satisfies, when prices are set to $p^d = (p_1^d, ..., p_J^d)$,

$$s_j^d(p^d) = \int s_j^d(e, u, p^d) dP_{E,\zeta}^d(e, u), \tag{1}$$

where $P_{E,\zeta}^d$ is the distribution of (E, ζ) for group d and

$$s_j^d(e, u, p^d) = \frac{\exp\left(\delta_j^d(p_j^d) + \mu_j^d(e, u, p_j^d)\right)}{\sum_{k=0}^J \exp\left(\delta_k^d(p_k^d) + \mu_k^d(e, u, p_k^d)\right)}$$

Now, we consider a Nash-Bertrand competition setting where firms are able to price discriminate by setting different prices to each of the n_D consumers groups. Letting $\mathcal{J}_f$ denote the set of products sold by firm f, the profit of f when the vector of all prices for group d is p^d satisfies

$$\pi_j = \sum_{d=1}^{n_D} P(D = d) \sum_{j \in \mathcal{J}_f} s_j^d(p^d) \times \left(p_j^d - c_j^d\right).$$

$s_j^d(p^d)$ is the market share of product j for group d when prices are equal to p^d. c_j^d is the marginal cost of the product j for group d.

The first-order condition for the profit maximization yields:

$$p_f^d = c_f^d + \left(\Omega_f^d\right)^{-1} s_f^d, \tag{2}$$

where p_f^d, c_f^d and s_f^d are respectively the equilibrium transaction prices, marginal costs and observed market shares vectors for firm f. Ω_f^d is the matrix of typical (i, j) term equal to $\frac{\partial s_j}{\partial p_i}(p^d)$. The optimal prices are the consequence of the firms making the traditional arbitrage between increasing prices and lowering sales. When a seller is able to price discriminate, it is less constrained than with a uniform pricing strategy since this arbitrage is made for each group separately. If a group is particularly price sensitive, the seller offers a low price and is still able to extract a lot of surplus for the less price sensitive group by setting a higher price for this group. Note that even if price discrimination is illegal, enforcing such a law may be difficult because consumers can hardly sue firms, at least if they pay less than the posted price. We thus assume that the posted prices $\overline{p} = (\overline{p}_1, ..., \overline{p}_J)$ satisfy

$$\overline{p}_j \geq \max_{d=1...n_D} p_j^d. \tag{3}$$

We shall reinforce this assumption below, for identification purpose.

3 Inference

We now turn to inference on this model. We assume that the econometrician observes the market shares s_j^d corresponding to each consumer group but not the discriminatory prices p_j^d paid by consumers. We do assume, on the other hand, that the econometrician observes the posted prices.

First, let us recall the standard case where the true prices are observed. Let

$$\theta_0^d = (\overline{\beta_0}^d, \overline{\alpha_0}^d, \pi_0^{X,d}, \Sigma_0^{X,d}, \pi_0^{p,d}, \Sigma_0^{p,d})$$

denote the true vector of parameters for group d. The standard approach for identification and estimation of θ_0^d, initiated by BLP, is to use the exogeneity of Z_j, which includes the characteristics X_j and other instruments (typically, function of characteristics of other products or cost shifters) to derive moment conditions involving θ_0^d. The exogeneity condition takes the form

$$\mathbb{E}\left[Z_j \xi_j^d\right] = 0. \tag{4}$$

The idea is then to use the link between ξ_j^d and the true parameters θ_0^d through Equation (1). Specifically, we know from Berry (1994) that for any given vector θ^d, the equation (1), where θ_0^d is replaced by θ^d, defines a bijection between market shares and mean utilities of products δ_j^d. Hence, we can define $\delta_j^d(s^d, p^d; \theta^d)$, where $s^d = (s_1^d, ..., s_J^d)$ denotes the vector of observed market shares. Once $\delta_j^d(s^d, p^d; \theta^d)$ is obtained, the vector $\xi_j^d(p^d, \theta^d)$ of unobserved characteristic corresponding to θ^d and rationalizing the market shares follows easily since

$$\xi_j^d(p^d; \theta^d) = \delta_j^d(s^d, p^d; \theta^d) - X_j \overline{\beta}^d - \overline{\alpha}^d p_j^d.$$

The moment conditions used to identify and estimate θ_0^d are then

$$E\left[Z_j \xi_j^d(p^d;\theta_0^d)\right] = 0. \tag{5}$$

Now let us turn to the case where the true prices are unobserved. First, remark that when observed prices are different from the true prices (for example when posted prices are used instead of transaction prices), the former approach is not valid in general. To see this, consider the simple logit model, where $\pi_0^{X,d}, \Sigma_0^{X,d}, \pi_0^{X,d}$ and $\Sigma_0^{X,d}$ are known to be zero. In this case $\delta_j^d(s^d, p^d;\theta^d)$ takes the simple form

$$\delta_j^d(s^d, p^d;\theta^d) = \ln s_j^d - \ln s_0^d$$

and does not depend on p^d. In this context, using posted prices $\overline{p}$ instead of the true prices amounts to relying on

$$\xi_j^d(\overline{p};\theta^d) = \ln s_j^d - \ln s_0^d - X_j\overline{\beta}^d - \overline{\alpha}^d\overline{p}^d,$$

instead of on $\xi_j^d(p^d;\theta^d)$. If the measurement errors $\overline{p}^d - p_j^d$ on the true prices were classical, then $E\left[Z_j(\overline{p}^d - p_j^d)\right] = 0$ would be credible and we would still have

$$E\left[Z_j \xi_j^d(\overline{p}^d;\theta_0^d)\right] = 0. \tag{6}$$

However, $\overline{p}^d - p_j^d$ is not a standard measurement error, in particular because the true price depends on the characteristics of the good. If for instance a group of consumer values particularly the horsepower of automobiles, powerful cars will be priced higher for this group. Thus, $\overline{p}^d - p_j^d$ will be negatively correlated with horsepower. Because horsepower itself is correlated with the instruments, as required for the instruments to be relevant, in general we will have $E[Z_j(\overline{p}^d - p_j^d)] \neq 0$, and (6) no longer holds. In the general random coefficient model, this issue also arises but on top of that, $\delta_j^d(s^d, p^d;\theta^d)$ generally depends on p^d. Thus, we also commit an error at this stage that is likely to make (6) invalid, even if the measurement error $\overline{p}^d - p_j^d$ were independent of p_j^d.

Instead of simply replacing p^d by $\overline{p}$, we use the supply model and reasonable identifying conditions on marginal costs and posted prices to recover, given a set of demand parameters θ^d, the corresponding transaction prices $p^d(\theta^d)$. Once these transaction prices are recovered, we can use the standard BLP method to compute $\xi_j^d(p^d(\theta^d);\theta^d)$ and then the moment conditions $E[Z_j\xi_j^d(p^d(\theta^d);\theta^d)]$, to check whether they are equal to zero or not.

The first identifying condition we impose is that the marginal cost of a product is identical for all buyers, so that $c_f^d = c_f$ for all d and f. This amounts to neglecting differences in the costs of selling to different consumers in the total cost of a product. This assumption is likely to be satisfied in many settings, such as the automobile market, where most costs stem from producing, not selling the goods. Second, we suppose that firms post the higher discriminatory price and then offer some discounts according to observable characteristics of buyers in order to reach optimal discriminatory prices. In other words, instead of having simply the inequality (3), we impose

$$\overline{p}_j = \max_{d=1...n_D} p_j^d.$$

This implies that, for each product j, there is a group $\overline{d}_j$, called the pivot group hereafter, that pays the posted price, $p_j^{\overline{d}_j} = \overline{p}_j$. This assumption is necessary since otherwise, we could shift all discounts by an arbitrary constant. It is also in line with empirical evidence on the automobile market (a survey conducted in 1990 in the UK reveals that 17% of car purchasers paid the posted price).[2] Note however that the pivot group is neither supposed to be known ex ante nor constant across different products.

Under these two additional restrictions, we have, by the first-order condition of the Nash-Bertrand equilibrium,

$$\overline{p}_j = c_j + \max_{d=1...n_D} \left[\left(\Omega_f^d\right)^{-1} s_f^d \right]_j,$$

where $[.]_j$ indicate that we consider the j-th line of the vector only. Then the discriminatory prices satisfy

$$p_j^d = \overline{p}_j - \max_{d=1...n_D} \left[\left(\Omega_f^d\right)^{-1} s_f^d \right]_j + \left[\left(\Omega_f^d\right)^{-1} s_f^d \right]_j. \tag{7}$$

Hence, the discriminatory prices are identified up to Ω_f^d. Now, using (1), we obtain, after some algebra,

$$\frac{\partial s_j^d}{\partial p_j^d}(p^d) = \int \left(\overline{\alpha}_0^d + \pi_0^{p,d} e + \Sigma_0^{p,d} u^p \right) s_j^d(e,u,p^d)(1 - s_j^d(e,u,p^d)) dP_{E,\zeta}^d(e,u) \tag{8}$$

We get a similar expression for $\partial s_j^d / \partial p_l^d(p^d)$. This shows that Ω_f^d depends only on the parameters θ_0^d and on the δ_j^d, through $s_j^d(e,u,p^d)$. Besides, we still can define, for a given set of parameters and transaction prices, the vector of mean utilities $\delta_j^d(s^d, p^d; \theta^d)$. Hence, to obtain the discriminatory prices for a given vector of parameter $\theta = (\theta^1, ..., \theta^{n_D})$, we have to solve a system of non-linear equations in $\delta = \delta^1, \ldots, \delta^{n_D}$ and the transaction prices $p = (p^1, \ldots, p^{n_D})$. We use, for that purpose, the following iterative procedure:

1. Start from $p^{d,0} = \bar{p}$ for all groups.

[2] See the UK Competition Competition report : *New cars: A report on the supply of new motor cars within the UK* (2000).

2. Given the current vector of transaction prices p^d, compute $\delta^d = \delta(s^d, p^d; \theta^d)$. We can use for that purpose the contraction mapping suggested by BLP.

3. Given the current vector of mean utilities, compute the corresponding matrix Ω_f^d and update the transaction prices, using (7).

4. Iterate 2 and 3 until convergence of prices and mean utilities.

It seems difficult to prove that the nonlinear system of equations mentioned above admits a unique solution, and that the previous procedure should necessarily converge. In practice, however, we always obtained convergence and did not face any dependence on the initial choice $p^{d,0}$ of p^d.

Finally, we can apply GMM to identify and estimate $\theta_0 = (\theta_0^1, ..., \theta_0^{n_D})$. Let

$$M_J^d(\theta) = \frac{1}{J}\sum_{j=1}^{J} Z_j \left(\delta_j^d(\theta) - X_j\beta^d - \alpha^d p_j^d(\theta)\right)$$

denote the empirical counterpart to the moment conditions in (5). Let also $M_J(\theta) = (M_J^1(\theta)', ..., M_J^{n_D}(\theta)')'$ and define

$$Q_J(\theta) = M_J(\theta)' W_J M_J(\theta),$$

where W_J is a positive definite matrix. Our GMM estimator of θ is

$$\widehat{\theta} = \arg\min_{\theta} Q_J(\theta).$$

It is also possible to take advantage of the structure imposed on the supply side to construct an additional set of moments. Let X^s be the vector of cost shifters. As usually, we specify the marginal cost as log-linear :

$$\ln(c_j) = X_j^s\gamma + \omega_j,$$

where ω_j stands for the unobserved cost shock. It is possible to recover c_j from any group d using the transaction price and the theoretical margin equation (2):

$$c_j = p_j^d - \left[\left(\Omega_f^d\right)^{-1} s_f^d\right]_j.$$

We can also recover $\omega_j(\theta, \gamma)$:

$$\omega_j = \ln\left(p_j^d - \left[\left(\Omega_f^d\right)^{-1} s_f^d\right]_j\right) - X_j^s\gamma$$

and construct the supply moment condition :

$$M_J^s(\theta,\gamma) = \frac{1}{J}\sum_{j=1}^{J} Z_j \left[\ln\left(p_j^d - \left[\left(\Omega_f^d\right)^{-1} s_f^d\right]_j\right) - X_j^s\gamma\right]$$

Then we can proceed as previously, simply replacing $M_J(\theta)$ by $M_J(\theta, \gamma) = (M_J^1(\theta, \gamma)', ..., M_J^{n_D}(\theta,\gamma)', M_J^s(\theta,\gamma))'$.

3.1 Particular cases

We investigate here two restricted versions of the general model where the identification of transaction prices together with demand parameters is obvious and the estimation procedure is simplified. The first model is the random coefficient model without random coefficient on the price variable (partial random coefficients model) and the second is the nested logit model.

3.1.1 The partial random coefficients model

In this model, we consider that all consumers inside a given demographic group have the same price sensitivity, $\alpha_i^d = \bar{\alpha}^d$. We can apply the same method as previously, by first recovering the transaction prices and then compute the corresponing error term to yield the moment conditions. The main difference with previously is that in this specification, the $\mu_j^d(e, u, p_j^d)$ function does not depend on the transaction price p_j^d anymore. As a result, $\delta(s^d, p^d; \theta^d)$ does not depend either on p^d. In this case, instead of solving jointly for δ and the transaction prices p, we can solve for both sequentially. We first compute, for a given set of parameters θ^d, the mean utilities $\delta(s^d, p^d; \theta^d)$. Then we compute the corresponding Ω_f^d and transaction prices. Thus, the computational cost is significantly reduced compared to the general method.

3.1.2 The nested logit case

The nested logit model is very frequently used to model the demand for differentiated goods. Its popularity comes from its simplicity to implement, the market share inversion is linear and does not rely on the use of the contraction mapping, which decreases the computational intensity. Yet, the nested logit does not impose too much restriction on substitution patterns, contrary to the logit model. Thanks to the nested structure, (products are segmented into homogeneous groups), we obtain more substitution inside groups than across groups of products.

However, when we introduce unobserved price discrimination, the model is no longer linear and must be estimated using the GMM approach. But the estimation method is simpler than the full random coefficients and partial random coefficients models presented before. The same remark applies of course to the simple logit model.

In the nested logit approach, consumers with same characteristics d are supposed to have homogeneous preferences, so that the term $\mu_j^d(e, u, p_j^d)$ is actually zero. On the other hand, in order to obtain still realistic substitution patterns, the error terms corresponding to products of a same segment g are assumed to be correlated through a parameter denoted σ^d. Aggregate market shares then take the simple form

$$\ln s_j^d / s_0^d = \delta_j^d + \sigma^d \ln s_{j/g}^d,$$

where $s_{j/g}^d$ denotes the market share of j inside segment g for group d. We also have the simple expression

$$\frac{\partial s_j^d}{\partial p_j^d} = \alpha^d s_j^d \left(s_j^d + \frac{\sigma^d}{1-\sigma^d} s_{j/g}^d - \frac{1}{1-\sigma^d} \right).$$

As a result, Ω_f^d, and therefore the transaction prices, do not depend on δ. The transaction prices only depend on the observed market shares and on the parameters $(\alpha^1, \sigma^1, ..., \alpha^{n_D}, \sigma^{n_D})$. Denoting them by $p_j^d(\alpha^1, ..., \sigma^{n_D})$, we obtain the moment equations

$$\mathbb{E}\left[Z_j \left(\ln \frac{s_j^d}{s_0^d} - X_j \beta^d - \alpha^d p_j^d(\alpha^1, ..., \sigma^{n_D}) - \sigma^d \ln s_{j/g}^d \right) \right] = 0.$$

This moment equation is far simpler to solve than in the general model because (i) each term in the left hand side is easy to compute and (ii) one has to optimize on $(\alpha^1, \sigma^1, ..., \alpha^{n_D}, \sigma^{n_D})$ only, since once they are fixed, β^d can easily be obtained by 2SLS on each group of consumers. This reduces the dimension on which to optimize to $2 \times n_D$ parameters.

4 Application to the French new car market

4.1 Description of the data

We apply our methodology to estimate demand and transactions prices in the new automobile industry using a dataset of the association of French automobile manufacturers (CCFA, Comité des Constructeurs Français d'Automobiles) that records all the registrations of new cars bought by households in France between 2003 and 2008. Each year, we observe a sample of about two million vehicles. For each registration, the following attributes of the car are reported: brand, model, fuel energy, car-body style, number of doors, horsepower, CO_2 emissions, cylinder capacity and weight. These characteristics have been complemented with fuel prices, so as to compute the cost of driving (in euros for 100 kilometers). Automobile sellers are well known to price discriminate, negotiate or to offer discounts to close the deal. As in the setting, we only observe posted prices that come from manufacturers catalogues.

We now turn to the construction of the consumer groups that we suppose firms use to price discriminate. Apart from cars attributes, the date of the registration and some characteristics of the owner are provided in the CCFA database, in particular his municipality of residence and his age. The age (or the age class) is presumably a strong determinant of purchase, and is easily observable by a seller even if he does not know the buyer before the transaction. We therefore suppose that it is used by the automobile makers to price discriminate. The income is also likely to affect preferences for different car attributes and price sensitivity. The income is however likely to be unobserved by the seller but instead inferred from the municipality the buyer lives in and the age class. We compute a predictor of buyer's income, namely the median household income in his age class and in his municipality using data from the French national institute of statistics (Insee).[3] It seems reasonable to assume that the seller does not have a far better prediction of the buyer's income in such anonymous market, where buyers and sellers do not know each other before the transaction. It is crucial for our approach that buyers cannot lie about their individual characteristics[4] and we believe that buyers have high incentive to buy a new car at a close dealer especially for the after-sale services. At the end, we define groups of buyers by interacting three age classes and two income classes.[5] The six groups and their corresponding frequency in the data are presented in Table 1.

Group	Characteristics	Frequency
1	Age $<$ 40, income $<$27,000	15.7%
2	Age $<$ 40, income $>$27,000	11.5%
3	Age $\in$ [40,59], income $<$27,000	16.3%
4	Age $\in$ [40,59], income $>$27,000	22.3%
5	Age $\geq$ 60, income $<$27,000	20.8%
6	Age $\geq$ 60, income $>$27,000	13.2%

Table 1: Definition of the groups of consumers and frequency

When defining the groups of consumers, we face a trade-off between realism (it is likely

[3] There are over 36,000 municipalities in France. Paris as well as the second and third biggest cities are split in smaller units (''arrondissement''). There is significant heterogeneity in the distribution of median incomes across municipalities.

[4] We do not consider that sellers discriminate according to the gender because if women (or men) are known to get higher discount it would be easy to send the husband (or wife), a relative or a friend to buy the car.

[5] We use another definition of demographic groups using six age classes. The results, not reported here, are similar.

that firms discriminate along several dimensions) and accuracy of the observed proportion of sales $\widehat{s}_j^d$ as estimators of the true market shares s_j^d. Contrary to linear models where standard measurement errors on outcomes are not problematic, here measurement error can bias the results because by Jensen's inequality, $E\left(\ln \widehat{s}_j^d\right) < \ln E(\widehat{s}_j^d) = \ln s_j^d$. Moreover, one can show that the relative bias increases as s_j^d goes to zero, implying that this measurement error becomes more problematic with a large number of groups. Finally, even if in theory $s_j^d > 0$ so that $\ln s_j^d$ is always well defined, we may observe $\widehat{s}_j^d = 0$, in which case $\ln \widehat{s}_j^d$ is not defined. The six groups that we consider are large enough to avoid in most cases this zero sale issue. When it occurs, however, rather than discarding the products, we replace $\widehat{s}_j^d$ by a predictor that minimizes the bias : $s_j^d = \frac{n_j^d + 0.5}{N_d}$, n_j^d denoting the number of sales of product j in group d and N_d the number of potential buyers with characteristics d.[6]

We define a product as a brand, model, segment, car-body style and fuel type. A total of 3205 products for the six years is obtained.[7] Table 2 presents the average characteristics of new cars purchased for each group of consumers. We find significant heterogeneity. The average price of vehicles increases with the income. Purchases of medium age class are, in average, more expensive. The chosen cars are also bigger (with a higher weight) and with higher horsepower. Young purchasers are more interested in smaller cars (lighter and/or with three doors) whereas station-wagons are more popular among the medium age class. The higher age group of consumers purchases lighter vehicles than medium age classes, but these vehicles are in average less fuel efficient (with higher fuel costs).

	Consumer group	Price	Fuel cost	HP	Weight	Three doors	Station Wagon
1	A $<$ 40, I $<$27,000	19,803	6.2	5.7	1,182	19.0%	9.7%
2	A $<$ 40, I $>$27,000	20,911	6.5	6.0	1,221	16.8%	12.9%
3	A $\in$ [40,59], I $<$27,000	21,521	6.5	6.1	1,231	14.3%	12.7%
4	A $\in$ [40,59], I $>$27,000	21,739	6.8	6.2	1,236	14.8%	13.1%
5	A $\geq$ 60, I $<$27,000	20,117	6.9	5.9	1,194	11.4%	8.9%
6	A $\geq$ 60, I $>$27,000	20,831	7.0	6.0	1,219	10.9%	10.5%

Lecture notes : A represents the age class and I the income class. Prices are in constant (2008) euros, fuel cost is the cost of driving 100 kilometers, in constant (2008) euros, HP stands for horsepower, weight is in kilograms.

Table 2: Average characteristics of new cars purchased across groups of consumers

[6] Gandhi et al. (2013) propose a method to deal with null market based on the use of inequality moments, however we believe that it is beyond the scope of the paper.

[7] See Table 13 in Appendix A.1 for the fraction of products with null market shares.

The dataset does not contain any information on the distribution network and the distribution sector is not modeled in this application. We make the traditional assumption that manufacturers and their dealers are perfectly integrated. Another related assumption is that there is no variation in competition across geographic market. In reality, a lot of dealers are independent and choose their own pricing and discounting strategy. Furthermore, dealers are not uniformly located on the French territory so there may exist some difference in competition intensity across geographic market. With sales at the dealers level, or the locations of dealers, we would be able to take into account heterogeneity of pricing strategy or competition intensity (see, e.g., Nurski & Verboven, 2012). Due to a lack of such available data, we abstract from these issues afterwards.

4.2 Estimation results

We present the estimations of different models. We estimate the standard logit model and the logit model with unobserved price discrimination. For the estimation of these two models, we do not use supply moment conditions. Then we estimate the standard BLP model and the BLP model with unobserved price discrimination. All the models are estimated using the GMM approach. For the logit models, we use only demand-side moment conditions while we use the supply-side moment conditions for the random coefficients models. For all specifications, we control for the main characteristics of the cars such as horsepower, weight, the cost of driving 100 kilometers (computed by introducing annual average fuel prices) in the demand function. We also introduce dummies for station-wagon car-body style and three doors. Finally, we introduce year and brand dummies that are constrained to be identical for all demographic groups. For the BLP models (standard and with unobserved price discrimination), we allow for unobserved heterogeneity of preferences inside groups of consumers for price, horsepower, fuel cost, weight and for the utility of holding a car (represented by the intercept). We constrain the heterogeneity parameters to be constant across groups.[8] In the marginal cost equation, we use horsepower, fuel consumption (in liter for 100 kilometers) and weight as cost-shifters. We also introduce brand dummies to control for manufacturer's specific unobserved quality of cars.

We construct demand and supply moment conditions using instrumental variables. In addition to exogenous characteristics we construct three sets of instruments. The first is composed by the sums of continuous characteristics (namely weight, horsepower and fuel cost) of other brands' products, the second is the sums of these characteristics over other

[8]This constraint helps to identify the parameters of heterogeneity. We observe 6 years and 6 groups of consumers which gives a total of 36 markets.

brands' products of the same segment, supposed to be closer substitutes. The third set of instrumental variables is the sums of these characteristics of the other products of the brand belonging to the same segment.[9]

The implementation of the estimation method uses the nested fixed point to inverse the market share, using a tight tolerance value of (10^{-12}). To compute the transaction prices, we also use a contraction mapping and set the tolerance to (10^{-6}) (it is equivalent to a cent of euro). To approximate the market share function, we prepare 1000 Halton normal draws. Specifically, we draw a set of 1000 individuals for each demographic group and each year. We use the simplex minimization algorithm. We carefully investigate potential convergence issues, as suggest Knittel & Metaxoglou (2014), using different starting values.

First of all, we compare the estimates of the model with unobserved price discrimination to the standard model for the simple logit case (columns (1) and (2)) and random coefficients model (columns(3) and (4)). In the standard models, we assume that the posted prices are paid by all consumers. For the standard BLP model, the supply-side model also differs from our model since it assumes that the posted prices are the optimal non-discriminatory prices, given the heterogeneous preferences of the different groups of consumers. The results for the different models are presented in Tables 3 and 4. The estimated parameters are generally similar for the logit models and the BLP models. In the logit case, the estimations are barely affected when we use our methodology or when we use the posted prices. However, for the BLP models, the estimation results imply much lower price sensitivity parameters with unobserved price discrimination. The less price sensitive group is however always the same, namely the group of old consumers with high income. The deviation of the price sensitivity parameter is higher for the BLP model with unobserved prices than for the standard model (1.09 versus 0.70). The intercept is negative, reflecting the fact that the major part of consumers choose the outside option (not to buy a car or buy it on the second-hand market). There is a significant heterogeneity in the utility of holding a car across demographic groups, the young purchasers being the ones with the highest utility. The mean parameters of preference for horsepower surprisingly appear to be negative in the standard BLP specification but become positive for the full model with unobserved price discrimination. The standard deviation from the means is also much higher for the BLP model (4.26) than for the model with price discrimination (-0.23). As expected, all groups of consumers dislike large fuel expenses. The parameters of sensitivity to the fuel cost are consistent with the sensitivity to the car price parameters. The old consumers with high income appear to be also the less sensitive to the cost of driving while the more sensitive

[9]See the segmentation used in Table 15, Appendix A.1.

consumers are also the young and middle-age groups with a low income. As weight is a proxy for the size and the space of the car, it is positively valuated by all the consumers. Three doors and station-wagon vehicles are negatively valuated, reflecting that most of the consumers buy sedan cars with five doors (four doors plus the trunk). The cost equation parameters have the predicted signs : the marginal cost of production is increasing in the horsepower and weight while it is less costly to produce inefficient cars (i.e with higher fuel consumption).

	Standard Logit		Logit + unobs. prices		Standard BLP		BLP + unobs. prices	
	Parameter	Std-err	Parameter	Std-err	Parameter	Std-err	Parameter	Std-err
Price sensitivity								
Age $<$ 40, I = L	-2.72**	0.21	-2.72**	0.21	-2.8**	0.241	-4.88**	0.395
Age $<$ 40, I = H	-2.54**	0.2	-2.54**	0.2	-2.72**	0.259	-3.92**	0.286
Age $\in$ [40,59], I = L	-2.22**	0.2	-2.22**	0.2	-2.16**	0.224	-4.3**	0.359
Age $\in$ [40,59], I = H	-2.07**	0.2	-2.07**	0.2	-2.15**	0.214	-4.04**	0.359
Age $\leq$ 60, I = L	-2.02**	0.2	-2.02**	0.2	-2.36**	0.229	-4.09**	0.372
Age $\leq$ 60, I = H	-1.88**	0.19	-1.89**	0.19	-1.91**	0.232	-3.46**	0.237
std deviation (σ^p)					0.7**	0.089	1.09**	0.081
Intercept								
Age $<$ 40, I = L	-7.77**	0.28	-8.21**	0.33	-4.13**	0.301	-5.87**	0.48
Age $<$ 40, I = H	-8.48**	0.27	-8.83**	0.32	-5.29**	0.301	-6.18**	0.338
Age $\in$ [40,59], I = L	-8.43**	0.26	-8.6**	0.31	-5.07**	0.298	-6.58**	0.43
Age $\in$ [40,59], I = H	-8.49**	0.26	-8.58**	0.31	-5.28**	0.295	-6.82**	0.416
Age $\leq$ 60, I = L	-8.05**	0.25	-8.11**	0.3	-5.04**	0.289	-6.19**	0.418
Age $\leq$ 60, I = H	-8.4**	0.25	-8.39**	0.25	-5.33**	0.322	-6.61**	0.341
std deviation (σ^x)					-0.18	0.111	-0.24	0.173
Horsepower								
Age $<$ 40, I = L	5.77**	0.56	5.77**	0.56	-5.68**	0.904	8.1**	0.952
Age $<$ 40, I = H	5.06**	0.54	5.05**	0.54	-5.58**	0.992	4.42**	0.432
Age $\in$ [40,59], I = L	4.14**	0.53	4.13**	0.53	-6.98**	0.876	6.41**	0.821
Age $\in$ [40,59], I = H	3.59**	0.52	3.59**	0.51	-6.97**	0.822	5.66**	0.768
Age $\leq$ 60, I = L	3.04**	0.52	3.05**	0.52	-7.12**	0.939	4.93**	0.783
Age $\leq$ 60, I = H	2.67**	0.5	2.69**	0.5	-8.39**	0.955	0.68†	0.407
std deviation (σ^x)					4.26**	0.244	-0.23	0.244
Fuel cost								
Age $<$ 40, I = L	-6.73**	0.27	-6.73**	0.27	-5.77**	0.291	-8.01**	0.653
Age $<$ 40, I = H	-5.86**	0.26	-5.85**	0.26	-4.96**	0.296	-5.83**	0.558
Age $\in$ [40,59], I = L	-5.83**	0.25	-5.83**	0.25	-4.9**	0.278	-7.18**	0.598
Age $\in$ [40,59], I = H	-4.94**	0.25	-4.94**	0.25	-3.88**	0.272	-6.03**	0.576
Age $\leq$ 60, I = L	-4.23**	0.25	-4.23**	0.25	-3.12**	0.275	-5.3**	0.579
Age $\leq$ 60, I = H	-3.61**	0.24	-3.62**	0.24	-2.16**	0.278	-2.88**	0.512
std deviation (σ^x)					-1.25**	0.155	0.65	0.699
Weight								
Age $<$ 40, I = L	4.8**	0.4	4.8**	0.4	5.92**	0.352	5.76**	0.615
Age $<$ 40, I = H	4.97**	0.39	4.96**	0.39	6.22**	0.364	4.86**	0.384
Age $\in$ [40,59], I = L	4.88**	0.38	4.88**	0.38	5.9**	0.358	5.77**	0.54
Age $\in$ [40,59], I = H	4.67**	0.38	4.67**	0.38	5.92**	0.334	5.61**	0.532
Age $\leq$ 60, I = L	3.68**	0.38	3.69**	0.38	5.07**	0.349	4.56**	0.527
Age $\leq$ 60, I = H	3.74**	0.36	3.75**	0.36	4.94**	0.369	4.84**	0.3
std deviation (σ^x)					-0.19	0.418	0.17	0.572

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

Table 3: Estimation of parameters : logit model, logit model with unobserved price discrimination, standard BLP model and BLP model with unobserved price discrimination

	Standard Logit		Logit + unobs. prices		Standard BLP		BLP + unobs. prices	
	Parameter	Std-err	Parameter	Std-err	Parameter	Std-err	Parameter	Std-err
Three doors								
Age < 40, I = L	0.00	0.12	0.00	0.12	0.25*	0.12	-0.23	0.165
Age < 40, I = H	-0.14	0.12	-0.14	0.12	0.03	0.12	-0.14	0.126
Age ∈ [40,59], I = L	-0.16	0.12	-0.15	0.12	0.12	0.115	-0.36*	0.149
Age ∈ [40,59], I = H	-0.29*	0.12	-0.29*	0.12	-0.06	0.117	-0.5**	0.146
Age ≤ 60, I = L	-0.61**	0.12	-0.61**	0.12	-0.46**	0.113	-0.8**	0.142
Age ≤ 60, I = H	-0.67**	0.12	-0.67**	0.12	-0.41**	0.112	-0.63**	0.114
Station-Wagon								
Age < 40, I = L	-0.72**	0.08	-0.71**	0.08	-0.79**	0.082	-0.84**	0.118
Age < 40, I = H	-0.58**	0.08	-0.58**	0.08	-0.65**	0.081	-0.59**	0.09
Age ∈ [40,59], I = L	-0.63**	0.08	-0.63**	0.08	-0.7**	0.08	-0.74**	0.103
Age ∈ [40,59], I = H	-0.68**	0.08	-0.68**	0.08	-0.79**	0.082	-0.81**	0.1
Age ≤ 60, I = L	-0.70**	0.08	-0.7**	0.08	-0.8**	0.081	-0.83**	0.096
Age ≤ 60, I = H	-0.68**	0.08	-0.68**	0.08	-0.78**	0.081	-0.83**	0.08
Marginal cost equation								
Intercept					-0.79**	0.069	-0.63**	0.173
Horsepower					0.52**	0.027	0.26	0.244
Fuel economy					-2.89**	0.356	-1.42*	0.699
Weight					1.06**	0.044	1.01†	0.572
Value of objective function	4213		4220		1601		1668	
Number of observations	19,230		19,230		22,435		22,435	

Table 4: Estimation of parameters : logit model, logit model with unobserved price discrimination, standard BLP model and BLP model with unobserved price discrimination (continued)

4.3 Comparison with the standard models

We compare the results implied by the standard model and the model with unobserved price discrimination in the case of the random coefficient. The results of the comparison of the models in the logit case is displayed in Appendix A.3.

Table 5 focuses on demand parameters and implied price elasticities for the standard model and the model with unobserved price discrimination. Price elasticities are higher, in absolute value for the model with unobserved prices than for the standard one. However, this hides two opposite effects. On one side, parameters of price sensitivity are higher (in absolute value) for the model with unobserved price discrimination, as suggested in Table 3. However, when posted prices are used instead of transaction prices, we overestimate prices and so are price elasticities. At the end, the underestimating of price sensitivity parameters is more important and implies smaller (in absolute value) price elasticities.[10]

[10]For the logit case, parameters of price sensitivities are estimated to be identical in the two models. At the end we overestimate price elasticities when we estimate the standard model instead of the model with unobserved discounts (we get an average of -4.6 with the standard logit model versus an average of -4.4

In our model we find average price elasticities varying from is -4.5 to -6.4, which are in the range of those obtained by BLP, who report elasticities between -3.5 and -6.5, but below those of Langer (2012) who finds, when using transaction prices, a range between -6.4 to -17.8.

	Full model		Standard BLP	
Group of consumers	α^d	ε^d	α^d	ε^d
Age $<$ 40, I $=$ L	-4.88 (0.4)	-6.44	-2.8 (0.24)	-4.32
Age $<$ 40, I $=$ H	-3.92 (0.29)	-5.14	-2.72 (0.26)	-4.41
Age $\in$ [40,59], I $=$ L	-4.3 (0.36)	-5.9	-2.16 (0.22)	-3.38
Age $\in$ [40,59], I $=$ H	-4.04 (0.36)	-5.62	-2.15 (0.21)	-3.33
Age $\leq$ 60, I $=$ L	-4.09 (0.37)	-5.3	-2.36 (0.23)	-3.48
Age $\leq$ 60, I $=$ H	-3.46 (0.24)	-4.51	-1.91 (0.23)	-2.69
Std deviation (σ^p)	1.09 (0.08)		0.7 (0.09)	
Average	-4.13	-5.51	-2.21	-3.56

Table 5: Comparison of demand parameters and average price-elasticities for the standard BLP model and the model with unobserved price discrimination

Table 6 displays average mark-ups for the different groups of consumers for the BLP model with unobserved price discrimination and the standard BLP model[11]. We compute average mark-ups weighting by actual sales in each group but also using the same weighting scheme for all groups of consumers, namely, the overall market shares ("basket-weighted" method). This allows to eliminate the demand composition effect. Average mark-ups for the same artificial basket of cars are between 16.6%, for young people with low income and 24.2% for old people with high income. The mark-ups corresponding to the standard BLP model are much higher and the average mark-up reaches 32% compared to 20% implied by the model with unobserved price discrimination. This shows that taking into account transaction prices in the model may have important consequences for the validity of counter-factual simulations.

with unobserved price discrimination. See Table 16 in Appendix A.3.

[11]See Table 17 in Appendix A.3 for the comparison of mark-ups in the logit case.

	Full model		Standard BLP
Group of consumers	Sales-weighted	Basket-weighted	Basket-weighted
Age $<$ 40, I $=$ L	16.9	16.6	31.9
Age $<$ 40, I $=$ H	20.9	21.1	31.2
Age $\in$ [40,59], I $=$ L	18.6	18.9	31.3
Age $\in$ [40,59], I $=$ H	19.5	19.8	31.4
Age $\leq$ 60, I $=$ L	20.8	20.33	33.2
Age $\leq$ 60, I $=$ H	24.2	24.2	32.7
Average	20.0	20.0	31.9

Reading notes: the "basket-weighted" discounts and mark-ups are obtained by using the same artificial basket of cars for all groups.

Table 6: Comparison of implied average mark-ups for the standard BLP model and the model with unobserved price discrimination (in %)

Finally, we perform non-nested tests to compare the standard model that assumes uniform pricing on the supply side and the model with unobserved price discrimination using the Rivers-Vuong test based on the lack-of-fit criteria (see Vuong, 1989 and Rivers & Vuong, 2002). This test use the estimated marginal costs implied by the two alternative models and cost shifters. Specifically, it allows to find which marginal cost has the best statistical fit given the observed cost shifters. It has been used a lot in differentiated markets to compare two alternative competition models (see Bonnet & Dubois, 2010 and Ferrari & Verboven, 2012 for examples). We start by estimating the marginal cost equation by assuming:

$$C_{jt}^h = exp(w_j^h + W_{jt}'\lambda^h)\eta_{jt}^h,$$

where C_{jt}^h represents the estimated marginal cost implied by model h. w_j^h represents the product specific cost, W_{jt} are the cost shifters and η_{jt}^h the unobserved cost shocks. We transform the previous expression using the logarithm :

$$\ln C_{jt}^h = w_j^h + W_{jt}'\lambda^h + \ln \eta_{jt}^h.$$

Assuming $\mathbb{E}(\ln \eta_{jt}^h | w_j^h, W_{jt}) = 0$, we estimate λ^h using least squares:

$$\min_{\lambda^h, w_j^h} \frac{1}{n} \sum \left(\ln C_{jt}^h - w_j^h + W_{jt}'\lambda^h\right)^2$$

In addition to those used in the estimation to construct supply-side moment conditions, cost shifters include prices of raw materials involved in the production process such as rubber, petrol, platina and aluminium. These prices are interacted with the weight of the car[12].

The statistic test use the values of the lack-of-fit criterion:

$$Q_n^h = \frac{1}{n} \sum \left(\ln C_{jt}^h - w_j^h + W_{jt}' \lambda^h \right)^2$$

The null hypothesis is that the two models are asymptotically equivalent while the first alternative hypothesis (H1) is that the model with price discrimination (*pd*) is asymptotically better than the model with uniform pricing (*up*) and the second alternative hypothesis (H1') is that the model with uniform pricing (*up*) is asymptotically better than the model with price discrimination (*pd*). Specifically :

$$\begin{aligned}
\text{H0 :} \quad & \lim_{n=+\infty} \left\{ Q_n^{pd}(\bar{\lambda}^{pd}, \bar{w}_j^{pd}) - Q_n^{up}(\bar{\lambda}^{up}, \bar{w}_j^{up}) \right\} = 0 \\
\text{H1 :} \quad & \lim_{n=+\infty} \left\{ Q_n^{pd}(\bar{\lambda}^{pd}, \bar{w}_j^{pd}) - Q_n^{up}(\bar{\lambda}^{up}, \bar{w}_j^{up}) \right\} < 0 \\
\text{H1':} \quad & \lim_{n=+\infty} \left\{ Q_n^{pd}(\bar{\lambda}^{pd}, \bar{w}_j^{pd}) - Q_n^{up}(\bar{\lambda}^{up}, \bar{w}_j^{up}) \right\} > 0
\end{aligned}$$

We compute Rivers and Vuong test statistic :

$$T_n = \frac{\sqrt{n}}{\hat{\sigma}_n} \left(Q_n^2(\bar{\lambda}^{pd}, \bar{w}_j^{pd}) - Q_n^1(\bar{\lambda}^{up}, \bar{w}_j^{up}) \right),$$

which is asymptotically a standard normal variable under the null. In our data, we reject the null hypothesis in favor of H1, meaning that the price discrimination model fits better our data than the uniform pricing model.

5 Importance of price discrimination

5.1 Discounts

Table 7 presents the average discounts for all groups of consumers in the BLP model. As before, we compute average discounts weighted by actual sales and using the "basket" weights. The results with both weighting methods are however similar. As expected, the pattern on average discounts across groups is similar to the one on price elasticity. The estimated pivot group (the group assumed to be paying the posted price) is identical for 99.7% of the products and corresponds to the group with lower price elasticity. These are the 15.7% people over 60 year old, and with income above 27,000 euros.[13] We find a

[12] These raw material price indexes are published by CCFA to analyze the evolution of production costs.

[13] For the logit model with unobserved price discrimination, the pivot group is always the group of old with high income purchasers.

different pivot group for only 9 cars in 2003, namely the young consumers with high income. This result is in line with those of Harless & Hoffer (2002), who show using observed dealer margins that older buyers generate higher vehicle gross profit.

On average, the sales-weighted discount is 5.2%, with a substantial heterogeneity across groups of consumers.[14] Clearly, income and age are both important determinants of the discount obtained. On average, young purchasers with a low income pay their car 9% less than the posted price whereas young with a high income gets only 3.8% discount. These percentages represent a gross gain of almost 2,000 euros for the low income group and only slightly above 800 for the high income group. Old purchasers get very small discounts, only 4.9% (around 1,000 euros) in average for the group with low income. Middle age consumers get large discounts (6.7% for the low income group and 5.6% for the high income group), which represents a gross savings of 1,200 to 1,400 euros. Note that we assumed that the posted prices are equal to the discriminatory prices for the pivot group, or, in other words that we normalize their discount to zero. This assumption partly explains why we find relatively small discounts.

In 2000, the UK Competition Commission investigated the competitiveness of the UK new car market and gathered data on average discounts by brand and segment. The dataset is very reliable since it was collected directly from dealers. The report (*New cars: A report on the supply of new motor cars within the UK* (2000)) reveals that the average discount lies between 7.5% and 8%, which is above our estimate of 5.2%. For our estimate to be consistent, it implies that the average discount associated to the pivot group is roughly slightly below 3% which is very plausible.

[14] For the logit case, we obtain an average of 4.2% See Table 18 in Appendix A.3.

	Average discount (in % of posted price)		Average gross discount (in euros)	
Group of consumers	Sales-weighted	Basket-weighted	Sales-weighted	Basket-weighted
Age $<$ 40, I = L	9.04	9.16	1,785	1,936
Age $<$ 40, I = H	3.81	3.95	844	872
Age $\in$ [40,59], I = L	6.7	6.55	1,516	1,425
Age $\in$ [40,59], I = H	5.55	5.49	1,309	1,213
Age $\leq$ 60, I = L	4.87	4.88	1,007	1,076
Age $\leq$ 60, I = H	0.01	0.01	6	10
Average	5.2	5.2	1,127	1,132

Reading notes: the "basket-weighted" discounts and mark-ups are obtained by using the same artificial basket of cars for all groups.

Table 7: Average discounts by groups of consumers

Table 8 presents average discount for the different segments, as well as the market shares associated. We can note that the luxury cars are those who offer the highest discount in average (7.8% for Executive cars and 5.2% for Allroad vehicles). These cars are very expensive and have to decrease their price by an important amount to sell to the more price sensitive groups of consumers. One might expect luxury brands not to offer discounts as part of their brand image, but our model does not incorporate this dimension. Moreover, our model is static and thus firms offer discounts as long as the effect of increasing sales compensate the effect of lowering the price. All the other segments have quite homogenous average discounts, roughly between 4% and 5%.[15]

Segment	Share	Discount
Supermini	44.48%	4.73%
Executive	1.02%	7.8%
Small Family	15.8%	4.14%
Large Family	7.36%	4.4%
Small MPV	18.5%	4.26%
Large MPV	0.86%	4.64%
Sports	6.53%	4.43%
Allroad	5.43%	5.19%

Table 8: Average discounts by segment (sales-weighted)

Figure 1 presents the distribution of discounts for all sales. 98% of the discounts lie between

[15]For the average discount by segment in the logit model, see Table 19 of Appendix A.3.

0 and 30%, and almost 16% of the cars were sold without discounts (purchases from the pivot groups or with a discount lower than 1%). The distribution rapidly reaches its mode, the value of 5%, and the frequency then decreases more slowly. Very few cars are sold with more than 15% discount.[16]

The UK Competition Commission report refers to a consumer survey conducted in 1995 which asked automobile purchasers whether or not they obtained a discount over the posted price. This survey reveals that 17% of purchasers paid the posted price whereas 37% bargained and obtained a discount and 29% were automatically offered a discount. This figure of 17% is very close to our estimation of 16% of the cars are sold with less than 1% discount. Furthermore, the fact that some purchasers were "automatically offered a discount" corroborates our assumption that discounts are used as a tool to price discrimination because the posted price is not optimal for some consumers.

[16]Similar histogram for the logit model is displayed in Figure A.3 of Appendix A.3.

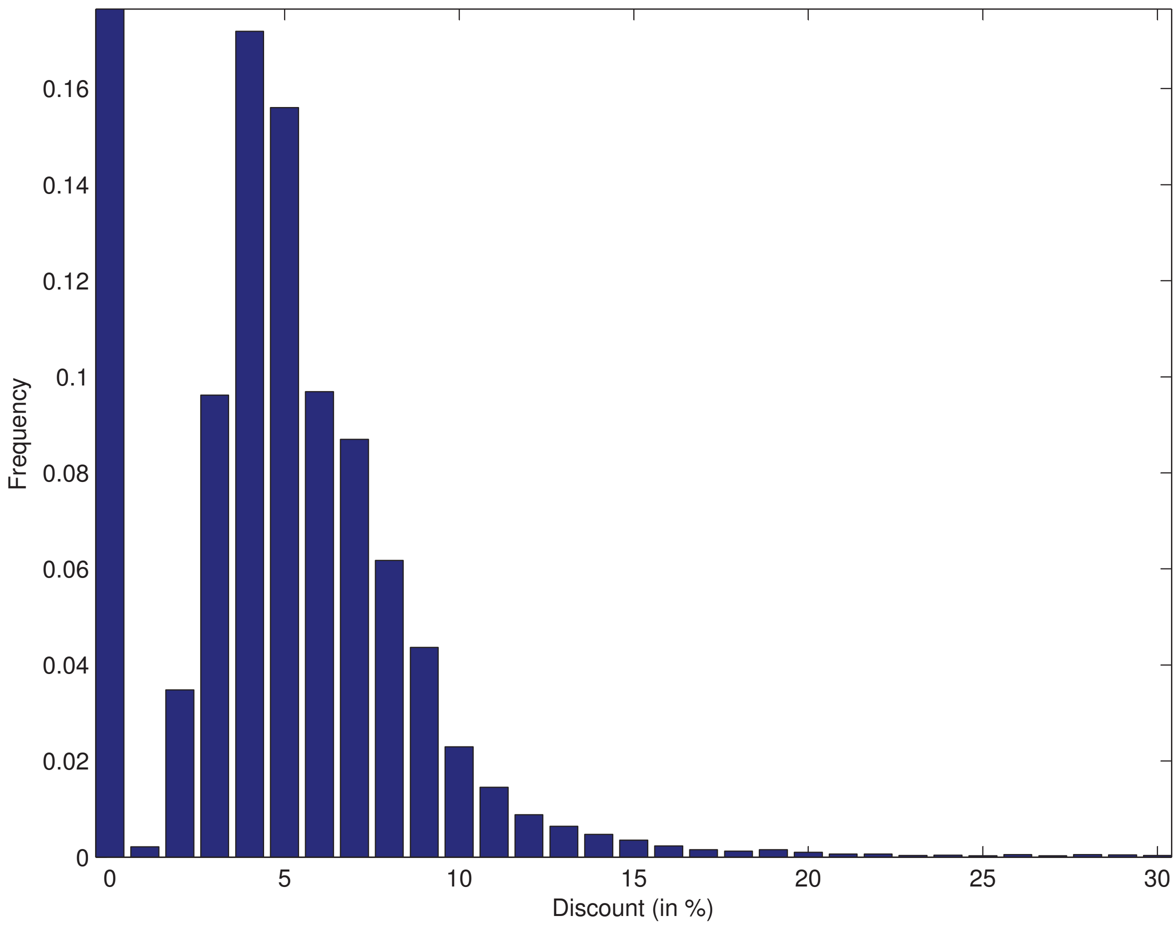

Figure 1: Distribution of estimated discounts

We could not find precise and reliable data on transaction prices in the French automobile market. Indeed, it is considered to be a strategic and confidential information for car dealers.[17] Specialized automobile magazines, online forum discussions and websites providing guidance and tips to purchase a new car always refer to the capacity of buyers to negotiate and obtain a discount. For example, when searching the keywords *how much discount for new car* with google search engine (in French), the second website referenced is states that *discount are generally between 5% and 20%.*[18]. The third website associated to the same key words search is a forum answering the question of how much discount one can expect

[17]A first anecdote is that some manufacturers claim that they have to survey their own dealers to get the information. The second evidence is that, on April 2014, an article from *Reuters* dealing with the importance of discounts due to the economic crisis in the European car market uses some figures obtained from *a confidential internal industry survey*.

[18]See `http://www.choisir-sa-voiture.com/concessionnaire/meilleur-prix-voiture.php`.

to obtain on the purchase of a new car. One reply states that it should be expected at most 20% discount while another one says that it is in average 6%.[19]. Our estimations are overall consistent with these anecdotal evidences.

A recent study by Kaul et al. (2012) investigates the effect of the scrapping policy on the magnitude of discounts in Germany using data collected to a sample of dealers. The average discount they obtain is 14%, which is much higher than our estimates. Once again, the assumption that the posted price is equal to the transaction price party explains this difference. Moreover, their study concerns the period 2007-2010, which corresponds to the beginning of the economic crisis and it is very likely that car dealers reacted to the bad economic climate by reducing their margins and increasing the discounts.

We are more cautious about comparing our estimates with data from the US market which has a very different organization. In the US, dealers are independent from the manufacturers and negotiation seems to be more common (see for example Morton et al. (2003) and Busse et al. (2012)). Moreover, it is documented that manufacturers offering rebates is a very common practice in the US market and is more comparable to our interpretation of discounts. Consumers are usually aware of these rebates[20]. In their study of the effect of scrapping policy on manufacturers rebates, Busse et al. (2012) report that 49% of cars were sold with a manufacturer rebate and it equals 2,700 dollars in average, which roughly represents 10% of the transaction price. This study provides evidence of heterogeneity in rebates according to manufacturers, the national ones such as General Motors and Chrysler offering higher than average rebates[21].

They are few papers that correlate the magnitude of discounts to individual characteristics. Using very precise data on dealer margins in Canada, Chandra et al. (2013) investigate the evolution of price discrimination against women in the car industry. In the US market, Harless & Hoffer (2002) using similar data and Langer (2012) using survey data on transaction prices also investigates the correlation between car prices, gender and some other demographic characteristics. All these studies report a significant heterogeneity in discounts with respect to the age and find that the dealers margins are the highest for the older age class, while the discounts are significantly higher for young purchasers. In the Appendix of her paper, Langer (2012) documents significant price discrimination with

[19]See `http://forum.hardware.fr/hfr/Discussions/Auto-Moto/negocier-voiture-concession-sujet_15899_1.htm`.

[20]The Kelley Blue Book website provides information on manufacturer's suggested retail prices (MRSP), rebates and target prices.

[21]See the regression Table 7 of the rebate on manufacturers dummies.

respect to the income, the high income groups of consumers (both men and women) are associated to higher margins.

5.2 The impact of price discrimination on firms and consumers

If price discrimination is always profitable for a monopoly, this may not be the case in an oligopoly, because price discrimination may reinforce competition among firms. At the end, all firms may be worse-off than they would be if they could commit to a uniform pricing strategy (Holmes, 1989, Corts, 1998). The effect on consumers is also ambiguous. Even for a given group of consumers, some products may be cheaper without price discrimination, and other more expensive. We investigate in this subsection the effect of price discrimination on firms and consumers. For that purpose, we compute, using our estimates of the model, the counterfactual prices and profits that would occur if firms could commit to set equal prices for all groups of consumers.

Results on firm profits are displayed in Table 5.2. Gains from price discrimination are rather small but heterogenous. We observe that if price discrimination is profitable for 68% of manufacturers, it makes 12 out of the 38 manufacturers worse off. The gains associated to price discrimination are particularly high for luxury brands such as Mercedes (+3.9%), Porsche (+7.2%) and Jaguar (+5.2%). Price discrimination appears to be also profitable for the major French manufacturers (+1.8%, +1.2% and +2.1% for respectively Renault, Peugeot and Citroen) but not profitable for Dacia (-1.5%). The total gain from price discrimination is rather small but significant, the industry profit increases by 1.2% compared to the uniform pricing equilibrium.

Brand	Profit with price discrimination (in M€)	Profit without price discrimination (in M€)	Gain from discrimination (in %)
Renault	644.64	633.48	1.8%
Peugeot	550.74	544.19	1.2%
Citroen	451.84	442.44	2.1%
Volkswagen	181.37	181.72	-0.2%
Toyota	164.98	163.37	1%
Mercedes	153.05	147.36	3.9%
Ford	135.06	133.69	1%
B.M.W.	112.85	111.01	1.7%
Opel	110.28	110.34	-0.1%
Audi	91.87	92	-0.1%
Fiat	64.75	64.55	0.3%
Dacia	58.61	59.52	-1.5%
Suzuki	54.64	54.26	0.7%
Seat	53.9	53.97	-0.1%
Nissan	50.57	50.28	0.6%
Mini	35.57	35.79	-0.6%
Honda	30.84	30.44	1.3%
Hyundai	29.94	29.85	0.3%
Skoda	23.48	23.53	-0.2%
Mazda	19.77	19.77	0%
Alfa Romeo	18.87	18.92	-0.2%
Kia	18.32	18.27	0.3%
Land Rover	16.88	16.66	1.3%
Smart	11.51	11.49	0.1%
Mitsubishi	10.54	10.49	0.5%
Porsche	9.66	9.01	7.2%
Jeep	7.36	7.36	0%
Chrysler	6.54	6.54	0%
Lancia	4.99	4.93	1.1%
Saab	4.46	4.4	1.2%
Dodge	3.41	3.44	-0.8%
Daewoo	3.36	3.35	0.4%
Jaguar	3.02	2.87	5.2%
Ssangyong	2.01	2.05	-1.8%
Daihatsu	1.99	1.97	0.9%
Subaru	1.98	2.01	-1.1%
Lexus	1.63	1.61	1.3%
Rover	0.05	0.05	0.6%
Total Industry	3145	3107	1.2%

Reading notes: Profits are annual profits, for the year 2007, in million euros. The gains from price discrimination represent the profit gains or losses of switching from the uniform pricing equilibrium to the equilibrium with price discrimination.

Table 9: Gains and losses from price discrimination by brand.

We also investigate the effect of price discrimination on consumers. In Table 10, we compute the average price differences between the uniform and the discriminatory prices for each group of consumers and report the number of products for which the discriminatory price is lower than the uniform one (see column 2). We can see that, except for the more sensitive group (the young consumers with low income), all groups have some products that become cheaper in the uniform pricing equilibrium. The middle-age groups experience significant gains with price discrimination (average gains are 270 and 59 euros for respectively low and high income groups). The highest average price difference between the two regimes is for the pivot group (old and high income) and it reaches around 1,100 euros.

			Average gains in purchases	
Group of consumers	Frequency	$\#\{j : p_j^d < p_j^{\text{uniform}}\}$	Sales-weighted	Basket-weighted
Age $<$ 40, I $=$ L	15.7	3,205	673	781
Age $<$ 40, I $=$ H	11.5	185	-326	-283
Age $\in$ [40,59], I $=$ L	16.3	3,107	262	270
Age $\in$ [40,59], I $=$ H	22.3	2,057	71	59
Age $\leq$ 60, I $=$ L	20.8	1,011	-49	-79
Age $\leq$ 60, I $=$ H	13.2	8	-1,072	-1,145

Reading notes: the third column indicates how much products (among the 3,205) have lower prices with the price discrimination regime. The "basket-weighted" discounts and mark-ups are obtained by using the same artificial basket of cars for all groups.

Table 10: Gains of price discrimination for groups of consumers.

In Table 11, we explore the effect of price discrimination on the total sales for each group of consumers. As expected, for some groups, pricer discrimination increases the total sales while it decreases for other groups. The overall effect is moderate but positive : price discrimination increases total sales by 1.2%.

	Price discrimination	Uniform pricing	Gain from discrimination
Age $<$ 40, I = L	110,945	94,619	14.7%
Age $<$ 40, I = H	95,006	98,940	-4.1%
Age $\in$ [40,59], I = L	104,788	100,501	4.1%
Age $\in$ [40,59], I = H	222,094	219,141	1.3%
Age $\leq$ 60, I = L	155,699	155,202	0.3%
Age $\leq$ 60, I = H	147,975	165,033	-11.5%
Total	836,508	826,226	1.2%

Table 11: Comparison of total number of purchases for groups of consumers with price discrimination and uniform pricing.

Finally, we measure the effect of price discrimination on consumers surplus. Overall, price discrimination is moderately benefic for consumers as it increases average individual surplus by 0.25%. Again, this moderate positive global effect hides some winners and some losers. The group that experience the highest welfare gain is the group of middle-age consumers with high income (+3.4%) and the group associated with the highest surplus loss is, not surprisingly, the pivot group (-1.9%).

	Price discrimination	Uniform pricing	Gain from discrimination
Age $<$ 40, I = L	13,738	13,276	3.37%
Age $<$ 40, I = H	19,748	19,871	-0.63%
Age $\in$ [40,59], I = L	15,970	15,768	1.26%
Age $\in$ [40,59], I = H	21,207	21,120	0.41%
Age $\leq$ 60, I = L	17,780	17,777	0.02%
Age $\leq$ 60, I = H	25,912	26,407	-1.91%
Average	18879	18833	0.25%

Table 12: Comparison of average individual surplus for the different groups of consumers with price discrimination and uniform pricing.

6 Conclusion

This paper investigates the recurrent problem of using posted prices instead of transaction prices in structural models of demand and supply in markets with differentiated products. We propose an approach based on the inclusion of some unobserved price discrimination by firms based on observable individual characteristics. This approach requires to have data on aggregate sales on the corresponding groups of purchasers and, as usually, characteristics of products. We use this model to describe the French new car market where price discrimination may occur through discounts. Our results suggest significant discounting by manufacturers. They appear to be very consistent with previous studies on price dispersions, dealers data collected by the UK Competition Commission and anecdotic evidences on the magnitude of discount in the French market.

References

Allen, J., Clark, R. & Houde, J.-F. (2014), Search frictions and market power in negotiated price markets, Technical report, National Bureau of Economic Research.

Ayres, I. & Siegelman, P. (1995), 'Race and gender discrimination in bargaining for a new car', *The American Economic Review* pp. 304–321.

Berry, S. (1994), 'Estimating discrete-choice models of product differentiation ', *The RAND Journal of Economics* **25**, 242–262.

Berry, S. & Haile, P. (2010), Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers.

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

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

Busse, M. R., Knittel, C. R., Silva-Risso, J. & Zettelmeyer, F. (2012), Did 'cash for clunkers' deliver? the consumer effects of the car allowance rebate system, Technical report, November 2012. Working Paper.

Chandra, A., Gulati, S. & Sallee, J. (2013), 'The gender gap in price negotiations: Evidence from new car sales'.

Charles, K., Hurst, E. & Stephens, M. (2008), 'Rates for vehicle loans: race and loan source', *The American Economic Review* **98**(2), 315–320.

Corts, K. S. (1998), 'Third-degree price discrimination in oligopoly: All-out competition and strategic', *The RAND Journal of Economics* **29**, 306–323.

Ferrari, S. & Verboven, F. (2012), 'Vertical control of a distribution network : an empirical analysis of magazines', *The RAND Journal of Economics* **43**(1), 26–50.

Gandhi, A., Lu, Z. & Shi, X. (2013), Estimating demand for differentiated products with error in market shares, Technical report, cemmap working paper, Centre for Microdata Methods and Practice.

Goldberg, P. (1996), 'Dealer price discrimination in new car purchases: Evidence from the consumer expenditure survey', *Journal of Political Economy* pp. 622–654.

Harless, D. & Hoffer, G. (2002), 'Do women pay more for new vehicles? evidence from transaction price data', *The American Economic Review* **92**(1), 270–279.

Holmes, T. J. (1989), 'The effects of third-degree price discrimination in oligopoly', *American Economic Review* **79**, 244–250.

Kaul, A., Pfeifer, G. & Witte, S. (2012), 'The incidence of cash for clunkers: An analysis of the 2009 car scrappage scheme in germany'.

Knittel, C. R. & Metaxoglou, K. (2014), 'Estimation of random-coefficient demand models: Two empiricists' perspective', *Review of Economics and Statistics* **96**(1), 34–59.

Langer, A. (2012), 'Demographic preferences and price discrimination in new vehicle sales'.

Morton, F., Zettelmeyer, F. & Silva-Risso, J. (2003), 'Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities?', *Quantitative Marketing and Economics* **1**(1), 65–92.

New cars: A report on the supply of new motor cars within the UK (2000), Technical report, UK Competition Commission, http://webarchive.nationalarchives.gov.uk/20111202195250/http://competition-commission.org.uk/rep_pub/reports /2000/439cars.htm#summary.

Nurski, L. & Verboven, F. (2012), 'Exclusive dealing as a barrier to entry? evidence from automobiles'.

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

Vuong, Q. H. (1989), 'Likelihood ratio tests for model selection and non-nested hypotheses', *Econometrica: Journal of the Econometric Society* pp. 307–333.

A Appendix

A.1 Additional descriptive statistics

Group	Characteristics	Frequency of null sale
1	Age $<$ 40, Income $<$27,000	11.6%
2	Age $<$ 40, Income $>$27,000	10.3%
3	Age $\in$ [40,59], Income $<$27,000	7.5%
4	Age $\in$ [40,59], Income $>$27,000	4%
5	Age $\geq$ 60, Income $<$27,000	7.8%
6	Age $\geq$ 60, Income $>$27,000	7.6%

Table 13: Fraction of product with null market shares in the final sample

A.2 Additional tables

Group	**Make**	Supermini	Small family	Large family	Executive	Sports car	Small MPV	Large MPV	Allroad/SUV
PSA	*Citroen*	C1, C2, C3, Saxo	Xsara	C5	C6	-	Berlingo, C4, Nemo, Xsara	C8	C-Crosser
	Peugeot	106, 107,1007,206, 207	306, 307, 308	406, 407	607	-	Bipper, Partner	807	4007
Renault	*Renault*	Clio, Modus, Twingo	Megane	Laguna	Vel Satis	-	Kangoo, Megane	Espace	Koleos
	Dacia	Sandero	Logan	-	-	-	-	-	-
B.M.W	*B.M.W*	-	1-Series	3-Series	5, 6, 7-Series	Z4	-	-	X3, X5, X6
	Mini	Mini	-	-	-	-	-	-	-
Chrysler	*Chrysler*	-	-	Sebring	300C, 300M, Crossfire	-	PT Cruiser	Voyager, G.Voyager	-
	Jeep	-	-	-	-	-	-	-	Compass, Cherokee, Commander, G.Cherokee, Wrangler
	Dodge	-	Caliber	Journey	Viper	-	-	-	Durango, Nitro
Daihatsu	*Daihatsu*	Cuore, Sirion, YRV	-	-	-	Copen	-	-	Terios
Daimler	*Mercedes*	-	A-Class	C, CLK-Class	E, CL, R, S, SL, CLS, SLR-Class	SLK-Class	B-Class, Vaneo	Viano	G, GL, GLK, ML-Class
	Smart	Fortwo, Forfour	-	-	-	Coupe, Roadster	-	-	-
Fiat	*Alfa Romeo*	Mito	147	156, 159, GT	166, Brera	GTV, Spider	-	-	-
	Fiat	500, Palio, Panda, Punto, Seicento	Bravo, Stilo	Croma	-	Barchetta	Doblo, Fiorino, Idea, Multipla	Ulysse	Sedici
	Lancia	Y	-	Lybra	Thesis	-	Musa	Phedra	-
Ford	*Ford*	Fiesta, Ka,	-	Mondeo	-	Puma	Focus, Fusion, T.Connect, Tourneo	Galaxy, S-Max	Kuga
	Jaguar	-	-	X-Type	S-Type, XJ, XK	-	-	-	-
	Land Rover	-	-	-	-	-	-	-	Freelander, Defender, Discovery, R.Rover
	Volvo	-	C30, V50	C70, S40, S60, V70	V40, S80	-	-	-	XC60, XC70, XC90
GM Europe	*Chevrolet*	Kalos, Matiz	Aveo, Lacetti, Nubira	Epica, Evanda	-	Corvette	Rezzo, Tacuma	-	Captiva, Tahoe
	Daewoo	Kalos, Matiz	Lanos	-	Evanda	-	Rezzo	-	Korando
	Opel	Corsa	Astra	Insigna, Signum, Vectra	Omega	Tigra, Speedster	Agila, Combo, Meriva, Zafira	-	Antara, Frontera
	Saab	-	-	9-3	9-5	-	-	-	-
Honda	*Honda*	Jazz	Civic	Accord	-	S2000	FR-V, Stream	-	CR-V, HR-V
Hyundai	*Hyundai*	Atos, Getz, I10	Accent, Coupe, I30	Elantra, Sonata	-	-	Matrix	Trajet	Tucson, Santafe, Terracan
	Kia	Picanto, Rio	Cee-d, Cerato	Magentis	-	-	Carens, Soul	Carnival	Sorento, Sportage
Lada	*Lada*	-	111, 112	-	-	-	-	-	Niva
Mazda	*Mazda*	2	3	6	RX8	MX5	5, Premacy	MPV	-
Mitsubishi	*Mitsubishi*	Colt	Lancer	Carisma	-	-	Spacestar	Grandis	Outlander, Pajero
Nissan	*Nissan*	Micra, Note	Almera, Qashqai	Primera	350Z, Maxima-Q	-	Almera	-	X-Trail, Murano, Pathfinder, Patrol, Terrano
Porsche	*Porsche*	-	-	-	911, Boxter, Cayman	-	-	-	Cayenne
Rover	*Rover*	25, Streetwise	45	75	-	-	-	-	-
Ssangyong	*Ssangyong*	-	-	-	-	-	-	Rodius, Stavic	Actyon, Korando, Kyron, Rexton
Subaru	*Subaru*	Justy	Impreza	Legacy	-	-	-	-	Forester, B9Tribeca
Suzuki	*Suzuki*	Alto, Ignis, Splash, Swift, SX4	Liana	-	-	-	Wagon-R	-	G. Vitara, Jimny, Samurai, Vitara
Toyota	*Toyota*	Aygo, IQ, Yaris	Auris	Avensis, Prius	-	Celica, MR	Corolla	Previa	RAV4, L.Cruiser
	Lexus	-	-	IS	GS, LS	-	-	-	RX
VW Group	*Audi*	A2	A3	A4, A5	A6, A8, R8	S3, S4, S6, S8 TT	-	-	Allroad, Q5, Q7
	Seat	Arosa, Ibiza	Cordoba, Leon	Toledo	-	-	Altea	Alhambra	-
	Skoda	Fabia	-	Octavia, Superb	-	-	Roomster	-	-
	Volkswagen	Fox, Lupo, Polo	Eos, Golf, Jetta, New-beetle	Scirocco, Passat	-	Phaeton	Caddy, Touran	Sharan	Tiguan, Touareg

Table 15: Segmentation of the automobile market

A.3 Results with the logit specification

Group of consumers	With unobs. prices	Standard Logit
Age $<$ 40, I = L	-4.91	-5.37
Age $<$ 40, I = H	-4.96	-5.33
Age $\in$ [40,59], I = L	-4.58	-4.77
Age $\in$ [40,59], I = H	-4.4	-4.51
Age $\leq$ 60, I = L	-3.98	-4.04
Age $\leq$ 60, I = H	-3.94	-3.92
Average	-4.4	-4.6

Table 16: Comparison of demand parameters and average price-elasticities for the standard Logit model and the logit model with unobserved price discrimination

	Logit with unobs. prices		Standard logit
Group of consumers	Sales-weighted	Basket-weighted	Basket-weighted
Age $<$ 40, I = L	23.74	22.96	25.76
Age $<$ 40, I = H	23.81	24.11	24.82
Age $\in$ [40,59], I = L	26.02	26.58	24.79
Age $\in$ [40,59], I = H	27.3	27.94	24.84
Age $\leq$ 60, I = L	29.25	28.42	26.49
Age $\leq$ 60, I = H	29.78	29.8	25.78
Average	26.81	26.79	25.39

Reading notes: the "basket-weighted" discounts and mark-ups are obtained by using the same artificial basket of cars for all groups.

Table 17: Comparison of implied average mark-ups for the standard logit model and the logit model with unobserved price discrimination (in %)

	Average discount (in % of posted price)		Average gross discount (in euros)	
Group of consumers	Sales-weighted	Basket-weighted	Sales-weighted	Basket-weighted
Age $<$ 40, I $=$ L	9.53	9.3	1673	1690
Age $<$ 40, I $=$ H	7.75	7.87	1419	1430
Age $\in$ [40,59], I $=$ L	4.47	4.62	829	838
Age $\in$ [40,59], I $=$ H	2.66	2.72	492	494
Age $\leq$ 60, I $=$ L	2.1	2.04	373	370
Age $\leq$ 60, I $=$ H	0	0	0	0
Average	4.15	4.15	748	753

Reading notes: the “basket-weighted” discounts and mark-ups are obtained by using the same artificial basket of cars for all groups.

Table 18: Average discounts by groups of consumers for the logit model

Segment	Freq.	Discount
Supermini	44.48%	5.36%
Executive	1.02%	1.5%
Small Family	15.8%	3.74%
Large Family	7.36%	2.46%
Small MPV	18.5%	3.36%
Large MPV	0.86%	2.19%
Sports	6.53%	3%
Allroad	5.43%	2.12%

Table 19: Average discounts by segment (sales-weighted) for the logit model

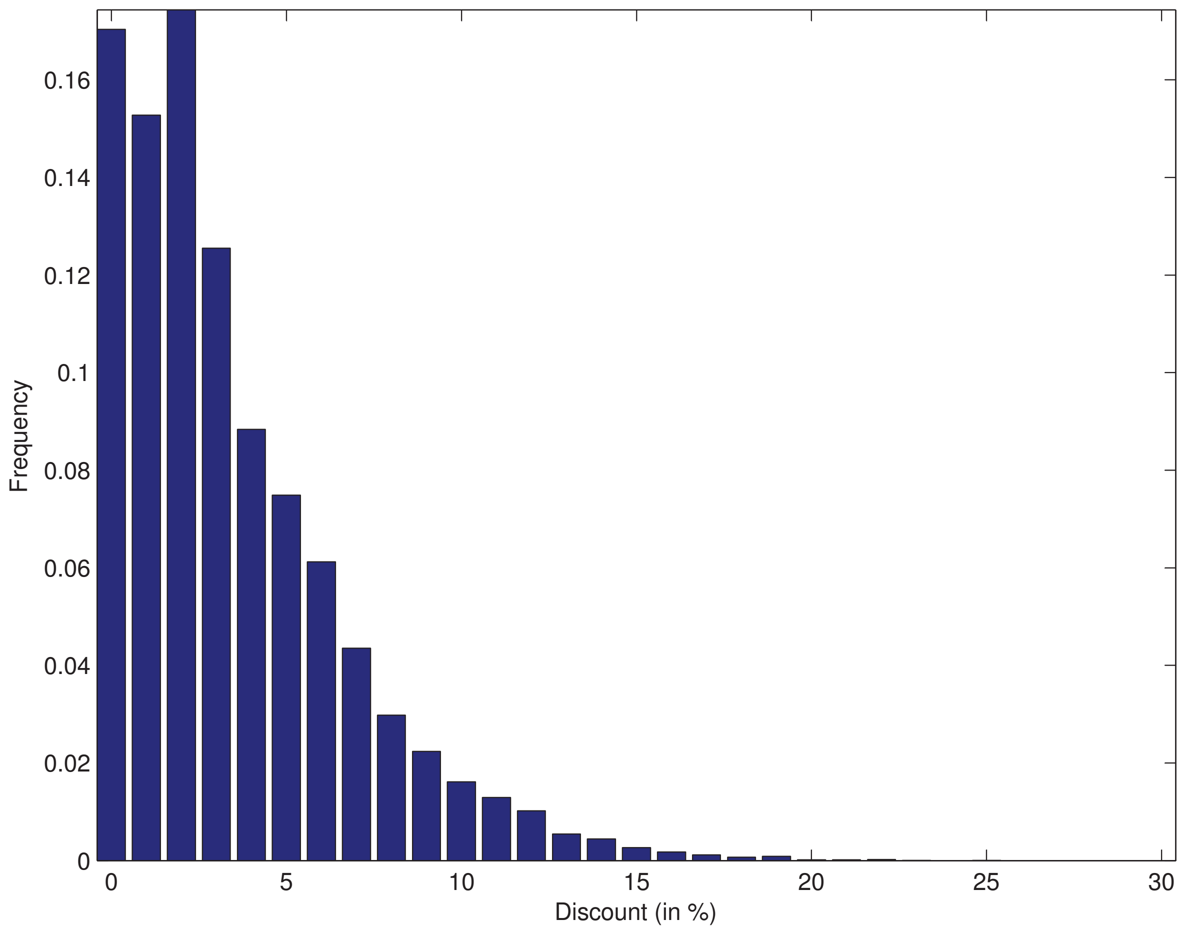

Figure 2: Distribution of estimated discounts for the logit model