An Analysis of Automobile Demand

Øyvind Thomassen
Exeter College
University of Oxford

MPhil thesis in Economics
Trinity term 2005

ABSTRACT

Using recent market-level data for the Norwegian car market, I simulate the price equilibrium that would result from letting car companies price discriminate between women and men. Industry profit turns out to be lower in the discriminatory equilibrium than with uniform price. This possibility is envisaged in the theoretical literature on oligopoly third-degree price discrimination, e.g. Holmes (1989), but does not appear to have been demonstrated empirically before. I use a random-coefficients logit model with an equilibrium pricing assumption (Berry, Levinsohn, and Pakes 1995) to estimate the demand system. My data specify sales by sex and age group. This allows me to estimate separate taste parameters for each demographic group, improving the fit of the model as well as enabling me to perform the discrimination experiment.

ACKNOWLEDGEMENTS

Thanks to my parents, Grete Riis and Hans Petter Thomassen, my supervisor, Howard Smith, and Pål Bruhn at the Information Council for Road Traffic in Norway.

CONTENTS

1. INTRODUCTION

This thesis uses recent market-level data for the Norwegian car market to simulate the price equilibrium that would result from allowing automobile companies to price discriminate between women and men. Industry profit turns out to be (slightly) lower in the discriminatory equilibrium than with uniform price. The theory of oligopoly price discrimination, e.g. Holmes (1989), envisages this possibility, but it appears that it has never been demonstrated empirically before. I use a random-coefficients logit model with an oligopoly pricing assumption (Berry, Levinsohn, and Pakes 1995) to estimate the demand system.

Estimating demand systems in markets with many differentiated goods is challenging because of the large number of cross-price elasticities. Instead of estimating price elasticities directly, most recent studies impose a structure where elasticities are functions of product attributes. The range of products is projected onto a space of product characteristics. A consumer utility function defined on this characteristics space is used to assign a choice probability to each product. Estimation essentially takes place by matching these choice probabilities to observed choices. This yields the parameters of the utility function, which are finally used to extract the elasticities. Automobiles are well suited to be analysed in this way, since many important characteristics are easily quantifiable. Berry et al. (1995), Goldberg (1995) and Petrin (2002) are some of several papers which apply this type of methodology to the US car market. With some modifications I follow the method used by Berry et al. I now set this forth in more detail.

Consumer utility depends linearly on product characteristics. In the simple logit model, heterogeneity in taste enters only through the extreme value iid error term. This restricts substitution patterns to depend only on market shares, and not directly on characteristics. To allow for more realistic substitution effects, a parametric family of distributions is assumed for the taste coefficients. The goal of the estimation is to find the parameters of these distributions.

It seems unlikely that the characteristics entering the utility function can include all those relevant to consumer choice. An error term for product characteristics that are unobserved by the econometrician, but known to manufacturers and consumers, is therefore added to utility. Since this term affects demand, it also enters the pricing decision, making prices endogenous. Direct application of instrumental variables methods to deal with this problem is not possible, because price and unobserved characteristics enter demand equations in a nonlinear way. Berry (1994) shows how the function defining market shares can be inverted to uncover the mean utility levels of products. By relating this mean utility to observed characteristics and price, the endogeneity problem can be solved by instrumental variables techniques. The random-coefficents case requires a numerical procedure to perform this inversion, whereas the fixed-coefficients model yields a closed form solution.

The data used here are in some ways richer than those used by Berry et al. (1995). There, only total sales are observed. In my data, sales to women and men in different age groups are recorded separately. By allowing the distributions of the taste parameters to vary between the different consumer groups, my model is more flexible. It is commonly observed that women and men buy different cars, and my results do indeed show that important differences exist between the processes that drive women's and men's choices.

These data permit me to perform a counterfactual experiment. Using the first-order conditions for Nash equilibrium in each market and the estimated demand functions, I simulate a new price equilibrium under the hypothesis that automobile companies can price discriminate between women and men. Contrary to the monopoly case, in oligopoly price discrimination can reduce industry profit, as shown by Holmes (1989), Corts (1998) and Armstrong and Vickers (2001) for different market structures. I find that the discriminatory equilibrium does in fact yield a slightly lower profit than with uniform pricing.

The method proposed by Berry et al. (1995) entails a heavy computational burden compared to other, traditional approaches to demand estimation for differentiated products, but gives more realistic results. I apply an extended version of their technique to a new type of data. The resulting demand elasticities yield the prediction that price discrimination between the sexes reduces industry profit.

The next chapter provides some background and descriptive statistics. The theoretical model is set out in Chapter 3. Chapter 4 describes the data and explains the estimation procedure. The estimation results are presented in Chapter 5. Chapter 6 briefly discusses the theory of third-degree price discrimination in oligopoly, and presents the simulated discriminatory equilibrium. Chapter 7 concludes.

2. GENDER AND CAR CHOICE

It is commonly believed that women and men have quite different tastes in cars. According to a spokesperson for Toyota Norway, when choosing a car, "men care about what their neighbour thinks, while women are concerned with functionality".[1] The same notion is expressed by a Ford representative: "Women don't find horsepower particularly interesting. If they have children, they care about whether it's easy to get the children in and out of the car".[2] Marketing, too, reflects sex differences. For instance, ads in women's lifestyle magazines are almost exclusively for small cars. These pieces of anecdotal evidence indicate that there are important differences between women's and men's preferences. As a backdrop to the subsequent demand analysis, some summary facts about car sales in Norway are presented here. Table 2.1 shows aggregate sales of new cars to individuals in Norway in the

Tab. 2.1: Total car sales by year and consumer group

	Women			Men			Total
Year	-34	34-57	57-	-34	34-57	57-	
2003	1,970	10,504	4,633	3,921	22,072	15,166	58,266
2002	1,984	9,931	4,250	4,187	21,523	13,960	55,835
2001	2,224	9,486	3,989	4,208	20,288	12,190	52,385
2000	2,905	11,497	4,256	5,541	22,278	13,235	59,712
Pop.[a]	483,995	738,407	535,071	499,777	769,072	420,904	3,447,229

[a] Mean population over 18 years.

[1] Okonomisk rapport, 21.05.2004.
[2] Verdens Gang, 19.04.2003.

period 2000-03. The numbers are based on the names of registered owners. A household with two adults does not necessarily choose to register its new car in the name of the person who will be the most frequent user of the car. Still, it is likely that the owner and user in most cases coincide. Accordingly, the owner is here assumed to be the decision maker. About the table, it suffices to note that men buy more cars, and that middle-aged people are more frequent buyers than young and old people. In table 2.2 some numbers

Tab. 2.2: Mean characteristics of cars sold 2000-03, by consumer group

	Women			Men		
	-34	34-57	57-	-34	34-57	57-
Price (1000 kr[a])	217.78	222.07	209.94	244.07	266.90	242.10
Length (m)	4.17	4.18	4.09	4.34	4.41	4.31
Fuel cons.(l/km)	0.58	0.58	0.57	0.60	0.62	0.60
Power (kW)	65.21	66.97	63.89	72.19	77.88	72.56
Weight (kg)	1130.00	1139.66	1098.84	1207.83	1256.13	1199.28
% with air-cond.	63.39	65.56	58.19	77.62	82.68	75.97

[a] 2003 kroner, using the consumer price index. 100kr is approx. £8.

are given which characterise the types of car bought by different consumer groups. The pattern is easily recognisable: Men buy bigger, more expensive and more powerful cars than women, and middle-aged people of both sexes do the same compared to old and young people of the same sex. Table 2.3 presents a sample of cars with characteristics and sales. The first six cars were chosen because they have particularly high market shares among women relative to men, and the final six because they are especially popular with men. The demand pattern revealed by the mean characteristics is well illustrated by the sample. The remainder of this thesis uncovers the preferences revealed by this observed behaviour, and discusses some of their implications.

Tab. 2.3: A sample of characteristics and sales by group, 2003

					Sales to women			Sales to men		
Model	Price	Length	kW	Air	-34	34-57	57-	-34	34-57	57-
Peugeot 206	139.90	3.83	44	no	88	422	197	55	185	199
Citroën C3	149.90	3.85	44	no	32	141	69	12	84	56
Toyota Yaris	204.00	3.88	62	no	121	611	490	76	310	376
VW Polo	141.59	3.90	40	no	117	521	300	55	251	235
Nissan Micra	170.40	3.72	59	yes	17	122	102	5	45	74
Opel Corsa	175.00	3.82	55	no	30	122	117	16	68	94
Peugeot 406	249.90	4.60	85	yes	34	178	29	134	708	368
Ford Mondeo	244.20	4.73	81	yes	25	144	39	113	912	313
Volvo V70	437.17	4.71	132	yes	13	92	16	55	624	150
Mazda 6	239.90	4.68	88	yes	35	199	41	159	864	496
Mercedes E	539.60	4.85	120	yes	2	78	27	58	537	241
Toyota Avensis	244.40	4.52	81	yes	79	373	74	226	1529	836

3. THE ECONOMETRIC MODEL

The random-coefficients logit method developed by Berry et al. (1995) employs market-level data to estimate realistic demand elasticities for a large number of products, while allowing for endogenous prices. I use an extended version of their model, where separate taste parameters are estimated for different consumer groups. This chapter discusses the structural model. The first two sections give an overview of the mechanics of discrete choice and logit models. The third section shows how to represent taste heterogeneity more realistically with random coefficients. In the fourth section I review a method for dealing with endogenous prices in discrete choice models. The fifth section discusses ways to include demographic information in market-level models. The supply side assumptions are covered in the sixth section, and the final specification of my model in the last section.

3.1 Discrete choice and logit

In a market of several products with interrelated demands, the sale of each good is a function of the prices of all the other goods. To estimate a demand system, a demand function for a representative consumer with all the prices as arguments can be specified for each product. All such models require the estimation of at least as many parameters as there are products, and many more in models that yield realistic substitution patterns. A well-known specification is the constant elasticities model

$$log(q_j) = \alpha_j + \sum_k \eta_{jk} log(p_j) + \epsilon_j,$$

where η_{jk} is the elasticity of good j with respect to the price of good k. With J products, this model requires the estimation of J^2 elasticities. The markets analysed here have close to two hundred different products, precluding direct estimation of a market-level demand function. In a market with highly differentiated products like the car market, it also seems necessary to explicitly take consumer heterogeneity into account. Although not impossible in principle, representation of consumer heterogeneity is generally not feasible in this type of models, because it would further increase the number of parameters.

3.1.1 Discrete-choice models

An alternative to the market-level demand functions is to project the range of products onto a space of product attributes.[1] A utility function is defined on this characteristics space. Discrete-choice models assume that each consumer chooses the one product that maximises utility, U (ties are assumed not to occur). The researcher does not observe utility, but can specify a representative utility function, $\delta_j = \delta(x_j)$, where x_j is a vector of characteristics of product j observed by the econometrician. δ depends on parameters that need to be estimated. Utility is decomposed as $U_j = \delta_j + \epsilon_j$, where the error term is simply the difference between true utility and representative utility. A good model approximates U as closely as possible by δ, and specifies a distribution for ϵ that captures the remaining error

[1] Anderson, de Palma, and Thisse (1989) discuss the relationship between these two ways of modelling demand.

well. Let $P(\epsilon)$ denote the joint distribution of $\epsilon = (\epsilon_1, \ldots, \epsilon_J)$, and let $A_j = \{\epsilon; \forall l \neq j, \delta_j + \epsilon_j > \delta_l + \epsilon_l\} = \{\epsilon; \forall l \neq j, U_j > U_l\}$. The probability (not conditional on ϵ) that a consumer choose product j is

$$s_j = Pr(\epsilon \in A_j) = \int_{A_j} dP(\epsilon). \tag{3.1}$$

$P(\epsilon)$ is interpreted as the distribution of the unobserved part of utility over consumers. That is, it represents variation in the unobserved factors in the population. s_j is the share of consumers who choose alternative j. Equation (3.1) defines a demand function $s_j = s(x_j)$ with observed product characteristics as arguments. Finding this demand function poses two challenges. The first is to estimate the parameters of the function $\delta_j = \delta(x_j)$ which determines A_j. This requires observations of x_j and market shares, S_j (note that a capital s is used for observed market shares). Estimation generally proceeds by matching observed market shares, S_j, and predicted market shares, s_j. The second challenge is to solve the integral over $P(\epsilon)$. The logit is a discrete-choice model that has a closed solution for this integral.

3.1.2 Logit

The logit model is defined by a utility specification as above, $U_j = \delta_j + \epsilon_j$, and the assumption that each ϵ_j is distributed independently, identically extreme value. The extreme value cumulative distribution function is $F(\epsilon_j) = e^{-e^{-\epsilon_j}}$. McFadden (1973) shows that with this distributional assumption on ϵ, the choice probability in equation (3.1) is

$$s_j = \int_{A_j} dP(\epsilon) = \frac{exp(\delta_j)}{\Sigma_l exp(\delta_l)}. \tag{3.2}$$

In most cases representative utility is specified to depend linearly on observed characteristics, i.e. $\delta_j = x_j\beta$ where β is a k-vector when there are k observed characteristics. This gives a market share function $s_j(x_j) = \frac{exp(x_j\beta)}{\Sigma_l exp(x_l\beta)}$. The logit market share function has a number of desirable properties. Market shares of all products included in the analysis is strictly between zero and one. The sum of market shares equals one. s_j approaches one when δ_j increases with other δs held constant, and approaches zero when δ_j decreases.

3.1.3 *Substitution patterns in the logit model*

The question of how demand changes when product characteristics change or when new products are introduced is one of the main issues adressed by discrete choice demand analysis. I will look at changes in price in particular, but the discussion generalises to other characteristics. The logit assumptions imply a particular substitution pattern. This can easily be seen from the form of the market share function in (3.2). The ratio of market shares of two products, j and l, is $s_j/s_l = exp(\delta_j)/exp(\delta_l)$. Since this ratio only depends on the characteristics of the two alternatives considered, the relative market shares are independent of all other products in the market. This property is called independence from irrelevant alternatives (IIA). As an illustration, consider an expensive Mercedes and a much cheaper Fiat. Suppose a new, expensive BMW is introduced in the market. One would expect that this new alternative would affect the Mercedes' sales more than the Fiat's, thereby increasing the ratio s_{fiat}/s_{merc}. Instead the IIA says that this ratio remains constant, implying that if Fiat and Mercedes have the same market share, just as many consumers substitute from Fiat to BMW as from Mercedes to BMW.

The price elasticities further illuminate this problem. Let observed utility be $\delta_j = -\alpha p_j + x_j \beta$, where p_j is price and x_j other characteristics of product j. The price elasticities of the market shares are

$$\frac{p_k}{s_j}\frac{\partial s_j}{\partial p_k} = \begin{cases} -\alpha p_j(1-s_j) & \text{if } k=j \\ \alpha p_k s_k & \text{otherwise} \end{cases}$$

Consider the example of the cars again (after the introduction of the BMW). If the price of the BMW goes up, the market shares of Fiat and Mercedes will see the same percentage change: $\alpha p_{bmw} s_{bmw}$. This is unrealistic. It is likely that the Mercedes is a much closer substitute for the BMW than the Fiat. The fact that the logit model a priori restricts this substitution effect to be the same, is a serious limitation. The problem is caused by the way consumer variation enters utility. Consumers choose their product because it gives them higher utility than any other product. When a characteristic is altered in this product, consumers substitute to other products if utility drops below that of the alternatives. In the logit model, consumers do rank the products differently, but this difference is entirely due to the error term. Consider a particular BMW-buyer, consumer c, and let the original observed utility be the same for all three models. Since BMW is picked when the observed valuation is the same, the consumer's unobserved valuation of the BMW must be larger than for Fiat or Mercedes. No further conclusions can be drawn about the realisation of the unobserved utility for this particular consumer. $\epsilon_{c,merc}$ and $\epsilon_{c,fiat}$ are realisations of independently and identically distributed random variables and so the latter could very well be larger than the former, leading to a switch from BMW to Fiat. Since the ϵ's are identically distributed, the proportion of BMW-buyers who rank Mercedes as second-best is just equal to the proportion of consumers in the population at large who rank Mercedes as their second-best. This means that although a consumer buys a particular type of car, according to the model, the likelihood that his second-choice is a similar car is no higher than it is for the average consumer.

3.2 Extensions of the logit model

The logit model represents variations in consumer taste, but in a highly simplistic way. The previous discussion demonstrates that to model substution patterns realistically, the unobserved portion of utility must be correlated across products. The model used for the empirical investigation in this thesis is the random-coefficients logit, which is treated in the next section. Here I briefly discuss two other approaches that have been used for demand analysis of differentiated products, the vertical differentiation model and the nested logit.

In the vertical differentiation model (Shaked and Sutton (1982) and Bresnahan (1987)) utility is $U_j = (x_j\beta)\nu - p_j$, where p_j is price and ν is a random term that reflects differing valuation of quality in the population. In the logit model the problem is that substitution effects do not depend on how close products are in characteristics space. The vertical differentiation model goes to the other extreme, allowing for substitution only between models that are neighbours in characteristics space.

Contrary to the vertical differentiation model, the nested logit (McFadden 1978) allows for substitution between all products. It also improves on the simple logit by letting substitution effects be higher between similar products. Products are divided into mutually exclusive groups. In a nested logit model of the European car market, Verboven (1996) uses class (small, large, luxury etc.) and country of origin to place cars in two levels of product groups. A general model can be specified as follows. Utility is $U_j = x_j\beta + \zeta_g + (1-\sigma)\epsilon_j$, where ζ_g is a random variable whose distribution depends on σ, with $\sigma \in [0,1)$. ζ_g is common to all products in group g. The parameter σ is estimated. When it has a value close to one, the correlation between utility of products in the same group is close to one. As σ approaches zero, correlation patterns approach those of the simple logit. This model has a closed solution of the integral in equation (3.1) and allows for substitution effects that are stronger between similar products. It has therefore found widespread applications (see for instance Goldberg (1995), Goldberg and Verboven (2001), and Ivaldi and Verboven (2004)). The problem is that the substitution patterns to some extent are imposed by the researcher through the grouping of products. The model allows for stronger substitution effects within groups, but exhibits IIA between groups.

3.3 Random-coefficients logit

The models considered so far suffer from the restrictions they impose on substitution patterns. Random-coefficients logit (RCL) models are much less restrictive.[2] The main drawback of this class of models is that the integral in the market share equation does not have a closed form solution. Estimation therefore requires the use of computationally demanding simulation techniques.

3.3.1 Taste variation

In the logit model, consumer heterogeneity enters utility only through the iid error term. The fact that consumers choose different cars is in other words caused by white noise. The vertical differentiation model and the nested

[2] See McFadden and Train (2000) for a discussion of the flexibility of RCL.

logit let variation in choice depend on characteristics, but only in a very restrictive way. In fact, consumer tastes vary along many dimensions. Some people look for a small, fast car and do not mind paying more to get the one they prefer. Others want a big car, but are very concerned with price. To take this multidimensional taste variation into account, RCL models interact the observed product characteristics with some form of variation across consumers. Utility is specified as $U_j = x_j\beta + \epsilon_j$, where the error term is iid extreme value as before. The βs are now random variables, where the distribution of β_k reflects the distribution of the taste for characteristic k in the population. A parametric family of distributions is assumed for the βs, and the goal of the estimation is to find the parameters of the distribution. The random variables β and ϵ are independent.

I will now derive the market share function for the RCL. Each consumer is formally defined by a realisation of the random vector (β, ϵ). Let $A_j = \{(\beta, \epsilon); \forall l \neq j, x_j\beta + \epsilon_j > x_l\beta + \epsilon_l\} = \{(\beta, \epsilon); \forall l \neq j, U_j > U_l\}$. A consumer chooses product j if and only if his or her realisation of (β, ϵ) is an element of A_j. Let $P(\beta, \epsilon)$ denote the distribution of (β, ϵ). The proportion of consumers who choose product j, i.e. the market share, is the probability that (β, ϵ) fall within the area A_j:

$$\begin{aligned}
s_j &= \int_{A_j} dP(\beta, \epsilon) \\
&= \int_{A_j} dP(\epsilon) dP(\beta \mid \epsilon) \\
&= \int \left[\int_{A_j|\beta} dP(\epsilon) \right] dP(\beta) \\
&= \int \frac{exp(x_j\beta)}{\Sigma_l x_l \beta} dP(\beta) \qquad (3.3)
\end{aligned}$$

The second equality follows from Bayes' rule, the third from the independence of β and ϵ, and the last from equation (3.2). A mixed function is a weighted average of several functions, with the density function that provides the weights called the mixing distribution. The RCL market share function is a mixture of the logit market share function evaluated at different values of the taste parameters, with $dP(\beta)$ as the mixing distribution. The fundamental structure of the choice process is still the logit, but the variation in taste now allows different consumers to have different logit choice probabilities. For the logit model, I illustrated the IIA property in terms of the ratio of market shares for two products. There, the denominators of the market share function cancelled, leaving the ratio to depend only on the characteristics of the two cars considered. For the RCL, this is no longer the case. Let $s_j(\beta)$ be the market share of product j conditional on β. The ratio of market shares of the products j and k is now

$$\frac{\int \frac{s_j(\beta)}{\Sigma_l s_l(\beta)} dP(\beta)}{\int \frac{s_j(\beta)}{\Sigma_l s_l(\beta)} dP(\beta)}.$$

Because of the integral, the ratio depends on all the data. In the logit, since the ratio was constant, changes in the characteristic of a third product had the same percentage effect on j and k. Now that the ratio is no longer fixed, the effects on market shares of changes in one product will vary across the other products.

To estimate the model, the parameters of the distribution of β need to enter explicitly. Let $\beta_k = \overline{\beta}_k + \sigma_k \nu_k$, where ν_k is an iid standard normal random variable representing the taste variation in the population for characteristic k. (This is the specification in Berry et al. (1995).) If price does not have a random coefficient, utility can now be written $U_j = -\alpha p_j + x_j \overline{\beta} + \Sigma_k \sigma_k x_{jk} \nu_k + \epsilon_j = \delta_j(\nu) + \epsilon_j$. The derivative of $\delta_j(\nu)$

with respect to p_j is $-\alpha$. Conditional on ν, market shares are $s_j \mid \nu = s_j^* = e^{\delta_j(\nu)}/\Sigma_l e^{\delta_l(\nu)}$. Still conditional on ν,

$$\frac{\partial s_j^*}{\partial p_j} = \left(\frac{e^{\delta_j}}{e^{\Sigma_l \delta_l}} - \frac{e^{\delta_j}}{(e^{\Sigma_l \delta_l})^2} e^{\delta_j}\right)\left(\frac{\partial \delta_j}{\partial p_j}\right) = -\alpha s_j^*(1 - s_j^*)$$
$$\frac{\partial s_j^*}{\partial p_k} = \left(-\frac{e^{\delta_j}}{(e^{\Sigma_l \delta_l})^2} e^{\delta_k}\right)\left(\frac{\partial \delta_k}{\partial p_k}\right) = \alpha s_j^* s_k^*.$$

The area that is integrated over with respect to β in the market share function does not depend on price, so the differential operator in the price derivatives can be moved under the integral sign. The price elasticities of the market shares are

$$\frac{p_k}{s_j}\frac{\partial s_j}{\partial p_k} = \begin{cases} -\frac{p_j}{s_j}\int \alpha s_j^*(1 - s_j^*) dP(\nu) & \text{if k = j} \\ \frac{p_k}{s_j}\int \alpha s_j^* s_k^* dP(\nu) & \text{otherwise.} \end{cases}$$

The formula for cross-price elasticities reveals how the RCL substitution patterns depend on taste variation. If two products j and k have similar characteristics, choice probabilities will be high for the same values of ν. This correlation means that the term $s_j^* s_k^*$ is on average higher than $s_j^* s_l^*$ where l is a very different product from j. The result is a higher cross-price elasticity between j and k than between j and l.

3.3.2 Simulation

In spite of the realistic representation of substitution patterns, there is a feature of the RCL that in many applications is a big disadvantage. The integrals in equation (3.3) and in the price elasticities do not in general have closed form solutions. One way to approximate the integral is to take *ns* draws, $(\beta_1, \ldots, \beta_{ns})$, from the distribution of β and let

$$\check{s}_j = \frac{1}{ns} \sum_{i=1}^{ns} \frac{exp(x_j \beta_i)}{\sum_l exp(x_l \beta_i)}. \tag{3.4}$$

Usually estimation of these models needs to be done by numerically optimising a likelihood or GMM-function. This means that the integral in the market share must be simulated for every product in every iteration of the optimisation algorithm, substantially increasing computation time, as well as complicating the software implementation of the estimation routine.

3.4 Demographic information

The various ways of representing heterogeneity in consumer tastes that have been discussed do not pose the question of where the heterogeneity comes from. Consumers have different tastes because people are different. Accordingly, various approaches have been developed where taste depends on quantifiable characteristics of the consumers, such as age, income, sex, and family size. This section first reviews a couple of examples from the literature that use market-level data and then discusses the way my own model incorporates demographic information.

3.4.1 Two examples from the literature

In a series of papers on the US ready-to-eat cereal industry, Nevo lets the random coefficients depend on income and age (Nevo 2000a) and (Nevo 2001). Let D_d denote element d of the vector of demographic variables. The coefficients are $\beta_k = \overline{\beta}_k + \sigma_k \nu_k + \sum_d \pi_{kd} D_d$. The distribution of the demographics is approximated by the empirical distribution of random draws of D from a population survey. This method lets the shape of the distribution $P(\beta)$ be affected by the distribution of consumer characteristics. For instance, if the distribution of taste for sugary cereals is skewed towards high sugar content, and this is partly due to an age distribution that is skewed towards young age, when people tend to like sweets more, Nevo's model will give a better fit.

In a study of the introduction of the minivan on the American car market, Petrin (2002) uses data that provide more information on the connection between demographics and product choice. Whereas Nevo's demographic data do not actually link consumers and choices, Petrin has information on the average demographics for buyers of each product. The utility specification is similar to Nevo's. The coefficients are $\beta_k = \overline{\beta}_k + \sigma_k \nu_k$. For some classes of cars - like the minivan - the σ depends on demographics: $\sigma_{k,minivan} = \sum_d \overline{\sigma}_{kd} \log D_d$. Like in the model above, the demographic data that enter utility are population survey data. The innovation here is primarily in the estimation. Petrin uses a moment constraint that matches average values of the Ds for households buying a particular class of car to the prediction of the model. In addition, like Berry et al. (1995), Petrin uses a simulated distribution of income based on the population income distribution. It enters utility as $\log(y - p_j)$. Furthermore, separate coefficients are estimated for three different income groups to reflect varying utility of money.

3.4.2 A model with sales by consumer group

My data contain sales and technical characteristics for all car models marketed in Norway in the period 2000-2003. Sales are given separately for men and women divided in three age groups, giving a total of six different demographic groups. This allows me to estimate separate distributions of the taste coefficients for different consumer groups. In contrast, the existing literature has used demographic information to better approximate the *average* distribution of the taste coefficients. This model should give a better fit, since tastes do seem to vary significantly between groups. In addition, the resulting elasticities can be used to investigate certain policy issues where distinctions between these groups are crucial.[3]

Utility is specified as $U_j = \alpha \log(y - p_j) + x_j\beta + \epsilon_j$. Different price coefficients are estimated for men and women, and different means of the βs for all six groups. That is,

$$\alpha = \sum_{d=1,2} \alpha_d I_d, \tag{3.5}$$

and

$$\beta_k = \sum_{d=1,..,4} \overline{\beta}_{kd} I_d + \sum_k \sigma_k x_{jk} \nu_k. \tag{3.6}$$

The ν_ks are iid standard normal and $d = 1, .., 4$, indicating "woman", "man", "young" and "old", respectively. I_d is a dummy variable indicating whether the observation has the demographic characteristic d. Note that six consumer groups are defined by combinations of the four demographic characteristics d (no age dummy means middle age). Effectively, the coefficients for each of the characteristics have a separate distribution in each of the six consumer groups. Standard deviations are assumed to be the same across consumer groups for each characteristic, but means are free to vary over groups. For example, $\beta_{k,youngwoman} \sim N(\overline{\beta}_{k,1} + \overline{\beta}_{k,3}, \sigma_k)$, where β_{kg} is the random coefficient on characteristic k for consumer group g.

[3] The application in Section 6 requires elasticities by consumer group.

3.5 Endogenous prices

It is unlikely that a discrete-choice model can include all product characteristics that affect consumer choice. Cars differ in ways that are inherently hard to quantify, such as style, status and durability. Besides, the number of characteristics that matter to consumers is most likely too large for it to be possible to estimate the separate effects of each of them. Although these characteristics are not observed by the econometrician, they are observed by manufacturers. Since they affect demand, they will also influence pricing. This means that prices are endogenous: Because of correlation between price and characteristics that are valued by consumers, but not controlled for in the model, demand curves can turn out to be upward-sloping. If demand has a linear form, as in equation (3.1), this problem can be handled by instrumental variables methods. In discrete-choice models, however, prices and characteristics enter the demand function in a nonlinear way, precluding the direct application of traditional instrumental variables techniques.

Berry (1994) proposes a method for solving the endogeneity problem in a broad class of discrete-choice models. The models above need to be modified by the addition of a utility term that captures unobserved characteristics. In the general formulation of utility, a term ξ_j is added:

$$U_j = -\alpha p_j + x_j\beta + \xi_j + \epsilon_j \tag{3.7}$$

Mean utility, the portion of utility that is the same for all consumers in the market, is now

$$\delta_j = -\alpha p_j + x_j\beta + \xi_j, \tag{3.8}$$

where β is fixed. A second modification that needs to be made, is the introduction of an outside good. Some measure, M, of the number of potential consumers in the market is assumed to be available. I will discuss how to obtain this number in section 4.1, but for now I assume it is observed. Observed market shares are defined as $S_j = q_j/M$, where q_j is sales. The outside good, indexed as 0, has market share $S_0 = (M - \sum_{l=1}^{J} q_l)/M$. Observed utility derived from the outside good is normalised to zero. The presence of the outside good therefore ensures that a general increase in prices reduces overall demand for the inside goods.

If the model is true, the market share function gives the true market shares. This defines the true value of mean utility:

$$S_j = s_j(\delta_j), \ \ j = 0, \ldots, J. \tag{3.9}$$

Berry (1994) proves an existence and uniqueness result for mean utility. Under weak regularity conditions on the density of consumer unobservables and for every possible S, there exists a unique δ^* that satisfies equation (3.9). This result means that (3.9) defines the mean utility as a function of observed market share: $\delta = s^{-1}(S)$. For the RCL, the function does not have a closed form, so mean utility must be solved for numerically. This will be discussed in section 4.2. In the simple logit, a closed form expression can be found. The fact that $\delta_0 = 0$, is essential. Using the logit market share function in (3.2) yields

$$\begin{aligned}\log(s_j) - \log(s_0) &= \log\left(\frac{e^{\delta_j}}{\sum_l e^{\delta_l}}\right) - \log\left(\frac{e^{0}}{\sum_l e^{\delta_l}}\right) \\ &= \delta_j.\end{aligned}$$

This, together with equation (3.8), produces a linear demand equation

$$\log(s_j) - \log(s_0) = -\alpha p_j + x_j\beta + \xi_j. \tag{3.10}$$

Regarding the unobserved characteristics term as a simple error term, we now have an equation that can be estimated with an instrumental variables regression.

3.6 The supply side

There are two reasons to specify a supply side when estimating a demand system. First, all applications that involve the computation of a new price equilibrium in a policy experiment require estimates of marginal costs. Secondly, the pricing rule depends on the true values of the demand parameters. Simultaneous estimation of demand and supply therefore improves efficiency. The disadvantage of using a supply side is that it requires more structure, in the form of an equilibrium assumption.

Marginal cost is assumed to depend on a vector of product characteristics w_j and a vector of unobservables ω_j:

$$\log(mc_j) = w_j\gamma + \omega_j. \tag{3.11}$$

The profit of a firm f that produces a subset F_f of the J products is

$$\Pi_f = \sum_{j \in F_f} (p_j - mc_j) M s_j(p) - C_f, \tag{3.12}$$

where M is the market size, C_f is fixed cost, and p is the price vector. A pure-strategies Bertrand-Nash equilibrium of strictly positive prices is assumed to exist.[4] The equilibrium first-order conditions are, for each p_j

$$s_j(p) + \sum_{k \in F_f} (p_k - mc_k) \frac{\partial s_k(p)}{\partial p_j} = 0. \tag{3.13}$$

Define a $J \times J$ matrix Ω, with elements

$$\Omega_{jk}(p) = \begin{cases} -\frac{\partial s_j(p)}{\partial p_k} & \text{if } \exists f : \{k, j\} \subset F_f \\ 0 & \text{otherwise.} \end{cases} \tag{3.14}$$

The first-order conditions in (3.13) can now be written on vector form:

$$s(p) - \Omega(p)(p - mc) = 0.$$

This defines the markup function

$$b(p) = p - mc = \Omega(p)^{-1} s(p). \tag{3.15}$$

Together with (3.11), this gives the pricing equation that can be used for estimation:

$$ln(p - b(p)) = w_j \gamma + \omega_j. \tag{3.16}$$

My data have separate observations of sales to six different consumer groups, indexed by $g = 1, \ldots, 6$. Each of these groups is treated as a distinct market, but all markets are served by the same companies and prices are uniform across markets. This requires a slight modification of the equations above. Profit is

$$\Pi_f = \sum_{j \in F_f} \left((p_j - mc_j) \sum_g M_g s_{gj}(p) \right) - C_f, \tag{3.17}$$

[4] Caplin and Nalebuff (1991) prove the existence of equilibrium for single-product firms. Like Berry et al. (1995) I assume that their result extends to multiproduct firms.

where M_g is the size of market g. First-order conditions for profit maximisation are now, for each p_j

$$\sum_g M_g s_{gj}(p) + \sum_{k \in F_f} \left((p_k - mc_k) \sum_g M_g \frac{\partial s_{gk}(p)}{\partial p_j} \right) = 0. \tag{3.18}$$

For each market g, the price derivatives matrix Ω_g is defined as in (3.14), but using group market shares, s_{gj}. Define $\Omega^*(p) = \sum_g M_g \Omega_g(p)$ and $s^*(p) = \sum_g M_g s_g(p)$. In the same way as in the one-market case, this gives a markup function

$$b(p) = \Omega^*(p)^{-1}(p - mc)s^*(p). \tag{3.19}$$

This is then used in the pricing equation (3.16) as before.

3.7 Final specification

This chapter has provided the theoretical background for the econometric model that I use to estimate a demand system for the Norwegian car market. This last section presents the final specification. The primitives of the model are the functional forms for consumer preferences and for marginal cost, and an equilibrium assumption. Utility has the form $u_j = \alpha \log(y - p_j) + x_j\beta + \xi_j + \epsilon_j$, where α and β are defined in (3.5) and (3.6). Correspondingly, the outside good has utility $u_0 = \delta_0 + \epsilon_0$, where $\delta_0 = \alpha \log(y) + \xi_0 + \sigma_0 \nu_0$. The mean utility of the outside good is normalized to zero by letting $U_j = u_j - \delta_0, j = 0, \ldots, J$. This yields the final specification of utility:

$$\begin{aligned} U_j &= \alpha \log(\frac{y - p_j}{y}) + x_j\beta + \xi_j \\ &= \alpha \log(1 - p_j/y) + x_j\beta + \xi_j, \end{aligned} \tag{3.20}$$

where I have abused notation slightly, since ξ_j and β do in fact also change. The notation has been left unchanged because these changes are immaterial, as ξ_j is unknown in any case and β is still normal, with unchanged and unknown mean and unknown variance. Observations from each of the six consumer groups enter the estimation as observations from different markets. The pricing equation is (3.16), repeated for convenience,

$$ln(p - b(p)) = w_j\gamma + \omega_j.$$

with the markup function $b(p)$ given by (3.19).

4. DATA AND ESTIMATION

4.1 Data

I use two types of data: Observations on the sales and characteristics of cars, with sales linked to demographic group, and summary statistics of the demographic groups.

4.1.1 The car sales data

The sales and characteristics data have been provided by the Information Council for Road Traffic in Norway. I have yearly sales of new cars, for every model marketed in Norway in the four-year period 2000-2003. Sales are recorded separately for six demographic groups: Women aged less than 34 years, women aged 34-57, women aged above 57, men aged less than 34, men aged 34-57, and men aged above 57. Modern cars come in many variants of each model, where by model I refer to all cars with the same principal model name, e.g. "Volkswagen Golf". It is common for one model to be marketed with more than thirty different combinations of engine size, transmission, body, or fuel type. The sales statistics in my data record sales for many different variants of each model, where the variant is characterised by technical specifications such as cylinder volume, power, and body type. It turns out that in spite of the many alternatives, sales of most models are highly concentrated in one variant, usually the most basic (with the smallest engine, sedan etc.). I have chosen to simplify the analysis by aggregating the sales of different variants of the same model, and to use the technical specifications of the variant with the highest sales. This clearly entails the loss of some information, but two considerations favour the choice that I have made: First, the fact that the variant with the highest sales in most cases sells much more than any other variant of the same model, means that the loss of information should not be too important. The second consideration is conditional on the available econometric models. The assumption of an iid error term on utility could be difficult to uphold if products that only differed by a tiny amount were included as separate products. Say, a three-door Golf with a 1.6 litre engine and a three-door Golf with a 1.9 litre engine are so similar in terms of idiosyncratic preferences, that the iid assumption is difficult to defend. The characteristics recorded are length, weight, power (kW), fuel consumption (l/km) and a dummy for whether air conditioning is standard. I use data for fuel prices to calculate fuel expenses, kr/km, where kr denotes Norwegian kroner. As a measure of acceleration, I use power over weight, kW/kg. The characteristics vector used in the estimation is $x = (length, kr/km, acc, air, const)$. It attempts to capture the following features: size, safety, space (length), costs, environment-friendliness (kr/km), acceleration (acc), luxury (air). The characteristics vector for marginal cost is $w = (\log(length), \log(acc), \log(l/km), air, const)$.

The sales statistics do not contain price information. To match prices with sales, I used a price list with technical specifications. The use of list prices instead of transaction prices is problematic. And it is of particular importance when I consider the issue of price discrimination. Using list prices amounts to assuming that everybody pays the same price. However, it is well known that buying a car frequently involves bargaining over price, the inclusion of extra equipment, the price offered for a part exchange of an old car, etc. The observed sales are outcomes of this bargaining, and it could be that some price discrimination does in fact occur there, whereas my analysis treats price discrimination as a counterfactual experiment. Nevertheless, list prices are usually the only option with market-level data. Besides, it is likely that correct price records would not fundamentally change the results, as deviation from list prices, even in the presence of bargaining, is limited.

4.1.2 Demographic data

Some additional information on the characteristics of the six demographic groups was obtained from Statistics Norway. For market sizes, I simply use the total number of persons belonging to the relevant demographic group in each year. It is a somewhat arbitrary choice. Attemps to identify the number of people who actually consider buying a car, however, will usually result in numbers that are endogenous to the analysis. Berry et al. (1995) use the total number of households. In any case, my results are very robust to changes in the estimates of market size. (The same is reported by Ivaldi and Verboven (2004), who use multiples of total sales).

A simulated income distribution is used in the utility specification. I have mean and standard deviation of income for persons in full-time employment in each of the demographic groups. The data were obtained from Statistics Norway, and have been computed from a large sample of tax forms. This is not entirely in keeping with the measure of market size, since I there use total population, and the proportion of people in full time employment varies between these groups. For instance, since many people in the highest age group are retired, using income information from the people who are actually in full-time employment could tend to overstate income for this group. On the other hand, many people in this age group also have high savings. Income enters utility as y in $\log(1 - p_j/y)$. Because cars are expensive in Norway, using income from only one year resulted in some negative observations of $1 - p_j/y$, making the log undefined. I therefore multiplied income means and standard deviations by three. It seems reasonable to use income for a longer period than a year, since people do not buy a new car every year.[1] The results are very robust to changes in the income specification.

4.2 Estimation

The estimation routine, which follows Berry et al. (1995), can be divided into several stages. Mean utility is recovered by matching observed and predicted market shares. Demand-side unobservables are then expressed as a function of mean utility and observed characteristics. The pricing assumption and the predicted market shares are used to solve for marginal cost. Together with the product characteristics, marginal cost defines the cost-side unobservables. The unobservables can now be regarded as functions of the parameters (the data are fixed). A set of orthogonality conditions between the unobservables and functions of the observed product characteristics is used to form a GMM objective function. The estimates are the parameter values that minimise the objective. This section discusses the procedure in more detail.

A remark on notation is in order. I observe four years and six demographic markets in each year, giving a total of twenty-four markets. Some variables are the same for all groups within a year, such as prices and product characteristics. Other variables, like market shares or income moments, vary across demographic groups as well as years. For simplicity, this depen-

[1] Variation is possibly slightly lower over a longer period, but I have no further information on this. I tried to vary standard deviations, but the final results differ only by negligible amounts even by large changes in income variation.

dence has been supressed in the notation. Variables are simply subscripted j if they are product specific, whether they vary over groups or not.

4.2.1 The simulation estimator for market shares

As discussed in section 3.3, the integral in the market-share function does not have a closed form. To compute market shares in the estimation procedure, we must therefore resort to a simulation estimator for the integral. The market share of product j is given by

$$s_j = \int \frac{\delta_j + \alpha \log(1 - p_j/(e^{\overline{m}+\overline{\sigma}\nu_y})) + \sum_k \sigma_k x_{jk}\nu_k}{\sum_l \delta_l + \alpha \log(1 - p_l/(e^{\overline{m}+\overline{\sigma}\nu_y})) + \sum_k \sigma_k x_l \nu_k} dP(\nu), \tag{4.1}$$

where $P(\nu)$ is the distribution function of the vector $\nu = (\nu_y, \nu_1, \ldots, \nu_K)$. $\overline{m}$ and $\overline{\sigma}$ are functions of the sample moments of income in the relevant demographic group. These functions transform the observed mean and standard deviation of income to parameters for the lognormal distribution which give it the same mean and standard deviation. For each market, I draw $ns = 50$ random numbers for each element of ν in each market. The draws are taken from an iid standard normal distribution. Random draws are indexed by i. The simulator for the integral is

$$\check{s}_j = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\delta_j + \alpha \log(1 - p_j/(e^{\overline{m}+\overline{\sigma}\nu_{yi}})) + \sum_k \sigma_k x_{kj}\nu_{ki}}{\sum_l \delta_l + \alpha \log(1 - p_l/(e^{\overline{m}+\overline{\sigma}\nu_{yi}})) + \sum_k \sigma_k x_{lk}\nu_{ki}}. \tag{4.2}$$

Using this formula, the derivatives of market shares can be found analytically for use in the pricing equation. Berry et al. (1995) discuss how to build a simulator with a smaller variance, and Berry, Linton, and Pakes (2004) provide asymptotic distributions for the simulation based estimators in this class of models. The exact properties of the estimators are beyond the scope of this thesis, however, and I compute standard errors without taking simulation error into account. (Berry et al. (1995) report that correcting for simulation error increases standard errors by 5-20%).

4.2.2 A contraction mapping for mean utility

The key to the estimation of this model is the method for uncovering mean utility developed in Berry (1994) and Berry et al. (1995). Berry proves that the equation

$$S_j = \check{s}_j(\delta_j), \; j = 0, \ldots, J \tag{4.3}$$

has a unique solution. This defines the mean utility as a function of observed market share: $\delta = \check{s}^{-1}(S)$. However, the function does not have a closed form, so mean utility must be solved for by an iterative procedure. The function

$$f(\delta_j) = \delta_j + \log(S_j) - \log\{\check{s}_j(\delta_j)\} \tag{4.4}$$

is a contraction mapping (Berry et al. 1995). A general property of contraction mappings is that they converge by recursive iteration. That is, if $m(\cdot)$ is a contraction mapping, the sequence $(z^1, z^2, \ldots)$ defined by $z^{h+1} = m(z^h)$, converges to a value z^H with the property that $z^H = z^{H-1}$. This means that for every j, the sequence defined by the contraction mapping in (4.4) converges to a value δ_j^H such that $f(\delta_j^H) = \delta_j^H$. It follows that $S = \check{s}(\delta^H)$. Using the result that, δ^*, the solution to (4.3) for a given S, is unique, we get $\delta^H = \delta^*$. In other words, δ^* can be found by recursively solving the contraction mapping in (4.4). Implementing the recursive iteration in a programming language is relatively straightforward. It is worth mentioning, however, that this introduces a new iterative process which will have to *converge* at every iteration of the minimisation process. The cost in terms of computation time is significant.

4.2.3 The residuals

Mean utility is defined as $\delta_j = x_j\overline{\beta} + \xi_j$, where $\overline{\beta} = \sum_d \overline{\beta}_d I_d$. (Price is not part of the portion of utility that is common to all consumers, since it enters nonlinearily with a random income term.) In the definition, mean utility contains an unobserved part. The goal of the method above was to recover mean utility from observed market shares in spite of the unobservable. Using the definition of mean utility and the recovered value , we can now solve for the demand-side unobservable:

$$\xi_j(\sigma, \alpha, \overline{\beta}) = \delta^*(\sigma, \alpha) - x_j\overline{\beta}. \tag{4.5}$$

This is the residual that enters the objective function. Note that the value of δ^* as defined by (4.3) depends only on the parameters σ and α, since $\overline{\beta}$ and γ do not enter $\check{s}$.

There is also a residual from the supply side. The simultaneous estimation of supply and demand improves efficiency because the pricing rule depends on the demand parameters, and therefore imposes further constraints on the minimisation. The pricing equation is $\log(p_j - b_j(p)) = w_j\gamma + \omega_j$, where $b(p)$ is as defined in section 3.6, but using the simulation estimator, $\check{s}(\delta^*) = \check{s}\{\delta^*(\sigma, \alpha)\}$, for market shares. Making this dependence explicit (and dropping the p in the argument), the supply side residual is

$$\omega_j(\sigma, \alpha, \gamma) = log\{p_j - b_j(\sigma, \alpha)\} - w_j\gamma. \tag{4.6}$$

The residual for product j, ω_j, is the same across groups within a year, since price and marginal cost only varies across years.

4.2.4 Instruments

In section 3.5 it was established that demand unobservables are correlated with price. Similarly, markups are probably correlated with cost unobservables in the pricing equation. The reason is again that price depends on unobserved (cost) characteristics. Since $b(p)$ is a function of price, it cannot be assumed to be uncorrelated with ω. This endogeneity problem is the reason that estimation of the demand and pricing equations uses moments restrictions instead of a maximum likelihood procedure. The procedure is to specify a list of variables z, that are mean independent of ξ and ω. The independence assumption is that the demand and supply unobservables are mean independent of both observed product characteristics and cost shifters (Berry et al. 1995). Formally

$$E[\xi_j \mid z] = E[\omega_j \mid z] = 0, \tag{4.7}$$

where $z_j = (x_j, w_j)$ and $z = (z_1, \ldots, z_J)$. The residuals expressed by (4.5) and (4.6) are functions of the parameters. At the true values of the parameters, $\theta = \theta_0$, these functions reproduce the true values of the unobserved characteristics. Therefore, (4.7) implies that at θ_0, the residuals are uncorrelated with any function of z.

The functions that are used as instruments are defined as follows. For each element k of the vector z_j, where product j is produced by firm f, the entries in the instrument matrix are the characteristic itself, the sum of the same characteristic over all other cars produced by the same company, and the sum of the characteristic over all cars produced by rival firms. That is

$$z_j, \sum_{r \neq j, r \in F_f} z_r, \sum_{r \notin F_f} z_r. \tag{4.8}$$

This yields a 15×1 instrument vector Z_j^x for x_j and a 15×1 vector Z_j^w for w_j. The intuition behind the choice of instruments and their correlation with price hinges on the oligopoly pricing rule[2]. The extent to which a car has close neighbours in product space is correlated with the sum of the characteristics of other products. Products with close substitutes, will tend to have low markups and thus low prices relative to cost. Since the markups depend on products produced by the same firm in a different way from those in the same firm, an ownership distinction is made in the instruments.[3]

4.2.5 Interactions with dummy variables

The linear part of the demand side residual is $x_j\overline{\beta}$, with $\overline{\beta} = \sum_d \overline{\beta}_d I_d$. I is the D-vector of dummy variables for the demographic characteristics. If the dummies are regarded as variables in their own right, this is a nonlinear specification. Still, it can be turned into a linear function of the data by defining a new $1 \times KD$ vector of variables, $\tilde{x}_j = x_j \otimes I$:

$$
\begin{aligned}
x_j \sum_d \overline{\beta}_d I_d &= \sum_k \sum_d \overline{\beta}_{dk} x_{jk} I_d \\
&= \sum_k \sum_d \overline{\beta}_{dk} \tilde{x}_{jk}^d \\
&= \tilde{x}_j \beta
\end{aligned} \tag{4.9}
$$

where β now denotes the KD-vector $(\overline{\beta}_1', \ldots, \overline{\beta}_D')'$. The characteristics vector now has $KD = 5 \cdot 4 = 20$ entries, while the instruments defined above sum up to $5 \cdot 3 = 15$. This means that the instruments vector does not satisfy the rank condition for identification of the parameters. The problem

[2] For a formal discussion on optimal instruments in this model, see (Berry et al. 1995).

[3] I treated all manufacturers owned by the same company as one firm. The number of firms is eighteen. The ownership structure was obtained from www.pommert.de/virtualia/garage.

is solved by interacting the instruments with the dummy variables. The new instruments vector is $\tilde{Z}_j^x = Z_j^x \otimes I_s'$, where I_s is the vector of dummy variables for men and women.

4.2.6 The objective function

With the instrument vectors $\tilde{Z}_j^x$ and Z_j^w the sample moment constraint is written as

$$\sum_j \begin{bmatrix} \tilde{Z}_j^x & 0 \\ 0 & Z_j^w \end{bmatrix} \begin{bmatrix} \xi_j(\theta) \\ \omega_j(\theta) \end{bmatrix} \equiv \sum_j Z_j r_j(\theta) = 0. \tag{4.10}$$

The GMM objective function is given by

$$\Big[\sum_j Z_j r_j(\theta)\Big]' \mathbf{W} \Big[\sum_j Z_j r_j(\theta)\Big]. \tag{4.11}$$

The weighting matrix is $\mathbf{W} = \sum_j \left[Z_j \hat{\hat{r}}_j\right]\left[Z_j \hat{\hat{r}}_j\right]'$, where the fitted residuals $\hat{\hat{r}}_j$ were obtained from a preliminary estimation with weighting matrix $\sum_j Z_j Z_j'$. The asymptotic covariance matrix of the estimates is

$$\left\{\Big[\sum_j Z_j \nabla_\theta r_j(\hat{\theta})\Big]' \mathbf{W} \Big[\sum_j Z_j \nabla_\theta r_j(\hat{\theta})\Big]\right\}^{-1}, \tag{4.12}$$

where ∇_θ denotes derivatives with respect to θ (Wooldridge 2002).

4.2.7 Optimisation

As can be seen from (4.5) and (4.6), the residuals have one part that is linear in the parameters, and one where they enter in a nonlinear fashion. What is more, one set of the parameters, $\theta_1 = (\beta, \gamma)$, enters only the linear part, whereas the remaining parameters, $\theta_2 = (\sigma, \alpha)$, enter only the nonlinear part. This means that the residual vector for product j, $r_j(\theta)$, can be written

$$r_j(\theta) = L_j\theta_1 + N_j(\theta_2) \tag{4.13}$$

with

$$L_j = \begin{bmatrix} -x_j \\ -w_j \end{bmatrix}, \; N_j(\theta_2) = \begin{bmatrix} \delta^*(\theta_2) \\ ln\{p_j - b(\theta_2)\} \end{bmatrix}.$$

Using the first-order conditions for minimisation of (4.11), $\hat{\theta}_1$ can be expressed as a function of the nonlinear parameters $\hat{\theta}_2$:

$$\hat{\theta}_1 = \Big\{\Big[\sum_j Z_j L_j\Big]' \mathbf{W} \Big[\sum_j Z_j L_j\Big]\Big\}^{-1} \Big[\sum_j Z_j L_j\Big]' \mathbf{W} \Big[\sum_j Z_j N_j(\hat{\theta}_2)\Big]. \tag{4.14}$$

The nonlinear search can now be limited to θ_2. To sum up, the steps of the estimation process are:

1. Temporary values of mean utility δ, and temporary values of σ and α are used to compute utility of every product for each simulated consumer.
2. Use the simulated utilities to compute predicted market shares, $\check{s}(\delta)$.
3. Match predicted market shares with observed market shares. The value of δ that equates the two is found by iteratively solving the contraction mapping, starting over from step 1 for each iteration until convergence.
4. Solve for new values of β and γ as a function of σ and α.
5. δ and β are used to solve for ξ.
6. δ is used to compute the markups implied by the market shares.
7. ω is found using the markups and γ.

8. The value of the objective function is computed, using ξ, ω and the instruments.

9. Let a minimisation algorithm suggest new values for σ and α, and repeat the process until the objective converges to a minimum.

The estimation algorithm was implemented in Matlab, using a simplex search method for the minimisation of the objective function.[4] I wrote the Matlab code myself. The minimisation took approximately three hours to converge. Standard errors were computed analytically.

[4] Gradient methods, both analytical and finite-differences, tended to go outside the parameter space and I could not make them converge. This is possibly a feature of RCL-models with a supply side, for Berry et al. (1995) use a simplex routine, whereas Nevo (2000b) (who does not have a supply side) is able to use a gradient method.

5. ESTIMATION RESULTS

This section presents the estimated parameters from four different utility specifications. The first is a simple logit model, the second a logit model with instrumental variables, while the third and fourth are versions of the random-coefficients model discussed in the previous chapter.

5.1 Logit and IV logit

The logit model can be estimated by a regression on the pricing equation $\log(s_j) - \log(s_0) = \alpha p_j + \tilde{x}_j \beta + \xi_j$, as discussed in section 3.5. $\tilde{x}_j$ is the vector of characteristics interacted with demographic dummies (see section 4.2). The IV logit estimates the same equation, but with an instrumental variables regression. The instruments are the same as in the full RCL model. Heteroscedasticity-robust standard errors are reported.

Starting at the top of table 5.1, price coefficients are much larger in the IV specification. This confirms the suspicion that prices are endogenous, and it indicates that the instruments work. Products with higher unobserved quality sell at higher prices. When this is not taken into account, the disutility of price seems smaller. In both specifications, the coefficient on price is higher for men than for women. This is perhaps somewhat surprising, given that men buy more expensive cars than women on average. However, the advantage of this type of demand modelling is precisely that it uncovers the role that each characteristic plays. Looking further down the list, it turns out that women have a lower valuation of size and acceleration, and a greater disutility from fuel expenses. These preferences all point in the direction of smaller, cheaper cars.

Younger and older people prefer smaller cars than middle-aged people of the same sex. Like the price terms, estimates of the length coefficients seem to improve when price endogeneity is accounted for. Acceleration is most valued by young men, in accordance with the stereotype, and more valued by men than by women. It should be noted that the model does not have distinct parameters for young men, young women etc. This means that the young coefficients represent an average deviation of the tastes of young men and women from those of middle-aged men and women respectively. It seems reasonable to expect that age works in the same direction for both sexes, but this assumption could be a limitation in some cases. Apart from this, there is a general problem with the precision of the age parameters, whereas fuel and air are the only variables that do not have significant parameters for the main groups men and women. Berry et al. (1995), too, find that the fuel parameter becomes insignificant when instruments are introduced. The coefficients on air-conditioning are all insignificant and appear to have the wrong sign. Presumably there is a problem with the strong correlation between air-conditioning (high standard) and fuel consumption (big engine) on one side and price on the other side.

The outside good has utility zero. In a sense, the negative constants for cars is the starting point that a car has to climb from by its characteristics to gain positive utility among some consumers. This starting point is higher for women, perhaps indicating a higher utility from simply having a car, independent of its characteristics. On the other hand, men place much higher value on certain characteristics (most notably size). In fact, more men than women end up with positive utility.

Tab. 5.1: LOGIT RESULTS, 3900 OBSERVATIONS

		OLS $ln(s_j/s_0)$ on x	s.e.	2SLS $ln(s_j/s_0)$ on x	s.e.
Price	Men	-0.327	0.033	-0.917	0.118
	Women	-0.213	0.027	-0.758	0.101
Length	Men	1.075	0.178	2.103	0.266
	Women	-0.267	0.166	0.683	0.235
	Young	-0.360	0.185	-0.360	0.196
	Old	-0.377	0.194	-0.377	0.205
Fuel	Men	-0.210	0.050	-0.016	0.069
	Women	-0.214	0.045	-0.034	0.063
	Young	0.101	0.052	0.101	0.071
	Old	0.065	0.055	0.065	0.073
Acc	Men	0.053	0.044	0.413	0.114
	Women	-0.029	0.039	0.305	0.098
	Young	0.023	0.047	0.023	0.093
	Old	-0.005	0.048	-0.005	0.093
Air	Men	0.078	0.160	-0.120	0.171
	Women	-0.055	0.159	-0.238	0.167
	Young	-0.091	0.185	-0.091	0.194
	Old	-0.149	0.196	-0.149	0.202
Const	Men	-12.838	0.698	-18.916	1.375
	Women	-7.565	0.656	-13.179	1.200
	Young	-0.104	0.691	-0.104	0.763
	Old	1.241	0.725	1.241	0.794

5.2 The full model

For the full model, the estimated parameters and a sample of the resulting price elasticities are reported.

5.2.1 Coefficients

The utility specification that determines demand is $u_j = \alpha \log(1 - p_j/y) + x_j\beta + \xi_j$, where β is a vector of normally distributed random coefficients. Means on all characteristics are estimated separately for each sex, with coefficients for age groups as above. Standard deviations are estimated for each characteristic. Demand is estimated jointly with a pricing equation derived from the assumption of a Bertrand-Nash equilibrium.

Precision does not appear to be a problem in the full model, but a few remarks are in order. First, standard errors have not been corrected for simulation noise and bias, but it is unlikely that doing so would dramatically change the picture. Secondly, the joint estimation of supply and demand improves efficiency by a substantial amount. This improvement is welcome as long as a Nash equilibrium in prices is a good representation of the real market structure, but it is a strong structural assumption.

The coefficients on the price term are now positive, since they reflect the value of money. Men value money more, confirming the results from the logit model. Looking at the means first, length is again more important for men than for women, and the difference is even larger than in the simpler model. Young people value length more, and older people value it less. The coefficient on fuel expenses is negative for both men and women, slightly more for women. It is not clear why both young and old people are much less concerned with fuel consumption. Acceleration is much more important to men and young people, whereas old people do not deviate from the middle group in their taste for this characteristic. The coefficient on air-conditioning still has the wrong sign, indicating that this variable captures something else than high standard of equipment, possibly some of the price variation. As before, the constant is lower for men than women. The supply side parameters are estimated jointly with the demand side. Size matters most for marginal cost.

Of the random coefficients standard deviations, all but the one on air-contitioning are significant. The disutility of fuel consumption exhibits the largest variation in relation to its mean. For instance, according to the

Tab. 5.2: RESULTS FROM THE FULL MODEL, 3900 OBSERVATIONS

			Estimate	Standard error
Price coefficients (α)		Men	136.897	4.916
		Women	64.212	5.322
Std. deviations (σ)		Length	0.301	0.042
		kr/km	0.299	0.018
		Acc	0.029	0.011
		Air	0.092	0.118
		Constant	0.447	0.120
Means ($\overline{\beta}$)	Length	Men	3.521	0.112
		Women	1.100	0.133
		Young	0.420	0.119
		Old	-0.486	0.110
	kr/km	Men	-0.298	0.047
		Women	-0.366	0.046
		Young	0.214	0.051
		Old	0.240	0.044
	Acc	Men	1.278	0.059
		Women	0.694	0.067
		Young	0.708	0.073
		Old	-0.075	0.061
	Air	Men	-0.543	0.090
		Women	-0.462	0.089
		Young	-0.415	0.111
		Old	-0.196	0.107
	Constant	Men	-27.872	0.523
		Women	-16.201	0.679
		Young	-6.854	0.475
		Old	1.135	0.434
Cost parameters (γ)				
		$\log(Length)$	3.348	0.330
		$\log(Acc)$	0.908	0.054
		$\log(l/km)$	1.285	0.095
		Air	-0.092	0.024
		Constant	-23.799	2.130

numbers, fifteen percent of men actually experience no disutility by high fuel consumption. The taste for length too varies substantially. Thirty percent of women have a length coefficient lower than 0.8 or higher than 1.4. Still, the overall impression is that the means conditional on demographic group capture most of the taste variation. The length coefficients illustrate this well. The coefficient for men has a much higher mean than the women's coefficient - 3.5 versus 1.1. The coefficients are both normally distributed with a standard deviation of 0.3. This means that the density functions cross

Tab. 5.3: RESULTS WITHOUT GROUP-SPECIFIC MEANS, 3900 OBSERVATIONS

		Estimate	Standard error
Price coefficient (α)		106.974	2.513
Std. deviations (σ)	Length	0.432	0.391
	kr/km	0.615	0.242
	Acc	0.005	0.054
	Air	0.020	0.451
	Constant	1.217	0.308
Means ($\overline{\beta}$)	Length	2.443	0.342
	kr/km	-0.774	0.341
	Acc	1.414	0.050
	Air	-0.755	0.060
	Constant	-25.422	0.330
Cost parameters (γ)			
	$\log(Length)$	3.015	0.114
	$\log(Acc)$	0.951	0.037
	$\log(l/km)$	1.234	0.076
	Air	-0.035	0.015
	Constant	-21.748	0.642

at a distance of four standard deviations from each of the means. In other words, intersex variation is much stronger than intrasex variation. Table 5.3 reports the results of the random-coefficients model without group-specific means. The means appear to be close to the averages of the group-specific means in the full model, as could be expected. The random coefficients now have higher variance than in the full model. For length, fuel and the constant, standard deviations increase by between 50 and 200 percent. This indicates that sex and age do capture a substantial part of the taste variation in the population.

Tab. 5.4: A SAMPLE OF PRICE ELASTICITIES FOR MEN AGED 35-57 YEARS, 2003

	Peugeot[a] 206	Citroën C3	Toyota Yaris	VW Polo	Nissan Micra	Opel Corsa	Peugeot 406	Ford Mondeo	Volvo V70	Mazda 6	Mercedes E-class	Toyota Avensis
206	-214.10	0.20	1.10	0.55	0.12	0.19	3.21	3.92	5.52	3.85	6.00	6.60
C3	0.41	-229.92	1.09	0.55	0.12	0.19	3.18	3.88	5.46	3.81	5.93	6.53
Yaris	0.45	0.22	-313.81	0.59	0.13	0.21	3.46	4.21	6.00	4.16	6.51	7.11
Polo	0.40	0.19	1.06	-216.63	0.12	0.19	3.11	3.80	5.32	3.73	5.79	6.39
Micra	0.41	0.20	1.08	0.54	-262.02	0.19	3.15	3.85	5.42	3.78	5.89	6.48
Corsa	0.42	0.20	1.12	0.56	0.12	-269.19	3.26	3.97	5.62	3.91	6.10	6.69
406	0.46	0.22	1.23	0.62	0.14	0.22	-384.18	4.42	6.22	4.34	6.77	7.42
Mondeo	0.45	0.22	1.19	0.60	0.13	0.21	3.51	-374.43	6.01	4.21	6.55	7.22
V70	0.51	0.24	1.35	0.67	0.15	0.24	3.95	4.80	-685.98	4.74	7.46	8.10
6	0.48	0.23	1.26	0.63	0.14	0.22	3.71	4.53	6.39	-367.38	6.95	7.61
E	0.51	0.25	1.37	0.68	0.15	0.24	4.00	4.87	6.94	4.80	-857.83	8.20
Avensis	0.45	0.22	1.20	0.60	0.13	0.21	3.52	4.30	6.04	4.22	6.58	-371.82

[a] Cell entries i,j, for row and column respectively, give one hundred times the percentage change in the market share of i with a one percent change in the price of j.

Tab. 5.5: A SAMPLE OF PRICE ELASTICITIES FOR WOMEN AGED 35-57 YEARS, 2003

	Peugeot[a] 206	Citroën C3	Toyota Yaris	VW Polo	Nissan Micra	Opel Corsa	Peugeot 406	Ford Mondeo	Volvo V70	Mazda 6	Mercedes E-class	Toyota Avensis
206	-147.64	0.14	0.88	0.49	0.14	0.15	0.33	0.26	0.32	0.36	0.34	0.66
C3	0.39	-158.73	0.87	0.49	0.14	0.14	0.33	0.26	0.32	0.35	0.34	0.66
Yaris	0.41	0.15	-217.29	0.50	0.14	0.15	0.34	0.27	0.33	0.37	0.36	0.69
Polo	0.39	0.14	0.86	-149.37	0.14	0.14	0.32	0.26	0.31	0.35	0.34	0.66
Micra	0.39	0.14	0.86	0.48	-181.08	0.14	0.32	0.25	0.31	0.35	0.34	0.65
Corsa	0.40	0.14	0.88	0.49	0.14	-186.10	0.33	0.26	0.32	0.36	0.34	0.67
406	0.43	0.15	0.95	0.53	0.15	0.16	-269.06	0.28	0.35	0.39	0.37	0.72
Mondeo	0.42	0.15	0.93	0.53	0.15	0.16	0.35	-262.72	0.34	0.38	0.37	0.72
V70	0.44	0.16	0.99	0.55	0.15	0.16	0.37	0.29	-486.41	0.40	0.39	0.74
6	0.43	0.15	0.96	0.54	0.15	0.16	0.36	0.28	0.35	-257.84	0.38	0.73
E	0.45	0.16	1.00	0.55	0.16	0.16	0.37	0.29	0.36	0.41	-611.36	0.75
Avensis	0.42	0.15	0.93	0.52	0.15	0.15	0.35	0.28	0.34	0.38	0.36	-262.50

[a] Cell entries i,j, for row and column respectively, give one hundred times the percentage change in the market share of i with a one percent change in the price of j.

5.2.2 Elasticities

Elasticities can be computed for all characteristics in addition to price. Here the focus is on pricing, with given product characteristics, so I concentrate on price elasticities. Profit maximising prices are on the elastic part of the demand curve. The own-price elasticities computed from the simulated market share function using the parameter estimates, are all greater than one in absolute value.

Tables 5.4 and 5.5 show a sample of estimated price elasticities (multiplied by one hundred for the sake of readability) for middle-aged men and women in 2003. The sample consists of six cars especially popular among women, and six cars especially popular among men. The most notable feature of the elasticities is that the effect of a price change for a given car is quite similar across the other products. To an extent, this is reasonable. Elasticities with respect to a price change in product j depend on the market share of product j. If a product has a large market share, a price rise of a given size means that many consumers will substitute away. Conversely, a very small market share means that even if a large proportion of consumers move to other products, their number will be insignificant. In the simple logit model, all products have the same cross-price elasticity with respect to product j. This means that the rows in the tables would be identical apart from the elements on the diagonal. As was discussed in Chapter 3, the random-coefficients model used here relaxes this constraint. When the pattern of the logit model is still clearly recognisable, it is partly due to the reduction in the variation of the random-coefficients that follows from estimating group-specific means. After all, the simple logit is exactly the random-coefficients logit with standard deviations equal to zero.

Another pattern is that cars with higher price (see Chapter 2) have higher elasticities. This is reasonable in the sense that the same percentage price change amounts to more money for an expensive car. However, in many cases this means that cross price elasticities are higher between cars that are far apart in product space than products that are close. This outcome is not necessarily unrealistic, since market shares also play an important role, but here it seems to be too prevalent. Possibly, this phenomenon is due to limitations in the data. Since I only observe four years, the changes in the choice set are perhaps too small to identify the full extent of taste variation. Berry et al. (1995) use twenty years, and get higher estimates for the standard deviations. Their cross-price elasticities are more in accord with a priori considerations of closeness in characteristics space. On the other hand, it is perhaps questionable whether tastes in cars can be assumed to remain the same over their sample period.

The differences between men and women are clear. As could be expected from their relatively small price coefficient, women have lower own-price elasticities. For women, the effects of a price rise in one of the typical male cars is much smaller than for men (generally less than a tenth). For the typically female cars, on the other hand, they have about the same elasticities as men, in spite of the difference in the price coefficient. This means that consumers in a given group have relatively high elasticities for cars that are popular in that group. So, in spite of the limitations observed above, a large part of the expected substitution patterns is captured by the group-specific elasticities.

6. PRICE DISCRIMINATION

This chapter discusses the effects of allowing car manufacturers to price discriminate between female and male customers. It is likely that discrimination would lead to some pro-forma changes in demand: Households would invariably register their cars in the name of a person of the low-price sex. This effect is ignored in the analysis that follows. It could be argued that gender-based price discrimination is politically, as well as practically, infeasible. Still, the experiment provides an example of a method that could be applied to more realistic cases of price discrimination. More importantly, it demonstrates empirically a phenomenon that has been discussed in the theoretical literature. The simulated discriminatory price equilibrium is presented in the last section. Before that, I briefly discuss some relevant results from the literature on price discrimination in oligopoly.

6.1 Oligopoly price discrimination

Most of the literature on price discrimination deals with the monopoly case. When a monopolist is given the opportunity to price discriminate, the optimisation problem is the same as before, but with the constraint of a uniform price removed. Accordingly, profit must be (weakly) higher when discrimination is allowed. In a noncooperative game, on the other hand, firms could possibly be worse off with a larger choice set.

Holmes (1989) analyses a simple oligopoly model with two firms producing different products (one each). The set of potential consumers are partitioned into two groups, the "weak" market and the "strong" market. Discrimination will raise the price in the strong market and lower it in the weak market. Using a strong assumption of symmetric demand for the two products, Holmes shows what factors determine the change in industry output when discrimination is allowed. Discrimination will tend to raise output when the weak market has a relatively high industry demand elasticity (tendency to substitute to the outside good) compared to the cross-price elasticity, or when strong-market demand is strongly concave compared to weak-market demand. Holmes also finds that profit can decrease with discrimination. When the weak market has higher cross-price elasticity, but lower industry elasticity than the strong market, a price rise in the weak market would be better for industry profit. Still, because the higher cross-price elasticity outweighs the industry demand elasticity in the weak market, the price goes down there, while it goes up in the strong market, where total profit suffers more. Armstrong and Vickers (2001) analyse this model with Hotelling demand. Corts (1998) shows how all prices can fall if firms rank consumers differently in terms of demand elasticities, i.e. if firms have different "strong" markets. While uniform prices tend to isolate firms from competition in their respective strong markets, the ability to price discriminate can trigger price wars in every market.

Concerning the analysis here, the main lesson of this literature is that discrimination can intensify competition in the market with the highest cross-price elasticities, and that the effect on total profit is a priori ambiguous.

Tab. 6.1: Comparison of uniform-price and discriminatory equilibria, 2003

		Women	Men	Total
Mean % price change		7.74	-3.21	
Mean price[a] (unweighted)	Without discr.	318.65	318.65	318.65
	With discr.	337.91	311.80	324.86
	Change	19.26	-6.84	6.21
	% change	6.04	-2.15	1.95
Mean price (weighted)	Without discr.	221.48	259.33	248.22
	With discr.	237.70	251.01	247.77
	Change	16.21	-8.32	-0.44
	% change	7.32	-3.21	-0.18
Output	Without discr.	17,107	41,159	58,266
	With discr.	14,143	43,973	58,116
	Change	-2,963	2,814	-149
	% change	-17.33	6.84	-0.26
Profit	Without discr.	1,236,253	2,851,950	4,088,203
	With discr.	1,280,483	2,805,881	4,086,365
	Change	44,229	-46,068	-1,838
	% change	3.58	-1.62	-0.04

[a] Prices and profits are in 1000 kr. Weighted mean prices are weighted by sales.

6.2 The discriminatory equilibrium

When price discrimination between women and men is allowed, the link between the two markets is broken. Firms now have two optimisation problems instead of one. The solution to each problem is derived in the same way as with uniform pricing. I repeat it here for convenience. For each market g, define a $J \times J$ matrix Ω_g, with elements

$$\Omega_{gjk}(p) = \begin{cases} -\frac{\partial s_{gj}(p)}{\partial p_k} & \text{if } \exists f : \{k, j\} \subset F_f \\ 0 & \text{otherwise.} \end{cases}$$

F_f denotes the set of products owned by firm f. The first-order conditions can now be written on vector form:

$$s^*(p) - \Omega^*(p)(p - \widehat{mc}) = 0, \tag{6.1}$$

where $\Omega^*(p) = \sum_g M_g \Omega_g(p)$ and $s^*(p) = \sum_g M_g s_g(p)$. The subscript g indexes consumer group. Marginal cost is estimated by using the cost side parameters from the estimation of the RCL model. For each of the two markets, the equilibrium price vector solves (6.1), with g going through the three age groups for women and men respectively. I solved (6.1) using a numerical zero-finding technique.[1]

Table 6.1 summarises the results of allowing discrimination in the 2003 car market. The first row shows that the average increase in prices offered to women is almost eight percent, while the prices offered to men on average go down by a little more than three percent. Women have lower own- and cross-price elasticities. When prices are set separately for women and men, this

[1] I used Matlab's fsolve routine. Convergence was rapid from widely different starting values.

means that competition is less intense in the women's market, and the resulting equilibrium prices higher. The next four rows show how the unweighted mean price of cars changes when companies start to discriminate. When prices change, consumers make new choices, and so market shares change. The sales-weighted means reflect this adaptation. The overall mean price paid is almost unchanged, while the mean price paid by women increases by seven percent and that for men decreases by three percent. Total output hardly changes. It falls slightly more for women than it goes up for men, giving a larger percentage change for women. Women have lower price elasticities, but the women's price rise is sufficiently larger than the men's reduction, for women's demand to change by more than men's. Finally, the profit figures show how increased competition in the men's market forces prices down and reduces overall profit. In the women's market, manufacturers can now take advantage of lower elasticities to extract a higher surplus. The effects for women and men nearly cancel each other. The net change in profit is negligible. Still, the fact that it does have a negative sign, while the changes in each market are substantial, illustrates the point made in the literature: The effect on total profit of allowing price discrimination is ambiguous and depends on the exact shape of demands.

Table 6.2 shows how prices change for a sample of cars. (The sample is the same as in Chapter 2 and Chapter 5.) The first six cars are models with relatively high market shares among women, and the final six are models particularly popular with men. In Chapter 5, the tables of elasticities showed that women have relatively high elasticities for the cars popular among women, and vice versa for men. Looking at the percentage changes in table 6.2, the overall impression is that prices change more for the models that are less popular: For men, the typically female cars become around ten

Tab. 6.2: Uniform and discriminatory prices for selected cars

	Uniform price	Discriminatory price		Percentage change	
		Women	Men	Women	Men
Peugeot 206	139.90	153.37	124.71	9.63	-10.86
Citroën C3	149.90	162.49	135.65	8.40	-9.51
Toyota Yaris	204.00	215.64	190.58	5.70	-6.58
VW Polo	141.59	153.27	127.17	8.25	-10.18
Nissan Micra	170.40	181.45	155.06	6.49	-9.00
Opel Corsa	175.00	187.84	161.14	7.34	-7.92
Peugeot 406	249.90	273.69	246.44	9.52	-1.38
Ford Mondeo	244.20	266.94	241.54	9.31	-1.09
Volvo V70	437.17	458.13	434.81	4.79	-0.54
Mazda 6	239.90	262.81	236.79	9.55	-1.30
Mercedes E	539.60	562.91	537.39	4.32	-0.41
Toyota Avensis	244.40	265.59	241.10	8.67	-1.35

percent cheaper, whereas the typically male cars drop about one percent in price. For women, the picture is less clear cut, but the typically male cars tend to rise more in price than the typically female ones. The reason for this pattern is most likely the following: For the cars that are mostly bought by men, the uniform price is already set almost without regard to women's preferences, and therefore it remains almost unchanged for men when discrimination is introduced. The cars mostly bought by women (and not so many men), however, have a uniform price that depends a lot on women's preferences. Therefore, a lot of scope remains for changing it in the men's market when effects in the women's market no longer need to be taken into account.

7. CONCLUSION

This thesis shows that if price discrimination between men and women were allowed in the Norwegian car market, industry profits would go down. The possibility of this outcome has been put forth in the theoretical literature on oligopoly third-degree price discrimination (Holmes 1989), but appears never to have been demonstrated empirically before. The reduction in profit is admittedly very small, but this is of less importance. My results show that in spite of substantial changes in equilibrium prices to adapt to the new environment, car companies are slightly worse off when they are able to discriminate. For a monopolist, price discrimination simply represents a way to extract more surplus from consumers. The crucial difference in oligopoly is that the removal of the uniform price constraint breaks the cross-market links that restrain aggressive pricing in the market with the highest price elasticities. When lowering prices in the high-elasticity market no longer reduces profits in the low-elasticity market, firms start to compete more fiercely in the high-elasticity market - in this case so much so that total profit suffers. As is well known from noncooperative game theory, having a larger choice set is not necessarily beneficial.

To estimate demand functions I use the random-coefficients logit method developed by Berry et al. (1995). I use a new type of data and a modification of their model to estimate separate taste parameters for different sex and age groups. This permits the investigation of certain issues, like price discrimination, where differences in demand functions between groups is crucial. In addition, it improves the fit of the model by capturing an important source of taste variation. This is of importance in other applications, such as merger analysis or policy predictions. The method is particularly valuable since this type of data could easily be made available in most countries (car buyers are usually required to submit their date of birth and name (sex) when registering a new car).

BIBLIOGRAPHY

ANDERSON, S. P., A. DE PALMA, AND J.-F. THISSE (1989): "Demand for Differentiated Products, Discrete Choice Models, and the Characteristics Approach," *Review of Economic Studies*, 56(1), 21–35.

ARMSTRONG, M., AND J. VICKERS (2001): "Competitive price discrimination," *RAND Journal of Economics*, 32, 579–605.

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.

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.

BRESNAHAN, T. F. (1987): "Competition and Collusion in the American Auto Industry: The 1955 Price War," *Journal of Industrial Economics*, 35, 457–482.

CAPLIN, A., AND B. NALEBUFF (1991): "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 59(1), 25–59.

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

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

Goldberg, P., and F. Verboven (2001): "The Evolution of Price Dispersion in the European Car Market," *Review of Economic Studies*, 68, 811–848.

Holmes, T. J. (1989): "The Effects of Third-Degree Price Discrimination in Oligopoly," *American Economic Review*, 79(1), 244–250.

Ivaldi, M., and F. Verboven (2004): "Quantifying the Effects from Horizontal Mergers in European Competition Policy," *International Journal of Industrial Organization*.

McFadden, D. (1973): "Conditional Logit Analysis of Qualitative Choice Behaviour," in *Frontiers of Econometrics*, ed. by P. Zarembka. Academic Press.

——— (1978): "Modelling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlqvist. North-Holland.

McFadden, D., and K. Train (2000): "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics*, 15(5), 447–470.

Nevo, A. (2000a): "Mergers with differentiated products: the case of the ready-to-eat cereal industry," *RAND Journal of Economics*, 31(3), 395–421.

——— (2000b): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 9(4), 513–548.

——— (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69(2), 307–342.

Petrin, A. (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110.

Shaked, A., and J. Sutton (1982): "Relaxing Price Competition Through Product Differentiation," *Review of Economic Studies*, 49, 3–13.

Verboven, F. (1996): "International Price Discrimination in the European Car Market," *RAND Journal of Economics*, 27, 240–268.

Wooldridge, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press.