



KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and
Applied Economics

Department of Economics

Automobile engine variants and price discrimination

by

Øyvind THOMASSEN

Econometrics

Center for Economic Studies
Discussions Paper Series (DPS) 10.15
<http://www.econ.kuleuven.be/ces/discussionpapers/default.htm>



April 2010

**DISCUSSION
PAPER**

Automobile engine variants and price discrimination

Øyvind Thomassen*
Katholieke Universiteit Leuven

April 15, 2010

Abstract

Using a structural model of demand for automobile engine variants, this paper finds that there is second-degree price discrimination: markups increase with engine size. Still, average markups are lower than when models have just one engine. The paper develops the first empirical demand framework suitable for markets with variants. There is an unobserved product characteristic and a consumer-specific logit term for classes of products, but both are fixed across variants. Fixed effects control for unobservables. The literature's assumption of orthogonality between unobserved and observed product characteristics is not needed. JEL: L11, L62, C25

Keywords: second-degree price discrimination, automobiles, discrete-choice demand models

*E-mail: oyvind.thomassen@econ.kuleuven.be. This is a revised version of chapter 1 of my 2009 Oxford DPhil thesis. Thanks to my thesis supervisor Howard Smith, examiners Ian Crawford and Philipp Schmidt-Dengler, and Frank Verboven, Johannes Van Biesebroeck and Laura Grigolon for comments.

1 Introduction

It is a widespread practice for firms to offer different qualities of what are otherwise the same products. iPhone 3GS has 16GB and 32GB storage variants. Internet service providers offer a choice of bandwidths. Airlines and trains have different travel classes. Car models come with several engine options. The existence of quality variants is often explained by price discrimination. In an early mention of the issue, Dupuit (1849) points out that railway companies let their third-class accommodation be “open carriages with wooden benches (...) to prevent the passengers who can pay the second-class fare from traveling third-class”.¹ Low-price variants have lower quality, so that consumers with higher willingness to pay will self-select into higher qualities. Mussa and Rosen (1978) show that a monopolist can charge higher markups by sorting consumers in this way.

More recent theories extend the basic monopoly model to include competition. Take a duopoly with horizontally differentiated firms, each offering two qualities. Markups are positive, because each firm has a comparative advantage in its horizontal location. But firms have no comparative advantage in supplying quality. Armstrong and Vickers (2001) and Rochet and Stole (2002) show that depending on the intensity of competition firms may no longer discriminate and instead charge the same markups for both qualities.² Let the duopolists play a game where they can offer one or two qualities. Offering two qualities may enable a firm to price discriminate. But it also increases competition, by moving closer to the preferences of some of the other firm’s customers.³ In this way, the outcome (‘one’,‘one’) may have higher profits than (‘two’,‘two’). Furthermore, if ‘two’ is a dominant strategy, the game is a prisoner’s dilemma. In this case we will observe quality variants

¹Quoted by Tirole (1988).

²Verboven (1999) finds constant markups for any intensity of competition. See discussion in Stole (2007) pp. 2274-6.

³This is particularly clear if firm A’s actions are to offer either low quality or both and firm B’s to offer either high or both.

and discrimination, but firms would be better off if they could commit to restrict quality ranges.⁴

The theoretical models provide useful benchmark results. But the results depend on specific assumptions (the degree of differentiation, symmetry). To understand quality-based price discrimination, it is therefore useful to let the data determine these specifics. This paper is an empirical study of the central questions from the theory literature discussed in the previous paragraph: is there quality-based price discrimination, and do the gains from discrimination outweigh the cost of increased competition at the industry level? The contribution of the paper is to address these questions by developing and estimating a new structural demand framework that is the first to accommodate quality variants, and to have the relevant theory models as special cases. The new demand model is a natural extension of existing empirical models, along the lines of the theoretical models. It has broad applications, in that it is suitable for horizontally differentiated and quantity variants as well as quality variants (e.g. fat levels of a milk brand, or small and large boxes of breakfast cereal). In particular, I ask whether there is price discrimination across automobile engine variants, and whether industry profits would be higher if car models had only one engine variant.

My results show that there is discrimination, but that it is not sufficient to raise average markups relative to a situation where car models have only one engine size. Markups are computed using a standard equilibrium assumption. Within car models, markups increase in engine size. In a counterfactual equilibrium where car models are restricted to have only one engine variant, average markups are slightly higher than in the multi-variant case. The benefits of discrimination do not offset the increase in competition from a large number of variants. Profits still rise slightly, however, since total sales go up because of increased product variety.

⁴See Gilbert and Matutes (1993) for this prisoner's dilemma. Ellison (2005) pp. 596-7 finds that both restricted and full quality ranges may be equilibria. Corts (1998) finds a similar prisoner's dilemma result for third-degree price discrimination: discriminating may be a dominant strategy leading to lower equilibrium profits than uniform pricing.

A natural point of departure for a demand model is the differentiated-products demand literature starting with Berry, Levinsohn, and Pakes (1995) (BLP). The utility a consumer gets from a product is a linear function of observable product characteristics. In addition, a product-specific constant captures unobserved product characteristics that affect consumers in the same way. Finally, utility includes a random shock that is iid across products for a given consumer, and across consumers. The shock represents likes and dislikes that are idiosyncratic in the sense that they are not correlated across products.

These assumptions are not entirely appropriate for markets with variants. My demand framework is designed to accommodate the fact that quality variants of products differ only with respect to observable characteristics (e.g. horsepower or hard-drive storage). The effect of observable product characteristics is fully accounted for by directly including these characteristics in the utility function. Since variants differ only by observable characteristics, a difference in utility between two variants must be fully explained by the terms involving the observable characteristics. Letting either the taste shock or the unobserved product characteristic vary across variants of a product is therefore not justified. The model in this paper holds both unobserved characteristics and taste shocks fixed across variants.

While my demand model is different, I maintain the substantive assumptions of the recent automobile demand literature⁵: there are unobserved differences and idiosyncratic tastes for models. The difference is in how I deal with the additional information in my data (engine variants).⁶ As in many markets, it is plausible that consumers have real idiosyncratic tastes for models. When this is the case, removing the iid shock altogether, as in Bajari and Benkard (2005) and Berry and Pakes (2007), would underestimate product differen-

⁵BLP, Petrin (2002), Berry, Levinsohn, and Pakes (2004).

⁶Of course the literature abstracts away from the variant issue: each consumer chooses a car *model* whose price and engine characteristics are chosen by the researcher from among the model's range of engine variants. Since the demand models depend on correctly estimating the coefficients on price and engine characteristics in utility, this may be a problematic simplification.

tiation. On the other hand, Petrin (2002) points out the idiosyncratic term introduces a new product characteristic for each new product. This may overestimate the benefits of increasing product variety. But in this paper shocks enter only at the model level. Therefore, when I compare equilibria with one and many variants per model, the number of product characteristics does not change.

A final point about the demand model is that it permits identification with weaker assumptions than in the literature because of the richer data. Goldberg (1995) restricts the unobserved product characteristic to be the same across car models of a given brand. Brand-level fixed effects control for unobservable characteristics. BLP relax this assumption by including model-specific dummies to control for unobservables at the model level. But these dummies capture the mean effect of observed characteristics as well as unobserved characteristics. To separate the two effects, BLP need to assume orthogonality between unobserved and observed characteristics. Any model that allows for alternative-specific unobservables needs this orthogonality assumption.⁷ However, the assumption is problematic because both types of product characteristics are choice variables of manufacturers.⁸ I relax both the orthogonality assumption of BLP and Goldberg’s restriction that unobservables are fixed across models of a brand. Since unobservables do not vary across variants, I can control for unobservables with model-level fixed effects. Price and product characteristics vary within models, so that their effects are separately identified.

In the empirical literature on price discrimination, McManus (2007) and Cohen (2008) are similar to this paper, in that they estimate structural demand models to decompose costs and markups, and thereby draw conclusions about second-degree price discrimination (in quantities of specialty coffee and paper

⁷BLP micro (Berry, Levinsohn, and Pakes 2004) in principle need it too, but instead use calibration to find the unidentified parameters.

⁸E.g. Akerberg, Benkard, Berry, and Pakes (2007) p. 4197 say: “There are plausible reasons to believe that product characteristics themselves are correlated with ξ [the unobservable]”.

towels, respectively). Both studies use random-coefficients logit models with idiosyncratic taste shocks for each alternative, which is serving size/package size. This introduces an element of horizontal differentiation between quantity variants that is unrelated to quantity itself.⁹ This assumption will influence the estimated markups and substitution patterns, and thereby the conclusions about price discrimination. In contrast, my model follows the theory literature on price discrimination in assuming that quality variants differ only in observed quality.

Verboven (1999) and Verboven (2002) look at price discrimination over automobile engine sizes and fuel types, respectively. Both papers find that there is discrimination across variants, as is confirmed in this paper. These studies differ from my paper in that they do not use fully structural models of demand for variants.¹⁰ Like Ellison (2005) and Ellison and Ellison (2009), Verboven (1999) emphasises the role of imperfect information about the prices of premium variants in explaining the presence of discrimination. In this respect, my results show that there may be price discrimination also when consumers are perfectly informed about prices.

The next section describes the utility model and derives the choice probabilities. Section 3 explains the identification strategy and compares it to the literature. Section 4 describes the data, and attempts to show that the assumptions of the model are satisfied by the data. Section 5 describes the estimator, section 6 the results. Section 7 looks at price discrimination across engine sizes.

⁹The distribution of the taste shocks is not estimated. Therefore the models cannot let the data determine the extent of horizontal differentiation between quantity variants.

¹⁰One paper uses a reduced-form specification, and the other conditions on the choice of model, and only models the choice of variant.

2 Model

2.1 Utility

I estimate the parameters of the conditional (on product choice) indirect utility functions of potential car buyers. Products are grouped into models. Products within a model, called variants, differ only with respect to observable characteristics. Products in different models differ with respect to characteristics that are not observable to the researcher, but known by market agents. Let $\mathcal{J} = \{0, \dots, J\}$ be the choice set, where $j = 0$ is the outside good.¹¹ Each product is described by a double $(x_j, m(j))$ - a vector of observable product characteristics and an index for the model to which the product belongs. Let the vector x_j^0 be the subvector of x_j consisting of the variant characteristics (all characteristics that vary across variants of a model) and let $x_{m(j)}^1$ be the remaining characteristics (model characteristics), so $x_j = (x_j^0, x_{m(j)}^1)$. Let p_j be price, y_i income, and z_i a vector of observable consumer characteristics. Conditional on choosing product j consumer i has indirect utility:

$$u_{ij} = -(\bar{\alpha} + \tilde{\alpha}_i)p_j + x_j^0(\bar{\beta}^0 + \tilde{\beta}_i^0) + x_{m(j)}^1\tilde{\beta}_i^1 + \delta_{m(j)} + \epsilon_{im(j)} \quad (1)$$

$$u_{i0} = \xi_0 + \epsilon_{i0} \quad (2)$$

$$\tilde{\alpha}_i = \tilde{\alpha}(z_i, y_i) \quad (3)$$

$$\tilde{\beta}_i^r = \tilde{\beta}^r(z_i, \nu_i), \quad r = 0, 1 \quad (4)$$

$$\epsilon_{im(j)} \sim \text{iid extreme value}, \quad (5)$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ are functions of consumer characteristics whose realisations (z_i) or distributions (y_i, ν_i) are known.

The model mean effects, δ , can in theory be decomposed into the mean effect of model characteristics and unobservable characteristics: $\delta_m = x_{m(j)}^1\bar{\beta}^1 +$

¹¹Individuals or products are not followed across time. I.e. they have different subscripts for different years of data.

ξ_m . For most applications, the decomposition of δ is not required, nor is it identified with the assumptions used in this paper. Unlike in the literature, δ does not contain the mean effects of price and variant characteristics.

$\xi_m + \epsilon_{im}$ captures the common and consumer-specific valuation of any aspect of model m that is not controlled for by the observable characteristics x . Since variants of a model differ only with respect to observable characteristics x^0 these terms have m -subscripts.

ϵ captures effects of unobservable characteristics which do not provide any information about the unobservable effect of another product on the same consumer, nor on the effect on a different consumer. An example is consumer A who likes the look of VW Golf, but dislikes the Ford Focus; consumer B who thinks the Golf is boring, and the Focus elegant; consumer C who likes the look of the Golf, and loves the Focus.

2.2 Choice probabilities

Denote the ‘observable’ part of utility $u - (\delta + \epsilon)$,

$$V_{ij} = -(\bar{\alpha} + \tilde{\alpha}_i)p_j + x_j^0(\bar{\beta}^0 + \tilde{\beta}_i^0) + x_{m(j)}^1\tilde{\beta}_i^1.$$

Conditional on the realisations of the random variables (y, ν) the probability that consumer i chooses model m is

$$Pr(m|i, y, \nu) = Pr(\max_{j \in \mathcal{J}^m} \{V_{ij}\} + \delta_m + \epsilon_{im} \tag{6}$$

$$> \max_{j' \in \mathcal{J}^{m'}} \{V_{ij'}\} + \delta_{m'} + \epsilon_{im'}, \quad m' \neq m|i, y, \nu)$$

$$= \frac{\exp(\max_{j \in \mathcal{J}^m} \{V_{ij}\} + \delta_m)}{1 + \sum_{m' \in \mathcal{M}} \exp(\max_{j' \in \mathcal{J}^{m'}} \{V_{ij'}\} + \delta_{m'})}, \tag{7}$$

where \mathcal{J}^m is the set of variants of model m , and \mathcal{M} is the set of models in the market. Conditional on (y, ν) , the $\max\{V\}$ -terms are not random variables.

The second equality therefore follows in a standard way by integrating over the distribution of the ϵ -term to obtain a logit choice probability.

Still conditional on (y, ν) , consumer i 's probability of choosing j is equal to $Pr(m|i, y, \nu)$ if j maximises $\{V_{ij}\}$ over \mathcal{J}^m , and zero otherwise. Letting the indicator function $1[j|m, i, y, \nu]$ be one in the first case and zero in the second case, we can write the choice probability of product j

$$Pr(j|i, y, \nu) = 1[j|m, i, y, \nu] \frac{\exp(\max_{j \in \mathcal{J}^m} \{V_{ij}\} + \delta_m)}{1 + \sum_{m' \in \mathcal{M}} \exp(\max_{j' \in \mathcal{J}^{m'}} \{V_{ij'}\} + \delta_{m'})}. \quad (8)$$

Integrating over the distributions of y and ν , we now get consumer i 's unconditional choice probability for product j

$$Pr(j|i) = \int \int Pr(j|i, y, \nu) f(y|i) f(\nu) dy d\nu, \quad (9)$$

where $f(\nu)$ is the joint density of ν , and $f(y|i)$ is the density of the empirical income distribution in the population for consumers with observed characteristics z_i in the time period $t(i)$ in which i enters the data.

3 Identification

The model is estimated by assuming orthogonality between the individual-level prediction error (difference between observed choices and choice probabilities) and explanatory variables. This error term is distinct from the literature's unobserved product characteristic. I control for unobserved product characteristics using fixed effects for car models in the utility function. Unobserved product characteristics therefore do not enter the error term. The error term is pure sampling error. Price and product characteristics are therefore exogenous. I do not need orthogonality between unobserved and observed product characteristics to identify the price parameter, since it is

identified by within-model variation separately from model fixed effects.

3.1 Unobserved product characteristics and identifying assumptions

In discrete-choice demand models consumers choose among alternatives by maximising a utility function that depends on observed product characteristics. But typically consumers also care about product characteristics that are not observed by the researcher. Berry (1994) and BLP solve this problem by introducing product-specific constants in utility.

The product-specific constants ensure that utility includes the effect of unobservables. But the product-specific constants also capture the mean (across consumers) effect of price and observed characteristics. In order to separate the two effects, the literature assumes that unobserved and observed product characteristics are orthogonal.¹² That is, the observed characteristics of a product provides no information that affects the expected value of the unobserved characteristic of that product. Since both unobserved and observed characteristics are choice variables of firms, this assumption is problematic.¹³

¹²When observations are at the level of products, like in BLP, the orthogonality assumption is in fact crucial to identify any parameters at all. Since the product-level constants are observation specific, they explain all the variation in the data. Berry (1994) shows that regardless of the values of the other parameters, we can find product-level constants that exactly matches the model's predicted market shares to the observed counterparts. Individual-level studies can in fact avoid the orthogonality assumption by letting the coefficients on prices and product characteristics be of the form $\bar{\beta} + \tilde{\beta}_i$ with $\bar{\beta} = 0$, where $\tilde{\beta}_i$ is an interaction with demographic variables, income or another random variable. This is, however, a dubious identification of the "marginal effects" of x , since marginal utilities are not identified if $\tilde{\beta}_i = 0, \bar{\beta} \neq 0$ instead.

¹³Akerberg, Benkard, Berry, and Pakes (2007) p. 4197 say: "There are plausible reasons to believe that product characteristics themselves are correlated with ξ [the unobservable]. After all the product design team has at least some control over the level of ξ , and the costs and benefits of producing different levels of the unobservable characteristic might well vary with the observed characteristics of the product." Akerberg and Crawford (2006) say: "Just like price, product characteristics are typically choice variables of firms, and as such one might worry that they are actually correlated with unobserved components of demand."

An alternative identification strategy, which does not require orthogonality between unobserved and observed characteristics, is to use variation across products for which the unobserved characteristic can be assumed fixed. In this way Goldberg (1995) uses fixed effects for automobile brands to control for unobservable product characteristics.

My identification strategy uses the same principle. I restrict unobservable characteristics to be the same across engine variants of each car model. Since I observe products at the variant level, I can control for unobservables at the model-market level (rather than brand) with fixed effects. I therefore do not need either Goldberg's restriction that unobservables are the same within a brand, or BLP's orthogonality assumption.

3.2 The error term

The model is estimated by assuming orthogonality between the individual-level prediction error (difference between choice probabilities and the indicator function for observed choices) and the explanatory variables (product and consumer characteristics).

When observed market shares are generated by a large number of consumers, sampling error (the difference between a market share generated by a small number of consumers and the choice probability of the true model) should go to zero, i.e. choice probabilities should exactly equal observed market shares. This paper uses individual-level data, but I observe only the sex and age of the buyer, so that the data can be regarded as market-share data for submarkets defined by the age and sex of the buyers. In each market there is only about 500 individuals distributing between approximately 300 (inside) alternatives. On average, therefore, each observation (group market share) is generated by less than two individuals. Since the true model's choice probabilities for each submarket will not be attained with such a low number

of draws, there will be a discrepancy between observed and predicted shares. The prediction error is therefore explained by sampling error.

The literature emphasises that price is endogenous because the error term is the unobserved product characteristic. In this paper, the error term is not the unobserved characteristic, but rather the individual-level prediction error. Unobserved product characteristics are controlled for by the model fixed effects. Therefore they do not enter the error term. I assume individual-level sampling error does not enter the structural pricing (supply) equation. Price is therefore exogenous.

4 Data

The first subsection describes the data. The assumption that there are no unobservable differences between variants is central to the paper's identification strategy. In subsection 4.2 results from a hedonic regression indicate that the assumption holds. Subsection 4.3 discusses how care has been taken to ensure that trim levels do not cause unobservable differences between engine variants. Subsection 4.4 discusses further issues concerning trim levels. The reader who does not care about this issue may skip subsections 4.3 and 4.4.

4.1 Data

The data set is constructed by combining two data sources: new vehicle registrations and price lists from Norway 2000-2007.¹⁴ The registration data give the number of units sold by the sex and age of the buyer.¹⁵ I define models

¹⁴Both were provided by *Opplysningsrådet for Vegtrafikken AS* (the information council for road traffic).

¹⁵In the sales data, products are defined by the following characteristics: *brand, model (nameplate), body type, cylinder volume, horsepower, fuel type, number of seats* and *drive-*

as products of the same brand (e.g. Toyota) and of the same nameplate (e.g. Corolla). Variants are defined by the engine, as specified by horsepower and whether it is diesel or not. The sales of a variant is the sum of sales over all products of a model which are the same in terms of horsepower and fuel type. (In the literature, for instance in BLP, sales are the sums over products with the same nameplate.)

Table 1: Descriptive statistics

year	variable	mean ^a	s.d.	Means by sex and age					
				Men			Women		
				18-35	36-51	52-	18-35	36-51	52-
2000	variants	241							
	models	82							
	price ^b	327	157	292	306	276	242	243	225
	tax	97	60	88	91	79	68	67	61
	unit sales	56300		7961	15433	15483	4156	8008	5259
2007	variants	375							
	models	94							
	price	351	194	318	335	310	273	279	245
	tax	154	127	123	133	119	100	103	86
	unit sales	78139		10907	21499	23064	4444	10245	7980

^aNot sales-weighted

^bPrice and tax in 1000s of 2004 kroner, adjusted by CPI.

Table 1 shows summary statistics for the first and last of the eight years in the data. The data have the sales of 491,853 units, spread over 2397 products, 661 models (giving an average of 3.6 variants per model), 48 age groups and two sex groups. There is considerable variation along dimensions useful for identifying the parameters: the number of products across years, average car tax in different years, and in the purchases of different consumer groups.¹⁶

wheels (two-wheel drive vs. four-wheel drive). Price lists include the further product characteristics *length*, *weight*, *fuel consumption at mixed driving*, *airconditioning*, *number of collision bags*, *number of gears*, *automatic vs. manual gears*, *whether frontwheel or rear-wheel drive if 2WD*, *number of doors*, and *styling package*: a set of features (predominantly aesthetic) summed up by a tag, like ‘sportline’ or ‘comfort’.

¹⁶Differences in average characteristics purchased between groups for a given choice set reveal the role of demographic variables. Correlations in sales across time within groups

To allow consumers the choice of not purchasing a new car I need an estimate of the total market size. Since my data are by individual rather than by household, I could use the total population in each demographic (sex-age) group to estimate market size. However, on average 98.5% of people do not buy a new car in a given period. Since the goal is to analyse substitution patterns for cars, it seems reasonable to focus on the preferences of a more closely defined group.¹⁷ I let the market size of a consumer group be twice the maximum (across periods in the data) number of people in that group who bought a new car.

The income distributions conditional on demographic group are from the population as a whole, not from the potential car buyers as I have defined them, but this is the best approximation available. I use data from Statistics Norway on the number of people in twelve sex-age groups belonging to each of nine income brackets to estimate a kernel-smoothed income probability density function for each population group.

4.2 A hedonic regression

The assumption that there are no unobservable differences between models is central to the identification strategy of this paper. Results from a hedonic regression indicate that the assumption holds. Since any unobservable difference that affects the average consumer's valuation of a good should show up in price, I look at how much of price variation can be explained by observed characteristics and model effects. Regressing prices (for all products)

when the choice set changes (number of products, price, characteristics) reveal the effect of unobserved heterogeneity. Differences in (groupwise) sales within a model reveal the mean effects of price and engine characteristics.

¹⁷Also, for a given number of simulation draws, the larger the share of the outside good, the fewer simulation draws will result in a choice of one of the inside alternatives, making the simulation of the inside market shares less accurate. The problem could be reduced by oversampling draws that land on inside goods, like BLP do, but it requires an initial estimation with standard simulation techniques, which is difficult to do with any accuracy given the large computational burden of the model.

on model(-year) dummies (2397 products, 660 dummies), horsepower, fuel cost and diesel, cylinder volume and squares of horsepower, fuel cost and cylinder volume, and a constant gives an R-squared of 0.9913, leaving little scope for variant-specific unobservables to explain price variation.¹⁸ Results are given in table 2.

Table 2: Regression of price on engine characteristics and model dummies. Unit of observation is variants. 2397 observations. Coefficients on the 660 model dummies not shown.

<hr/> <hr/>		
R-squared	0.9913	
explanatory variable	estimate	p-value
horsepower	-0.005	0.547
fuel cost	0.099	0.000
diesel	0.035	0.000
cylinder volume	0.023	0.002
horsepower squared	0.085	0.000
fuel cost squared	-0.026	0.031
cyl. vol. squared	0.009	0.000
constant	0.066	0.000

4.3 Trim level: characteristics that vary within product units

In general, tractability concerns dictate that not all product characteristics can be included in the econometric model. Typically, for a given product unit as defined in the econometric model (e.g. nameplate/engine) the consumer faces a choice of other characteristics (e.g. transmission, leather interior), which affect the price of the product.

Denote the bundle of characteristics not included the ‘trim level’.¹⁹ To assign one price to each variant it is necessary to choose a trim level for each product.

¹⁸Including further engine-related characteristics (weight, co2 emissions), interactions and third- and fourth-order terms increases R-squared to as much as 0.9972, but this is open to charges of overfitting.

¹⁹Depending on specific modelling assumptions, my definition of the expression ‘trim’ may be wider than common usage.

Denote this the ‘baseline’ trim level. A central identifying assumption in this paper is that there are no unobserved differences between variants. In the following I discuss trim levels in some detail in order to show that this issue does not cause any violation of the restriction on unobservables.

Trim levels do not cause any unobserved difference between variants as long as three conditions are satisfied: the same set of trim upgrades over the baseline is available for all variants of a model; the price of upgrades is the same for all variants; and variants within a model are assigned the same baseline trim level.²⁰ Together with additively separable utility, this ensures that consumers choose trim upgrade independently of variant.²¹ The next three paragraphs look at how each condition is satisfied.

In some cases not all variant/trim-combinations show up in the price lists. But examination of car brands’ national web pages show that every engine variant is available with the full choice of trim levels in almost all cases. The few exceptions are mostly variants with engine sizes that are outliers relative to the model range, whose market shares are extremely small, which are not offered with the cheapest trim levels.

The price lists show that a given trim upgrade almost always costs the same regardless of engine size. That is, if you have to pay \$321 to upgrade from “basic” to “super” in the 2.0litre variant, you also pay \$321 for the same upgrade with the 2.4litre variant. I use this regularity to infer prices for variant/trim-combinations that are missing in the data. For the variants

²⁰In fact a weaker condition on availability is sufficient: let $J_i^m \subset J^m$ be the subset of variants in model m such that consumer i prefers any variant in J_i^m to any variant in $J^m - J_i^m$ regardless of trim level (even if the variants in the first set have the worst possible trim and those in the second the best possible). Then it is sufficient that the trim level that i prefers is available for all variants in J_i^m , since the availability of trim will not change the choice of variant. It follows that for instance unavailability of low-end trim with top-end engine variants is unlikely to contaminate the estimated tastes for engine size.

²¹Because of heterogeneity in price sensitivity, this framework is consistent with more expensive variants being sold with more expensive trim: people who do not mind paying extra for more horsepower may not mind paying extra for leather seats.

that still do not have a price at the baseline trim after imputing prices in this way, I reassign sales to nearest neighbour (according to horsepower, and if tied, fuel type).

The final condition, that all variants of a model are assigned the same baseline trim is up to the researcher as long as the corresponding prices are available. Since I do not have prices for every variant-trim combination, I chose as baseline the package for which I have prices for the variants corresponding to the highest sales (if tied, the cheapest package).²²

This discussion shows that the issue of baseline trim does not cause any unobserved differences between variants. Still the choice of baseline may be of some consequence. But as I discuss in the next subsection, these consequences are likely to be negligible - especially by comparison to the literature, where baseline trim also includes engine characteristics.

4.4 Trim level here and in the literature

A consumer's valuation of the baseline trim level will enter the unobserved characteristic. Choosing "super" instead of "basic" as baseline for model m will increase the prices of all variants in m by an equal amount. To ensure that m keeps its market share δ_m will adjust upwards. For any given consumer the effect is exactly the same on all variants, and therefore does not create unobserved differences between variants. But the choice of baseline may still have an effect: consumers with a low price sensitivity will substitute into model m (which has become more attractive since they exchange money

²²The raw data have total sales of 523,702 over eight years. For 499,167 of these matches are found between price lists and sales data. Failures may be due to privately imported obsolete products. Observations that cannot be matched are discarded. The number of sales for which there is a match between the two data sources, *and* for which I can find a price (at a baseline trim level) is 440,821. Remaining sales are reattributed to nearest neighbours. Finally I discard models with sales in a given year of less than 100 units. This reduces the number of products from 3,588 to 2,397 and total sales to 491,853.

against δ at a higher rate than the average consumer; and an equal amount with high price sensitivity will substitute out of the model). This may in principle affect the estimated distribution of price sensitivity. But the effect of these trim differences is likely to be negligible compared to the effects of engine characteristics, and model characteristics such as segment and size.

The choice of baseline is of less consequence in this paper than in the model-level literature. There are three reasons for this: first, in model-level studies, trim includes engine upgrades over the baseline engine. Since this makes trim account for a larger share of utility, the magnitude of the problem is greater. Secondly, I do not estimate tastes for components of trim, while model-level studies do, since engine characteristics are in trim. Increasing baseline engine size therefore changes components of x . This makes consumers with low tastes for engine size substitute out of the model, which may affect estimates further. Thirdly, engine upgrades may be correlated (in the population of car models) with baseline engine sizes (in ways that may depend on the choice of baseline). Since the engine upgrades enter the unobserved characteristic, ξ , they may be an additional source of dependence between ξ and x , violating the orthogonality assumption used in the literature.

5 Estimation

5.1 The objective function

Parameters are estimated by GMM, with moments

$$g_{ij}(\theta) = Z_{ij}[d_{ij} - Pr(j|i, \theta, \delta(\theta))], \quad (10)$$

where Z is a vector of functions of the explanatory variables.²³ d_{ij} is one if consumer i chooses alternative j , zero otherwise. $Pr(j|i, \theta, \delta(\theta))$ is consumer i 's choice probability for product j as defined in the model section (equations (7-9)). The model-market fixed effects, δ , are a function of the parameters, found by setting the aggregate (across consumers) model choice probabilities equal to model market shares: $s_m = P(m|\theta, \delta)$.

The objective function is

$$\left[\sum_{i=1}^n \sum_{j \in \mathcal{J}^{t(i)}} g_{ij}(\theta) \right]' W \left[\sum_{i=1}^n \sum_{j \in \mathcal{J}^{t(i)}} g_{ij}(\theta) \right], \quad \text{where} \quad (11)$$

$$W = \left[(n)^{-1} \sum_{i=1}^n \sum_{j \in \mathcal{J}^{t(i)}} Z'_{ij} Z_{ij} \right]^{-1},$$

where $\mathcal{J}^{t(i)}$ is the choice set of consumer i , depending on the time period where i is observed, and n is the number of individuals observed (sum of market sizes over groups and years). Efficiency could be improved by approximating the ideal instruments using initial estimates. But because of the large computational burden, and because the size of the data set makes efficiency relatively less of a worry I choose to use estimates from (11) as final.

The vector δ that sets predicted and observed model market shares equal is found by the BLP contraction mapping.²⁴ Conditional on the parameters, the model choice probabilities are an average over logit choice probabilities, like in BLP.

²³20 parameters to estimate, 22 instruments: *price, price², kw, kw², kw · age, kw · age², kw · wom, fuelcost, fuelcost², fuelcost · age, fuelcost · age², fuelcost · wom, diesel, diesel · age, diesel · age², diesel · wom, length, length², age, age², wom, constant.*

²⁴ $\delta_m^{t+1} = \delta_m^t + \log s_m - \log P(m|\theta, \delta^t)$

5.2 Simulation and asymptotic properties

The integral with respect to the density of ν and y in the choice probabilities is computed by a frequency simulator.²⁵ The simulator is not smooth in the parameters. I analytically integrate over the distribution of the logit term, but the choice probabilities still contain discontinuous functions of the parameters: the indicator function $1[j|m]$, and a max function for every model.²⁶ Indicator functions or max functions may change in jumps or not at all for a given change in their argument, making the same true for the objective function that depends on them. McFadden (1989) (Theorem 1, p. 1014) shows consistency and asymptotic normality for an estimator like (11) where g is simulated by a function allowed to have jumps.²⁷ A consistent estimator for the asymptotic covariance matrix is

$$\hat{A}var(\hat{\theta}) = (\hat{G}'W\hat{G})^{-1}\hat{G}'W\hat{\Lambda}W\hat{G}(\hat{G}'W\hat{G})^{-1}/n \quad \text{where} \quad (12)$$

$$\hat{G} = n^{-1} \sum_{i=1}^n \sum_{j \in \mathcal{J}^t(i)} \nabla_{\theta} g_{ij}(\hat{\theta}) \quad (13)$$

$$\hat{\Lambda} = n^{-1} \sum_{i=1}^n \sum_{j \in \mathcal{J}^t(i)} \sum_{j' \geq j} g_{ij}(\hat{\theta}) g_{ij'}(\hat{\theta})'. \quad (14)$$

²⁵Using 30 quasi-random draws (scrambled Halton, generated by Matlab's *haltonset*) for each of the 96 consumer groups in each of the 8 years. See Train (2003) for a discussion of Halton draws.

²⁶I did not want to use a smoothed simulator (such as a logit-smoothed accept-reject) as discussed in McFadden (1989), because it would effectively remove the feature of my model that variants differ only with respect to observed characteristics. In principle one could create a smooth simulator by decomposing the unconditional probability into a sum of probabilities conditional on which variant is chosen in each model, times the probability of the conditioned-on event. If there are M models and each model has V variants, the sum will have M^V terms, i.e. approximately $661^{3.6} = 1.4\text{E}+10$ in this paper. The computational burden is roughly proportional to the number of separate choice probabilities which need to be computed, which is proportional to the number of draws and to the number of terms in the sum above. Therefore the computational burden goes down with the smooth simulator relative to the one I use only if we can reduce the number of draws by a factor of M^V to obtain an equally good simulator. A second alternative would be to use a

²⁷See his assumption A12 p. 1018 for a regularity condition on the behaviour of the simulator.

Berry, Linton, and Pakes (2004) show that under assumptions typical in the IO demand literature, consistency requires that the number of simulation draws must grow as the square of the number of products. There are two reasons that the result does not apply to the model in this paper (so that I can rely on the results of McFadden (1989)). First, they analyse a situation of dependence between observations in market-level data, because one market share depends on the shares of other products. In this paper observations are individual purchases which do not depend on each other. Secondly, they look at a model where the error term (the unobservable characteristic) is a nonlinear transformation of the predicted market shares, so that simulation errors in the choice probabilities will cause a bias. In this paper the error term is a linear function of the choice probabilities, so that simulation error cancels across observations and disappears as the number of observations goes to infinity with the number of simulation draws held fixed.

Because of the irregularity of the objective function I need an optimisation algorithm that does not require continuity and is robust to local optima. I use the differential evolution genetic algorithm.²⁸

6 Results

6.1 Parameter estimates

Table 3 shows the estimates from the full model, and table 4 gives variable definitions and units. The mean taste coefficients, obtained by interacting with the mean values of the demographic variables, are: horsepower -7.7, fuel cost 18.7, diesel 0.6, and price -16.6. (For length and the constant for the outside good the means are contained in δ and therefore unknown.)

²⁸Developed by Kenneth Price and Rainer Storn, implemented for Matlab in the *devec3* code. The code can be found on www.icsi.berkeley.edu/~storn/code.html.

Table 3: Parameter estimates from the full model. 491,853 observations.

product characteristic	interacted with	est.	s.e.
horsepower		4.767	0.007
fuel cost (kroner/km)		4.932	0.007
diesel		-0.011	0.001
horsepower	age	-12.303	0.024
	age sq.	-24.685	0.033
	woman	1.425	0.010
fuel cost (kroner/km)	age	12.172	0.021
	age sq.	30.062	0.042
	woman	-1.604	0.013
diesel	age	0.044	0.003
	age sq.	2.158	0.004
	woman	0.117	0.008
horsepower	std.norm.	0.435	0.005
fuel cost (kroner/km)	std.norm.	-0.167	0.004
diesel	std.norm.	1.229	0.015
length	std.norm.	-0.102	0.007
inside good (const.)	std.norm.	7.719	0.010
price ^a	α_1	2.935	0.014
	α_2	47.082	0.074
	α_3	3.614	0.007

^aThe price coefficient is $-\alpha_1 + \alpha_2 \exp(-\alpha_3 \cdot \text{income}_i)$.

Table 4: Variables and units

Variable	Unit	mean	st.dev.	min	max
price ^a	kroner*1E-6	0.34	0.18	0.12	1.87
horsepower	kW*1E-2	0.98	0.39	0.37	3.75
fuel cost	(kr/litre)*(litres/km)	0.77	0.19	0.38	1.86
diesel	1 if diesel, 0 if petrol	0.34	0.47	0	1
length	metres	4.40	0.32	3.41	5.08
age ^b	age*1E-2	0.50		0.23	0.70
age sq.	age squared *1E-4	0.27		0.05	0.49
woman	1 if woman, 0 if man	0.29		0	1
income	kroner*1E-6	0.34		1E-3	3.28

^aprice, fuel cost and income in 2004 kroner, adjusted by CPI. 100 kroner, abbr. 'kr', is about 12 euros or 17 US dollars.

^bPersons of age<23 are assigned age 23 and with age>70 assigned age 70.

The first two do not have the expected signs. Clearly the analysis does not succeed in disentangling these highly correlated effects. However, since the variables represent aspects of the same underlying feature, engine size, this does not necessarily compromise the model's ability to predict substitution patterns. A consumer usually cannot change one of these characteristics without changing the other, and so the effect of having a different engine in practice works through both characteristics. Gramlich (2009) discusses this technological frontier in detail.

I regress horsepower on fuel cost over all products to obtain a rough measure of the technological connection between the two. The slope coefficient is 1.45. So if, starting from the average car, we move to a car with fuel cost 0.1 (kr/km) higher this car would typically have horsepower 0.145 (14.5 kW) higher. The average consumer would be willing to pay $(-7.7 \cdot 0.145 + 18.7 \cdot 0.1)/16.6 = 0.0454$, i.e. 45,540 kroner for this change. To see that this is roughly in line with market conditions, a regression of price on horsepower and fuel cost yields coefficients 0.43 and -0.13, respectively. Plugging in the changes in the two variables, we can expect price to go up by $0.43 \cdot 0.145 - 0.13 \cdot 0.1 = 0.0493$, i.e. 49,300 kroner.²⁹

6.2 Price elasticities

I compute price elasticities for models by taking derivatives with respect to a change in the prices of all variants in a model as a percentage of the sales-weighted mean price. Own-price elasticities range from -3 to -7, and absolute value tends to increase with price. For semi-elasticities this pattern

²⁹Slope is 1.69 if I include length and diesel in the horsepower regression. Any characteristic in the regression is conditioned on, so in this case the slope is the technological trade-off between fuel cost and horsepower keeping length and fuel type constant. Willingness to pay is 34,300 if length and fuel type are kept constant. Including length and diesel in the price regression gives coefficients of 0.35 and 0.19, resulting in an expected price change of $0.35 \cdot .169 + 0.19 \cdot 0.1 = 0.0781$, 78,100 kroner. All regression coefficients are significant to 95% except that (-0.13) on fuel cost in the first price regression.

is reversed.

Table 5 shows own- and cross-price elasticities (multiplied by 100 for readability) for a sample of models in the 2007 market (found by ordering models by price and picking every fifth product). Products are ordered by price with highest price top right. As usual the table shows the elasticity of demand of the row entry with respect to the price of the column entry. There is a clear pattern of higher cross-elasticities close to the diagonal (from top-right to bottom-left), and lower as we move away from the diagonal: cross-elasticities are higher among products in the same price category.

Variant own-price elasticities are larger in magnitude than those for models, since they include substitution within the model. Unlike model elasticities, variant-own price elasticities (and semi-elasticities) decrease in absolute value with price, presumably because intra-model substitution is larger in cheaper models.

Table 6 shows variant elasticities for several variants of the same models with the purpose of highlighting intra-model substitution patterns. The products were picked by taking a few arbitrary blocks of products adjacent in the data. Elasticities between variants of the same model can be high, but not all are, and some are zero. Restricting the idiosyncratic shock to be the same for all variants of a model permits, but does not impose, high intra-model elasticities. Substitution is spread less evenly across all products compared to demand models with idiosyncratic tastes for all products.

Table 5: Model elasticities (x100) - w.r.t. price change for all variants of a model

Model																
Mercedes-Benz E	-452	1.4	18.3	0.9	0.6	0.5	2.5	5.3	3.3	0.4	9.5	0.4	3.2	0.9	2.2	0.2
Saab 9-5	3.5	-510	20.3	1.0	0.7	0.5	3.1	6.5	4.0	0.6	12.3	0.6	4.6	1.3	3.4	0.4
Toyota RAV4	3.1	1.3	-455	1.2	0.9	0.6	3.9	8.3	5.0	0.7	16.3	0.8	6.6	1.9	5.2	0.6
Jeep PATRIOT	3.0	1.3	24.1	-481	0.9	0.6	4.0	8.5	5.1	0.8	16.7	0.8	6.8	2.0	5.4	0.6
Hyundai TUCSON	2.9	1.3	24.4	1.2	-487	0.7	4.3	9.0	5.4	0.8	18.1	0.9	7.6	2.2	6.1	0.7
Kia SPORTAGE	2.5	1.1	19.8	1.0	0.8	-490	5.0	8.5	5.9	1.0	19.5	1.0	9.2	2.6	6.8	1.0
Mazda 6	2.0	1.0	18.3	0.9	0.7	0.7	-483	9.2	6.7	1.2	23.4	1.2	12.1	3.4	9.2	1.4
VW TOURAN	2.3	1.1	21.0	1.1	0.8	0.7	5.0	-479	6.2	1.0	22.3	1.1	10.3	3.1	8.9	1.1
Volvo V50	2.1	1.0	18.6	0.9	0.7	0.7	5.3	9.1	-488	1.1	22.8	1.2	11.4	3.3	9.2	1.4
Dodge CALIBER	1.7	0.9	18.0	0.9	0.7	0.7	6.0	9.6	7.1	-474	25.3	1.3	13.5	3.9	10.6	1.6
VW GOLF'	1.8	0.9	18.4	0.9	0.7	0.7	5.7	10.0	7.0	1.2	-453	1.3	13.2	4.0	11.4	1.6
Hyundai I30	1.6	0.9	18.0	0.9	0.7	0.7	5.9	10.0	7.2	1.2	26.7	-452	13.9	4.2	11.9	1.8
Toyota COROLLA	1.4	0.8	16.5	0.8	0.7	0.7	6.5	10.2	7.7	1.4	29.3	1.5	-441	4.8	13.7	2.1
Kia CEE'D	1.3	0.7	16.5	0.8	0.7	0.7	6.3	10.5	7.7	1.4	30.0	1.6	16.4	-432	14.5	2.2
Toyota YARIS	1.2	0.7	16.6	0.9	0.7	0.7	6.3	11.2	7.9	1.4	32.1	1.7	17.4	5.4	-407	2.4
Honda JAZZ	0.9	0.6	13.0	0.7	0.6	0.7	6.8	9.7	8.2	1.5	32.6	1.7	19.2	5.8	17.0	-406
Mazda 2	0.8	0.5	12.1	0.6	0.5	0.6	6.9	9.8	8.3	1.6	34.5	1.8	20.6	6.4	18.9	3.5
Kia PICANTO	0.8	0.6	15.9	0.8	0.7	0.6	6.3	14.0	8.2	1.4	39.5	2.0	21.8	7.1	24.4	3.1

Table 6: Intra- and inter-model price elasticities (x100)

Variant																				
Citr.	C4	1.4	l	87	hp	P	-6669	501	5720	0.0	0.0	0.2	1.3	0.0	0.0	17.1	0.0	0.0	14.5	
Citr.	C4	1.6	l	88	hp	D	203.6	-3680	220	2698	99.5	1.8	0.0	0.0	0.0	1.0	0.1	0.0	0.0	
Citr.	C4	1.6	l	107	hp	P	1004	75.2	-1704	65.7	98.6	0.0	1.5	0.0	0.0	10.5	0.0	0.0	5.5	
Citr.	C4	1.6	l	107	hp	D	0.0	5890	311	-9381	3243	1.5	0.0	0.0	0.4	0.6	4.2	0.0	0.0	
	Citr.	C4	2	l	134	hp	D	0.0	90.0	152	734	-1637	0.8	0.1	0.0	0.3	0.8	2.7	0.0	0.0
	Citr.	C5	1.6	l	107	hp	D	0.1	2.4	0.2	0.8	1.7	-627	164	0.0	0.2	1.4	2.1	0.0	0.1
	Citr.	C5	1.7	l	121	hp	P	0.7	0.0	4.3	0.0	0.2	142	-626	0.0	0.0	10.6	0.0	0.0	6.2
	Toy.	COROLLA	1.4	l	95	hp	P	2.6	0.0	0.0	0.0	0.0	0.3	-11176	0.0	4236	0.0	0.0	42.4	
	Toy.	COROLLA	2	l	114	hp	D	0.0	0.0	0.0	0.0	6.8	1.8	0.0	0.0	-52847	303	25431	0.0	0.0
	Toy.	COROLLA	1.6	l	122	hp	P	1.2	0.1	4.0	0.0	0.2	0.2	1.5	14.0	10.4	-681	69.1	9.3	10.8
	Toy.	COROLLA	2	l	125	hp	D	0.0	0.2	0.0	1.2	3.4	1.2	0.0	0.0	816	416	-2914	0.0	0.0
	Toy.	COROLLA	1.8	l	127	hp	P	0.0	0.0	0.7	0.0	0.2	0.0	0.3	0.0	0.0	3121	0.0	-3680	0.2
	Toy.	YARIS	1.3	l	86	hp	P	1.7	0.0	3.8	0.0	0.0	0.0	1.5	0.1	0.0	19.3	0.0	0.0	-651
	Toy.	YARIS	1.4	l	88	hp	D	0.2	3.4	0.4	0.6	1.0	1.5	0.1	0.0	0.3	3.0	1.3	0.0	219
	VW	GOLF	1.4	l	79	hp	P	2.0	0.0	3.8	0.0	0.0	0.0	1.6	0.0	0.0	21.5	0.0	0.0	15.8
	VW	GOLF	1.9	l	88	hp	D	0.3	4.7	0.3	0.3	0.3	1.7	0.0	0.0	0.3	2.9	0.4	0.0	0.0

7 Second-degree price discrimination with engine variants

This section establishes that there is price discrimination across engine variants, in the sense that markups increase with engine size within models. However, a counterfactual shows that average markups would be higher if firms restricted each model to have only one engine variant. The benefits of price discrimination therefore do not offset the increased competition when there are many variants of each model.³⁰

7.1 The supply-side model

Since I have no data on transactions between car manufacturers, import companies and dealerships, I follow the literature and treat the supply side as consisting of vertically integrated car companies selling directly to consumers. The sales price can be decomposed as

$$p_j = (\textit{markup}_j + MC_j)(1 + \textit{vat}) + \textit{tax}_j \quad (15)$$

where MC is marginal cost (the incremental cost of producing one more unit of product j). I assume constant marginal costs within the relevant output ranges (no economies of scale). \textit{tax} is the engine tax (an increasing convex function of horsepower, weight and co2 emissions; in 2007 the tax on average accounts for 40% of price, ranging from 20 to 70%). $\textit{markup} + MC$ is the amount the car company is left with after paying the engine tax and the value-added tax, \textit{vat} , of 25%.

I follow the literature and assume car companies play a Bertrand game in which there is a unique equilibrium, where each company f chooses prices of

³⁰The analysis in this section uses the demand system estimated on the full data set, but markups and the counterfactual are computed only for the 2007 market.

its set F_f of products to maximise profits

$$\Pi_f = \sum_{j \in F_f} \left[\frac{p_j - tax_j}{1 + vat} - MC_j \right] Ms_j(p) - C_f, \quad (16)$$

where M is the size of the market (summed over all consumer groups). In equilibrium the first-order conditions for profit maximisation of each product must hold:

$$0 = \frac{s_j(p)}{1 + vat} + \sum_{j' \in F(j)} \left[\frac{p_{j'} - tax_{j'}}{1 + vat} - MC_{j'} \right] \frac{\partial s_{j'}(p)}{\partial p_j}, \quad j = 1, \dots, J \quad (17)$$

where $F(j)$ denotes the set of products owned by the company which owns product j . The only unknowns are the marginal costs, which I find by solving this system of J linear equations in the J unknowns. Once the system of equations is solved, we know markups and marginal costs.

7.2 Is there price discrimination over engine variants?

It is common to offer several quality variants of what is otherwise the same product. Some examples are different CPUs and hard-drives for PCs, bandwidths for internet services, business and economy class air travel, and automobile engine variants. Mussa and Rosen (1978) show how a monopolist may use qualities (or equivalently quantities) to sort consumers according to willingness to pay.

In competitive markets second-degree price discrimination over quality may not be possible. Verboven (1999), Armstrong and Vickers (2001), Rochet and Stole (2002) find that depending on the intensity of competition, markups may be fixed across quality variants for each firm. Since the theory results depend on specific assumptions about differentiation and symmetry, I let the data determine these specifics. The theory models are special cases of the

demand model used in this paper.³¹

Taking the average of the smallest engine variant over all models (with at least two variants) gives a markup of 52,800 kroner, while the average of the biggest variant over all models gives a markup of 83,800 kroner. For comparison, overall sales-weighted mean markup is 57,300 kroner. Average percentage markups are 18.36 for smallest variants and 19.75 for biggest variants. Figure 1 plots markups against horsepower for 9 of the 10 models in 2007 that have seven or more engine variants, with a line fitted by OLS. Models are arranged in order of increasing mean price, from left to right, then top to bottom. Markups clearly increase in horsepower for all models, with overall markups higher for the more expensive products.

To find the average (over models) effect of engine size on markups within a model, I regress markups (and percentage markups) on horsepower, horsepower squared and cubed, a constant, and model dummies. The model dummies isolate the effect of horsepower, so that higher markups for models with overall higher horsepower does not contaminate the results. Table 7 shows the results. All higher order terms in horsepower are highly significant. Figure 2 plots the polynomial functions of markups and percentage markups obtained from the regression. Markups clearly increase in horsepower within a model, while percentage markups are approximately constant.

The conclusion is clear. Consumers who buy larger engine variants of a car model pay a higher premium over marginal cost than consumers who buy smaller engine variants of the same model. In this sense there is second-degree price discrimination. Percentage markups are approximately constant across engine variants.

³¹My model has many products, each firm offers several models with quality variants, there are several dimensions of taste heterogeneity, and there is an outside good. In the theory literature there are two firms that each offer one model with quality variants, and there is one vertical (taste for quality) and one horizontal (taste for models) dimension of taste heterogeneity. Some theory models do not have an outside good.

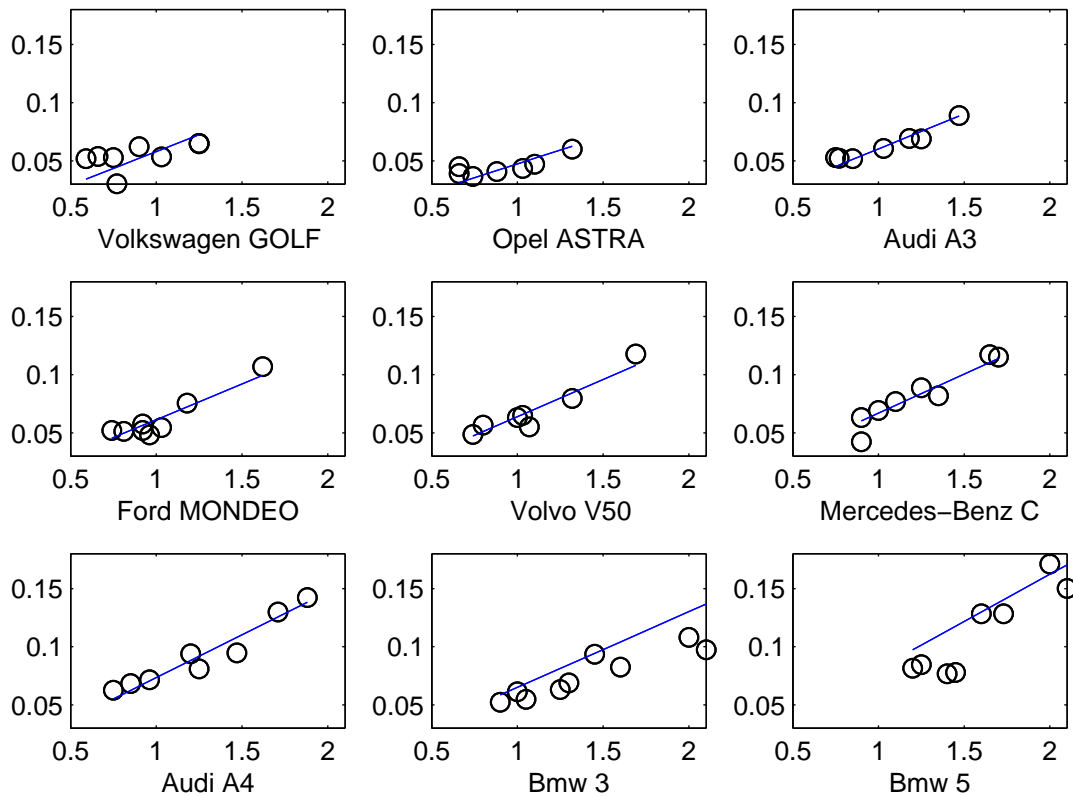


Figure 1: Markups (in $1e-6$ kroner, vertical axis) plotted against horsepower ($kw*1e-2$, horizontal axis) for models that have seven or more engine variants. Line fitted by OLS. Models arranged in order of increasing price. For the last two products some large engine variants do not show.

Table 7: Results from regression of markups and percentage markups (markup/price) on horsepower, higher order terms of horsepower and model dummies in year 2007. 375 observations, 94 model dummies. Coefficients on model dummies not shown.

Dependent variable	Explanatory variables			
	hp	hp ²	hp ³	constant
markup	-0.029	0.056	-0.007	0.038
(p-value)	0.205	0.001	0.047	0.000
markup/price	-0.214	0.129	-0.024	0.295
(p-value)	0.000	0.000	0.000	0.000

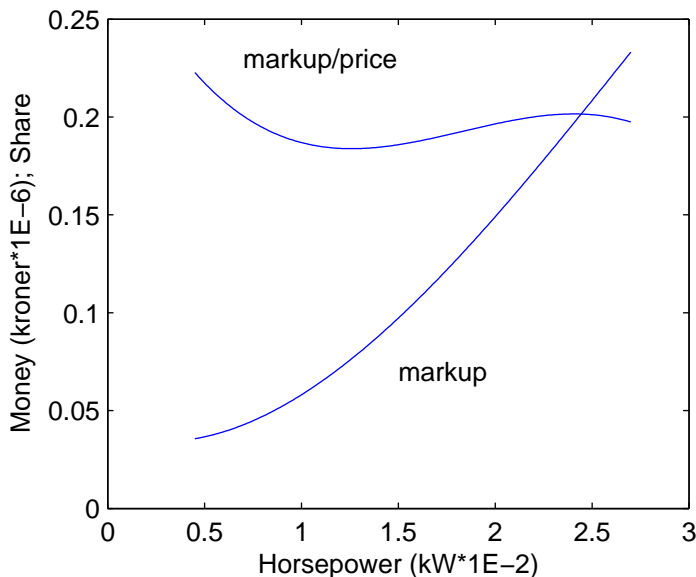


Figure 2: Within-model differences in markups and percentage markups as a function of engine size. Horsepower in the market ranges from 0.44 to 3.09, sales-weighted mean 0.92kW. 10th to 90th percentile (not sales-weighted) is (0.64,1.65)

Several theory papers find that pricing will be cost-plus-fixed-fee, i.e. no price discrimination. Since my empirical result is different, it is of interest to see what assumptions in the theory literature generates the cost-plus-fixed-fee results. I simulate some simple duopoly models that relax the assumptions of the models of Verboven (1999) and Ellison (2005) (by including an outside good, making qualities and marginal costs asymmetric, and introducing heterogeneity in one or both the price and quality parameters).³² I find that constant absolute markups are the exception rather than the rule.³³

³²I compute two-firm two-quality Nash equilibria in price in models where the taste for firms is logit; the price coefficient is either 0.5, $U(0,1)$ or $N(0.5,0.5)$; quality coefficient either 1 or $U(0.5,1.5)$; marginal costs are (3,4) for both firms, or (1,2)/(3,4); qualities are (7,9) for both, (6,8)/(7,9), or (3,5)/(7,9). If an outside good is included, it has utility -1.5 plus logit term. Equilibria are computed by iterating on the best-response functions. Best responses are computed using a genetic optimisation algorithm, constraining each firm's high quality price to be weakly higher than its low quality price.

³³Increasing markups is the most common result, while large asymmetries between firms may give decreasing absolute markups for one of the firms. For instance, pricing is no longer cost-plus-fee in Verboven's model if we include an outside good. Presumably, the participation constraint puts a downward pressure on prices, and more on the low quality product which is the closest substitute to the outside option.

7.3 Does discrimination compensate for increased competition?

The previous subsection establishes that there is price discrimination across engine variants of car models. But this does not imply that a market with engine variants is beneficial to firms. Multi-variant models crowd the product space and increase competition. Gilbert and Matutes (1993) find that increased competition from broad quality ranges may outweigh the benefits of quality-based price discrimination, so that firms would be better off if they could commit to restrict quality ranges. However, offering a full quality range is an individual best response regardless of the product line offered by the other firms. Firms are therefore unable to coordinate on the optimal outcome, and find themselves in a prisoner's dilemma.

To check whether car companies would prefer a market with single-variant models, I compare the actual market with equilibrium in a counterfactual market where each model is offered with only one engine size. I make two simplifying assumptions: first, that there are no fixed cost investments involved in expanding the range of engine sizes offered within a model (no economies of scope).³⁴ Secondly, when companies offer only one engine variant of each model, they choose the one that generates the highest profits in the current market.³⁵

Table 8 compares outcomes for the counterfactual single-variant equilibrium with the outcomes in the actual multi-variant equilibrium.³⁶ Profit increases,

³⁴Often a new variant is created by taking the baseline engine of one of the brand's higher segment models and putting it in the a lower segment model. Also, developing a more powerful version is less costly than developing an entirely new engine.

³⁵This is a limitation, and justified only because it simplifies computation. Endogenising the choice of baseline variant would require the computation of price equilibria for every possible configuration of engine sizes, which is not feasible. Allowing two choices for each firm would require the computation of 2^{94} equilibria. Relaxing the assumption of perfect information about the choices of other firms could simplify the computation of equilibria, but creating a model of such a game is outside the scope of the paper.

³⁶The numbers for the actual market are based on the model's predictions of sales. While these predictions are exactly correct for the sales of each model, and therefore aggregate sales, the model does not guarantee a perfect fit at the individual and variant level.

Table 8: Market outcomes with single-variant and multi-variant models

	i. single-variant models (counterfactual)	ii. multi-variant models (actual market)	change	% change
Profits ^a	4356	4460	104	2.4
Tax revenue	8341	9086	744	8.9
from Engine tax	5644	6290	646	5.3
from v.a.t.	2698	2796	98	3.6
Consumer surplus	63820	66597	2777	4.4
Unit sales	75618	77916	2298	3.0
Mean price ^b	288.7	296.6	7.9	2.7
Mean markup	57.6	57.3	-0.3	-0.5
Mean horsepower (kW)	87.3	91.6	4.4	5.0

^aProfits, tax and consumer surplus in million kroner.

^bPrice, markup and horsepower are sales weighted.

but only by a small amount, 2.4 per cent. In fact markups (average profit per unit) are marginally reduced, and profit goes up only because the increased product variety recruits some (3.0%) new consumers who otherwise would choose the outside good. It is therefore not price discrimination that increases profits, but rather increased variety attracting new consumers. Discrimination does not increase profits, which confirms the result of Gilbert and Matutes (1993).³⁷ Making the characteristics of the remaining engine variants endogenous may change the results, which should therefore be regarded with some caution. However, choosing the highest-profit variant of each model, as I currently do, means that chosen variants tend to be similar for models in the same segments. Models therefore compete head on in the current counterfactual. Allowing firms to choose characteristics would open up for more specialisation, and thereby increase markups even more, reinforcing my conclusion.

³⁷This result may depend on the way the engine tax increases the effective marginal cost of different engine sizes. Since the tax increases convexly in engine size it reduces the scope for high markups for large engines. The tax succeeds in appropriating a larger amount of consumer surplus when moving from single-variant to multi-variant models. (Tax revenue increases by 8.9 per cent, mostly due to the engine tax.) Depending on how much would be competed away, at least some of this would go to firms if there was no tax.

Consumers pay more, but are better off, implying that the non-price part of utility increases more than the disutility of price.³⁸ It has been pointed out that logit models tend to overestimate the benefit to consumers of product variety (see Petrin (2002)), because each new product introduces a new dimension of differentiation for which some consumers have a very high idiosyncratic taste. This is not the case here: no new dimension of product differentiation accompanies the introduction of new engine variants, because for each consumer, new variants have the same idiosyncratic shock as the baseline variant of the model. The crowding of characteristics space that results from multi-variant models is therefore taken into account.

8 Conclusion

This paper studies the extent and profitability of price discrimination with automobile engine variants, by developing and estimating a new empirical demand model that has the relevant theory models as special cases. The question of what engine variants to offer if restricting the quality range is related to the recent product-choice literature, e.g. Mazzeo (2002) and Gramlich (2009). I simplify this question by assuming that firms would choose their current highest-profit engine variant. Further work could look more closely at the interaction between discrimination and firms' product choice. While I focus on approximately vertically differentiated variants, the demand framework also fits horizontally differentiated variants and package sizes. I estimate the model on individual level data, but it could also be used for market level data, as long as a nonzero prediction error could be justified as sampling error or in another way.

³⁸See Train (2003) p. 266 for how to compute consumer surplus. In my utility specification income enters the price coefficient linearly, but disposable income enters only linearly. Total consumer surplus is computed over all consumers, including those choosing the outside good in either case, but these do not experience any change in consumer surplus since their choice or its attributes do not change.

References

- ACKERBERG, D., C. L. BENKARD, S. BERRY, AND A. PAKES (2007): “Econometric Tools for Analyzing Market Outcomes,” in *Handbook of Econometrics*, ed. by J. Heckman, vol. 6. North-Holland.
- ACKERBERG, D. A., AND G. S. CRAWFORD (2006): “Estimating Price Elasticities in Differentiated Product Demand Models with Endogenous Characteristics,” .
- ARMSTRONG, M., AND J. VICKERS (2001): “Competitive price discrimination,” *RAND Journal of Economics*, 32, 579–605.
- BAJARI, P., AND C. L. BENKARD (2005): “Demand Estimation With Heterogenous Consumers and Unobserved Product Characteristics: A Hedonic Approach,” *Journal of Political Economy*.
- BERRY, S. (1994): “Estimating Discrete-Choice Models of Product Differentiation,” *RAND Journal of Economics*, 25, 242–262.
- BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): “Automobile Prices in Market Equilibrium,” *Econometrica*, 63, 841–890.
- (2004): “Differentiated Products Demand Systems from a combination of Micro and Macro Data: The New Car Market,” *Journal of Political Economy*, 112(1).
- BERRY, S., O. B. LINTON, AND A. PAKES (2004): “Limit Theorems for Estimating the Parameters of Differentiated Products Demand Systems,” *Review of Economic Studies*, 71(3), 613–654.
- BERRY, S., AND A. PAKES (2007): “The Pure Characteristics Discrete

Choice Model of Demand,” *International Review of Economics*.

COHEN, A. (2008): “Package size and price discrimination in the paper towel market,” *International Journal Of Industrial Organization*, pp. 502–516.

CORTS, K. S. (1998): “Third-Degree Price Discrimination in Oligopoly: All-Out Competition and Strategic Commitment,” *RAND Journal of Economics*, 29(2), 306–323.

DUPOIT, J. (1849): “On Tolls and Transport Charges. Translated in *International Economics Papers* (London: Macmillan, 1952). Original version in *Annales des Ponts et Chaussées* 17,” .

ELLISON, G. (2005): “A Model of Add-on Pricing,” *Quarterly Journal of Economics*, 120(2), 585–637.

ELLISON, G., AND S. F. ELLISON (2009): “Search, Obfuscation, and Price Elasticities on the Internet,” *Econometrica*.

GILBERT, R., AND C. MATUTES (1993): “Product-line rivalry with brand differentiation,” *Journal of Industrial Economics*, 41, 223–240.

GOLDBERG, P. (1995): “Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry,” *Econometrica*, 63, 891–951.

GRAMLICH, J. (2009): “Gas Prices, Fuel Efficiency, and Endogenous Product Choice in the U.S. Automobile Industry,” .

MAZZEO, M. (2002): “Product choice and oligopoly market structure,” *RAND Journal of Economics*, 33, 1–22.

McFADDEN, D. (1989): “A Method of Simulated Moments for Estimation of

- Discrete Response Models Without Numerical Integration,” *Econometrica*, 57(5), 995–1026.
- MCMANUS, B. (2007): “Nonlinear pricing in an oligopoly market: the case of specialty coffee,” *The Rand Journal of Economics*, (38), 512–532.
- MUSSA, M., AND S. ROSEN (1978): “Monopoly and Product Quality,” *Journal of Economic Theory*, 18, 301–307.
- PETRIN, A. (2002): “Quantifying the Benefits of New Products: The Case of the Minivan,” *Journal of Political Economy*, 110(4), 705–729.
- ROCHET, J.-C., AND L. STOLE (2002): “Nonlinear Pricing with Random Participation,” *Review of Economic Studies*, 69(1), 277–311.
- STOLE, L. A. (2007): “Price Discrimination and Competition,” in *Handbook of Industrial Organization*, ed. by M. Armstrong, and R. Porter, vol. 3, chap. 34. Elsevier.
- TIROLE, J. (1988): *The Theory of Industrial Organization*. The MIT Press.
- TRAIN, K. E. (2003): *Discrete Choice Methods with Simulation*. Cambridge University Press.
- VERBOVEN, F. (1999): “Product line rivalry and market segmentation - With an application to automobile optional engine pricing,” *Journal of Industrial Economics*, 47(4), 399–425.
- (2002): “Quality-based price discrimination and tax incidence: evidence from gasoline and diesel cars,” *RAND Journal of Economics*, 33(2), 275–297.

