Household Gasoline Demand in the U. S.: An Estimation of a Multiple-Discrete/Continuous Choice Model

Zava Aydemir*

February 14, 2002

Abstract: This paper presents a multiple-discrete/continuous choice model. In this model consumers maximize utility by choosing the number and types of vehicles, several different fuel types and the quantity of gasoline. We estimate the model using household-level data. Compared to the traditional literature that uses aggregate data our estimates on own price elasticities are substantially lower and insignificant. Whereas time-series studies report elasticities of around -0.24 the corresponding figures in our study range between -0.17 and 0.16. Price variations virtually play no role when taking the multiple-discrete feature of gasoline demand into account.

Keywords: Multiple-Discrete/Continuous Choice Model, Product Differentiation, Gasoline Demand, Nested Logit

JEL Classification: C5, D1, Q4

---

*Affiliations: Zava Aydemir, Socioeconomic Institute, University of Zurich, Hottingerstr. 10, 8032 Zurich, Switzerland, zaydemir@soi.unizh.ch.
I would like to thank Johannes Binswanger and Stefan Bühler for useful comments.

1 Introduction

From a theoretical and econometric point of view, gasoline has some distinct features which make it an interesting subject for demand analysis. Gasoline is a complementary good to vehicles owned. Moreover, it is a differentiated product. Further, households make multiple-discrete choices by spreading their purchases over different types of gasoline. At the same time households naturally consume a continuous amount of each type they have selected.

The aim of the paper is twofold. It first develops a theoretical model of household demand that captures the multiple-discrete/continuous feature of gasoline consumption. Besides, it accounts for the complementarity between vehicle ownership and gasoline consumption. Second, in this paper we estimate the model by using micro-level data on households in the United States. The econometric specification of the model thereby takes the multiple-discrete/continuous feature of the gasoline demand fully into account.

The gasoline demand model exactly mirrors the discrete/continuous feature in gasoline consumption: It consists of a discrete choice part and of a continous choice part. According to the data used in this study, the majority of households own more than one car. Further, they usually drive each of their car by at most one fuel type. We thus model the discrete choice part as follows. The household chooses her vehicle stock. This decision process consists of the choice of the number of vehicles and of the choice of the vehicle type for each car. The vehicle type can either be a passenger car or a light truck. Additionally, the household selects for each of her car exactly one fuel type. In our estimation, we consider three different fuel types, namely regular, midgrade and premium gasoline.[1] The direct link in the model between the choice of the number of vehicles and the choice of the fuel type for each of the vehicles allows one to capture the multiple-discreteness in gasoline type selection. For instance, a household with two cars may purchase premium gasoline for its first car and regular gasoline for its second car.

In the continuous part of the model, the household chooses a continuous amount for each selected gasoline type. The multiple-discrete choice influences gasoline demand in two different ways. First, the characteristics of the cars

[1] By dropping households consuming leaded gasoline or diesel from the sample, I model fuel type choice as a matter of consumer tastes and not of the technical pecularities of the vehicles.

chosen by a household determine her gasoline consumption. For example, an old light truck with many cylinders consumes *ceteris paribus* more gasoline than a new passenger car with a high fuel efficiency. Second, households' heterogeneity regarding fuel types choice influences the price they face in the gasoline demand, i.e. a household having selected premium gasoline pays more for a driven mile than another household who drives regular gasoline.

In the econometric specification, the discrete choice part of the model is developed within the random utility framework. It represents household preferences over products as a funtion of individual household characteristics and product characteristics. The model assumes that the choice of the vehicle stock (i.e. the number and types of cars) mainly relates to household characteristics whereas the choice of the fuel types mainly depends on the characteristics of the cars owned and on fuel prices. The discrete choice alternatives are thereby conveniently categorized to allow for multiple-discrete choices. In the model, the choice sets consist of mutually exclusive alternatives which, each of them, contain a vehicle stock and fuel types for each of the cars in the vehicle stock.

This work relates to various contributions in the literature analyzing the discrete/continuous choice problem. McFadden and Dubin (1984) model the choice of appliances and the demand for electricity by appliance. Train, McFadden and Ben-Akiva (1987) analyze the discrete choice of local telephone services and the number of calls. Cardon and Hendel (2001) estimate a structural model which consists of a discrete health insurance choice and a continuous health care consumption. Hanemann (1984) provides a theoretical discrete/continuous choice model and discusses the corresponing estimation strategies. These papers share the common feature that only one alternative is chosen among mutually exclusive alternatives. The most important difference between this paper and the aforementioned studies is that it allows many alternatives to be chosen at the same time. That is, households are modelled to have the possibility to make multiple-discrete choices. The main contribution of this study is that, by combining multiple-discrete and continuous choice, zero expenditures for some commodities under consideration can be appropriately dealt with. The modelling of such a consumption behavior still seems to be "one of the ... pressing [problems] in applied demand analysis (Deaton, 1984, p. 1809)". The model generates consumption patterns for every utility maximizing household in which the non-negativity constraint is binding for some fuel types and the consumed quantities are strictly positive for alternative fuel types.

From the point of view of modelling consumption behavior, my contribution comes closest to Hendel's (1999). He estimates firms' demand for personal computers by considering the fact that buyers purchase multiple brands and at the same time multiple units of each brand that they have chosen. In contrast to my contribution, he models the latter part likewise as a discrete choice. Further, he estimates firm behavior within the random coefficients framework which puts less structure on demand behavior than the nested logit model employed in this paper. However, the assumption of generalized extreme value distribution on the error term requires a lower computational burden for the estimation and allows to write the probabilities in closed form which substantially simplifies the analysis.

This paper provides results on gasoline demand in the United States from estimating the model using houshold-level data. This contrasts to the traditional literature. Most of the work in this literature, surveyed by Dahl and Sterner (1991) and Espey (1998), is largely based on aggregate data in form of various time-series.[2] However, as Schmalensee and Stoker (1999) show in their study, individual household and demographic characteristics play an essential role in gasoline demand.[3] Whereas their focus is particularly on the impact of income on gasoline consumption, this analysis puts weight on determining the role of fuel prices in gasoline consumption with simultanously controlling for its multiple-discrete/continuous feature. In this way, the paper provides genuine *ceteris paribus* results on the price effects on fuel consumpton, in contrast to the traditional work, in which due to data aggregation the essential multiple-discrete/continuous feature is not captured.

Own price elasticities produced by our model are substantially lower in absolute terms than the short-run elasticities estimated in various time-series models. In those models price elasticity equals to around -0.24 according to Dahl and Sterner. Depending on whether further controlling for household and vehicle characteristics besides the multiple-discrete choice feature, estimated values for price elasticities range between -0.17 and 0.16. That is, elasiticity even may become positive (but highly insignificant). This particularly arises when households' residence and the number of cars owned are included in the model – in this case, price effects become virtually negligible. These results conform to intuition since controlling for vehicle stock hold and for demographic characterstics makes households become less flexible in adjusting their driving behavior to price changes.

The paper is organized as follows. Section 2 describes the data set. In section

[2] Exceptions are the contributions of Archibald and Gillingham (1980, 1981), Hausman and Newey (1995) and Jorgenson, Sleznick, and Stoker (1988).

[3] The authors estimate gasoline demand using the same data as this paper, but from previous years.

3, the main features of the discrete choice part of gasoline demand is presented by means of descriptive statistics. Section 4 introduces a consumer demand model that is able to generate multiple-discrete/continuous consumption patterns in the context of gasoline demand. Section 5 presents the econometric specification of the model. Section 6 provides the results of the discrete fuel types choices and of the gasoline demand. Finally, section 7 concludes.

2 The Data

The estimation of fuel types choice and gasoline demand uses data from the 1994 *Residential Transportation Energy Consumption Survey* (RTECS), published by the Environmental Information Administration (1997). This is a detailed survey of representative 3002 households on their driving behavior, vehicle ownership and choice of fuel types for each car in their vehicle stocks. The survey is conducted by the Department of Energy. The households are further interviewed on their demographic and socioeconomic characteristics, and on the characteristics of each of their vehicles. 87.6% of the households in the sample own at least a car. 0.04% of these households have at least a car in their vehicle stock which is owned by their employer. These observations were eliminated and attention was only paid to households which own the cars themselves. Further, to model fuel types choice as a matter of consumer tastes, only households whose vehicles are driven by unleaded gasoline (regular, midgrade, premium) were considered. 96.6% of all vehicles in the original sample are driven by these gasoline types.[4] A full listing of variable names, definitions, and summary statistics on the remaining households, the vehicles and gasoline consumption is given in Appendix B.

Annual household income is provided as a categorical variable. There are 25 income categories and each category represents a range in thousands of dollars. Using the maximum-likelihood method for grouped lognormal distributions (see Aitcheson and Brown, 1957), within-cell means for each of the ranges were estimated, and each household's income were set to the mean of the range in its according category. The resulting household incomes were then treated as continuous variables in the gasoline demand estimation.

In the survey, observations on driven mileage were collected for each car from odometer readings. Gallons of gasoline consumption were calculated by combining miles driven with estimated miles-per-gallon figures. Various miles-

[4] 61 vehicles of a total of 5519 vehicles are driven by leaded gasoline, 60 by diesel, and 61 by alternative fuels (ethanol, gasohol, natural gas, propane, alcohol).

Table 1: Absolute Frequency of Vehicles by Fuel Types

Fuel Types\Vehicle	First Vehicle	Second Vehicle	Third Vehicle	Forth Vehicle	Total
REGULAR	1678	1171	488	179	3516
MIDGRADE	347	212	91	24	674
PREMIUM	459	259	108	43	869
REG *and* PREM	3				3
Total	2487	1642	687	246	5062

per-gallon variables are provided in the data set. For the estimation of the model, the in-use miles-per-gallon observations were used.

Two price variables, constructed as follows from the original data set, are used in the estimation. The first price variable varies over fuel types and regions. The RTECS survey reports values of gasoline expenditure for each car of every household. Besides, in this survey, average prices for each of the three gasoline types are compiled for each region and then used to calculate the aforementioned household gasoline expenditures. Thus, taking the ratio of gasoline expenditure and gasoline consumption for each of the three gasoline types and for each region results in the first price variable which is denoted by $PGAS$ (see Appendix B for a definition). This price variable is an explanatory variable in the estimation of fuel types choice. The second price variable varies over households and is used in the estimation of gasoline demand. This price, denoted as P_{aggr}, represents an aggregate gasoline price for each household and is constructed as a weighted mean of the prices ($PGAS$) of the differentiated gasoline types that are chosen by the household. The weights are calculated for every household by taking the ratios of driven miles with a specific fuel type and total of miles driven.

3 Descripitve Statistics of Fuel Types Choice

To support the multiple-discrete part of the model, I first present some particular features of fuel type choice which underlies the demand for differentiated gasoline. Table 1 shows the distribution of vehicles for the different fuel types. The columns report the ranking of the cars in the households' vehicle stocks. By ranking, I mean whether the vehicle under consideration is the household's "first car", "second car" and so on. The figure in each cell describes the number of cars in the n'th rank in the vehicle stock that are driven by a specific fuel type (rows). For example, 1642 of all cars are households' second car (see last

Table 2: Joint Frequency of Number of Fuel Types Chosen and Number of Cars in Vehicle Stock

Fuel Types\Vehicles	1	2	3	4	Total
1	833	694	257	86	1870
2		244	143	41	428
3			21	34	55
Total	833	938	421	161	2353

row) and 212 of which are driven by midgrade gasoline.

The figures of Table 1 show that the vast majority of the vehicles in the sample are driven by exactly one fuel type. The number of vehicles which are driven by two fuel types (regular and premium) amount to only three of a total of 5062 vehicles. Households' choice of exactly one fuel type for each of their vehicles thus seems to be an essential feature in fuel type choice.

In Table 2, the joint frequency of the number of fuel types (rows) and the number of vehicles held by the household (columns) is reported by taking only households which use exactly one fuel type for each car into consideration. For instance, of the 421 households owning three cars, 143 have chosen two different fuel types for their three vehicles, i.e. regular for two cars and midgrade for one car, or regular for two cars and premium for one car etc.

The data in Table 2 makes two further essential features in fuel type choice evident. First, households with two or more cars may make multiple-discrete choices when buying fuels for their vehicles. Restricting estimation of gasoline demand at the micro level to households choosing only one fuel type would imply to disregard some of the available information and probably to produce distorted results since such a model would not be consistent with actual demand behavior. Second, the share of households making multiple-discrete choices varies considerably depending on the number of vehicles they own. Whereas 26% of housholds with two cars spread their fuel type choice over different types, this figure amounts to 39% for households with three vehicles and to 47% for households with four vehicles. Hence, in estimating gasoline demand, one has to account for this heterogeneity in household behavior, particularly by differentiating fuel types choice according to the number of vehicles owned.

4 The Model

This section outlines the main elements of the model. First, the products of the model are presented, followed by a modelling of gasoline demand. Finally, the behavioral implications are briefly discussed.

4.1 The Commodities

The commodities for the discrete choice are described as follows. There are $c = 1, ..., C$ car types and $g = 1, ...G$ gasoline types. Let c_n^h denote the n'th car of household h being of type c. Analogously, let g_n^h be the gasoline of type g which is used for the n'th vehicle of household h. The vector $s^h = \left(c_1^h, ..., c_{N^h}^h\right)$ denotes the vehicle stock of household h. The vehicle stock s^h contains two pieces of information that are important for the model. First, it describes the number N^h of vehicles the household's vehicle stock is made up of. Second, it ranks each car in the household's vehicle stock by determining at which position of the vector the car enters. This notation makes it possible, later, to differentiate among choices of vehicle stocks which differ with respect to their car rankings but otherwise consist of the same number of cars and the identical types of vehicles. This may be important to be taken into account because households may reveal different fuel types choice and gasoline demand behavior depending on what type the first, second etc. car is. Vector $f^h = \left(g_1^h, ..., g_{N^h}^h\right)$ denotes the fuel vector assigning to each of the N^h vehicles in stock s^h the type of gasoline that is used for. Note that the fuel vector f^h has exactly the same number of entries as vehicle stock s^h. Thus, it takes the feature of fuel type choice into consideration that only one gasoline type is chosen for each car (see Table 1). Additionally, and more importantly, this notation allows different gasoline types to be components of the fuel vector f^h. It is precisely in f^h where the possibility of multiple-discreteness in fuel type choice enters the model.

We observe a household h choosing an alternative $\left(s^h, f^h\right)$. We omit henceforth the household index for the sake of simplified presentation. Further, we slightly abuse the notation by letting s and f also denote an index with $s = 1, ..., S$ and $f = 1, ..., F$. Thus, households face a set of $S \cdot F$ mutually exclusive alternatives in their vehicle stock and fuel types choice.[5]

[5] To give a simple illustration of this notation for vehicles, consider an economy where households own at most two vehicles and face two types of cars, indexed 1 and 2. Then, for instance, $s = 1$ may correspond to $s = (1)$, $s = 2$ to $s = (2)$, $s = 3$ to $s = (1, 1)$ and so on until $s = S = 6$ may correspond to $s = (2, 2)$. Gasoline vectors f and alternatives (s, f) can analogously be illustrated.

To introduce the notion of product differentiation, I adopt the characteristics approach which assumes that each alternative (s, f) can entirely be described by a bundle of attributes over which households have preferences. Further, consumer behavior is modelled within the random utility framework. A random utility model arises when one assumes that a consumer's utility function, which is fully known to her, contains some components that are unobservable to the researcher, and hence are treated as random variables.

Here, I assume that there are three sets of attributes that determine consumer's utility. The first two are observed, the last one is unobserved. Besides these attributes presented below, household's preferences may further be influenced by her own characteristics (age, education etc.). To simplify the notation, I will suppress household characteristics in the presentation of the model. They are, of course, incorporated in the estimation.

The first set of attributes contains characteristics of the vehicle stock. These are denoted by $a_s = (a_{1,s}, ..., a_{K^s,s})$ for all $s = 1, ...S$, where a_s is a vector of K^s entries, and entry $a_{k,s}$ describes the amount of attribute k for vehicle stock s. I allow the number of components in the attribute vector to vary with the vehicle stock since it is natural to assume that the number of attributes rises with the number of cars in s. Thus, a vehicle stock consisting of two vehicles has potentially at least two entries for cylinders in its characteristics vector, one entry for cylinders of the first vehicle and a second entry for cylinders of the second vehicle (and possibly some interactions between them), whereas a vehicle stock consisting of only one car has only one entry for cylinder. The second set of attributes contains characteristics that vary both over the vehicle stock and its corresponding fuel vector. These attributes are denoted by $b_{sf} = \left(b_{1,sf}, ..., b_{L^{sf},sf}\right)$ for all $s = 1, ..., S$ and $f = 1, ..., F$, where vector b_{sf} consists of a variable number L^{sf} of components $b_{l,sf}$ which describe the amount of attribute l in alternative (s, f). Underlying this notation is the assumption that, when only considering the differentiated types of unleaded gasoline, households do not care about fuel characteristics *per se* but about the combination of the fuel type with the characteristics of their vehicles. Premium gasoline, for instance, does not seem to be valuable on its own right, but only in combination with the vehicle's characteristic of being a Rolls Royce. Mostly, these attributes will come in the form of interaction terms between vehicle characteristics and gasoline type dummies in the empirical implementation of the model.

The last set of attributes contains components which are deterministic for a household facing a choice among alternatives, but unobservable to the econometrician. The unobservable, denoted by the scalar ε_{sf}, may capture effects due to unobserved characterstics of the household and/or unobserved attributes of the chosen alternative, often labelled in the related literature as "quality" of the object to be selected. An estimation problem that may arise in the context of this unobserved quality is an endogeneity bias, even in the case when micro data are available (Berry, 1994).[6]

4.2 Gasoline Demand

Given the stock of vehicles and its corresponding fuel vector, each household chooses a vector of gasoline quantities $X = (X_1, ..., X_G)$ which enters a utility function of the following form:

$$U\left(X; a_s, b_{sf}, \varepsilon_{sf}\right). \tag{1}$$

Utility directly depends on gasoline consumption levels and on the bundle of attributes representing the given alternative.

Note that each gasoline consumption level refers only to one fuel type, but not to the cars in the vehicle stock. Hence, with this modelling of consumer behavior, I implicitly assume that households already optimized the allocation of consumption of a gasoline type across their vehicles driven by this gasoline type. The utility function in (1) is thus an indirect one.[7] A further comment has to be made with respect to the attributes. The model above allows consumers to reveal different gasoline demand patterns varying with the characteristics of the chosen alternative (s, f). With this assumption, I want to capture the idea that the characteristics of the vehicle stock, and the ranking of the cars and of fuel types within the vehicle stock may influence the choice of the quantity of gasoline consumption.

4.3 Behavioral Implications of the Model

Households maximize their utility for a given alternative by choosing an optimal gasoline consumption vector under the constraint of their budget set $B\left(p_f, y\right)$, where p_f is the fuel price vector corresponding to the gasoline types that appear in alternative (s, f) and y denotes household income.

Define the "indirect utility" which arises to the household from the optimal

[6] The unobserved factors that ε_{sf} might represent in our context of gasoline estimation is discussed in section 5. See also this section for a discussion of the identification assumptions in this model.

[7] The model thus implicitly makes a separability assumption.

gasoline demand given any alternative (s,f) as

$$v\left(p_f, y, a_s, b_{sf}, \varepsilon_{sf}\right) = \max_{X \in B(p_f, y)} U\left(X; a_s, b_{sf}, \varepsilon_{sf}\right). \tag{2}$$

Let (s^*, f^*) represent the alternative that gives a household the highest utility. Then the optimal discrete choice (s^*, f^*) can be summarized as follows:

$$v\left(p_{f^*}, y, a_{s^*}, b_{s^*f^*}, \varepsilon_{s^*f^*}\right) = \max_{(s,f)} \left\{v\left(p_f, y, a_s, b_{sf}, \varepsilon_{sf}\right)\right\}. \tag{3}$$

The model has behavioral implications which are consistent with the essential features of gasoline demand. First, through a convenient categorization of the choice set for alternatives (s,f) the model allows for multiple-discreteness in fuel types choices. Second, depending on the properties of the indirect utility function in (3), the model may be able to deal with the multiple-discrete/continuous dimension of gasoline demand behavior. In particular, if the indirect utility function is quasiconvex in p_f, decreasing in all the components of p_f and increasing in y, it satisfies Roy's identity:[8]

$$X_g^* = -\frac{\partial v\left(p_{f^*}, y, a_{s^*}, b_{s^*f^*}, \varepsilon_{s^*f^*}\right)/\partial p_g}{\partial v\left(p_{f^*}, y, a_{s^*}, b_{s^*f^*}, \varepsilon_{s^*f^*}\right)/\partial y}, \text{ for all } g \text{ appearing in } f^*, \tag{4}$$

where X_g^* is the optimal amount of consumption of type g gasoline and p_g is its price (and is one entry in the fuel price vector p_f). Since the indirect utility only depends on the prices of gasoline types which actually are chosen, the application of Roy's identity is accordingly restricted to the same gasoline types. More concretely, the model generates a gasoline consumption pattern of the form $X^* = \left(0, ..., 0, X_g^*, 0, ..., 0, X_{g'}^*, ..., 0\right)$ for a household that has chosen types g and g'.

The third property of the model is that in the estimation of price and income elasticities, the effects of the characteristics of the vehicle stock on gasoline demand are controlled for. Because households' preferences and the quality of the vehicle stock are incompletely observed, the quantities X_g^* are random variables (as can be seen in (4), X_g^* depends on ε_{sf}). In order to develop the probability formulas for the continuous choices, it is convenient to introduce the sets $A_{sf} = \{\varepsilon | v\left(\cdot, \varepsilon_{sf}\right) \geq v\left(\cdot, \varepsilon_{s'f'}\right), \forall s'f'\}$, for $s = 1, ..., S$ and $f = 1, ..., F$. The conditional mean quantity can then be obtained by the regression $E\left[X_g^* | \varepsilon \in A_{sf}\right]$ which yields predicted gasoline demand conditioned on the selected vehicle stock characteristics.

[8] For the application of Roy's Identity, some further technical assumptions must be satisfied besides the null-homogeneity in prices and income. See Mas-Colell, Whinston, Green (1995).

5 Econometric Specification and Estimation

5.1 Specification of Multiple-Discrete/Continuous Choice

To estimate the model, I have to choose functional forms of the attributes representing the alternatives and of the indirect utility function (2). Further, I have to specify the distribution of ε_{sf} in order to generate the discrete choice probabilities and the predicted gasoline demand functions conditioned on the selected vehicle stock.

As to the discrete choice characteristics, I assume that a vehicle stock with its associated gasoline vector can be summarized by a scalar which may be interpreted as an index of the overall quality of (s, f). To fix ideas I employ the following specification for this quality index which is denoted by ϕ

$$\phi\left(a_s, b_{sf}, \varepsilon_{sf}\right) = \exp\left(\alpha' a_s + \beta' b_{sf} + \varepsilon_{sf}\right). \tag{5}$$

The quality index contains a deterministic part consisting of the observed attributes and a random part which is represented by the unobserved ε_{sf}. In the deterministic part the vehicle stock and gasoline type attributes enter linearly and additively seperably into ϕ. The vectors α and β are parameters to be estimated.

An essential implication of the random utility model is that discrete choices depend on the characteristics of the different alternatives and on their prices, but not on the parametric form of the indirect utility function. The following specification of the indirect utility function satisfies this requirement of the random utility framework and has, thus, been used for estimation:

$$v\left(p_f, y, \phi\left(a_s, b_{sf}, \varepsilon_{sf}\right)\right) = \frac{\mu}{\rho - 1}\left(\frac{\phi\left(a_s, b_{sf}, \varepsilon_{sf}\right)}{\prod_{n=1}^{N} p_{g_n}^{\gamma_n}}\right)^{\rho - 1} + \frac{1}{1-\eta} y^{1-\eta}, \tag{6}$$

where p_{g_n} is the price of type g gasoline used for the household's n'th vehicle, y denotes household income and μ, ρ, η and γ_1 to γ_N are parameters to be estimated. This specific form of indirect utility is a similar one to a suggestion of Hanemann (1984). However, I modified it to allow for multiple-discrete choices.

From (6), the probability of choosing alternative (s, f) can be written as

$$\Pr\{\alpha' a_s + \beta' b_{sf} - \sum_n \gamma_n \ln p_{g_n} + \varepsilon_{sf} \geq \alpha' a_{s'} + \beta' b_{s'f'} - \sum_n \gamma_n \ln p_{g'_n} + \varepsilon_{s'f'}, \text{ for all } (s', f')\}.$$

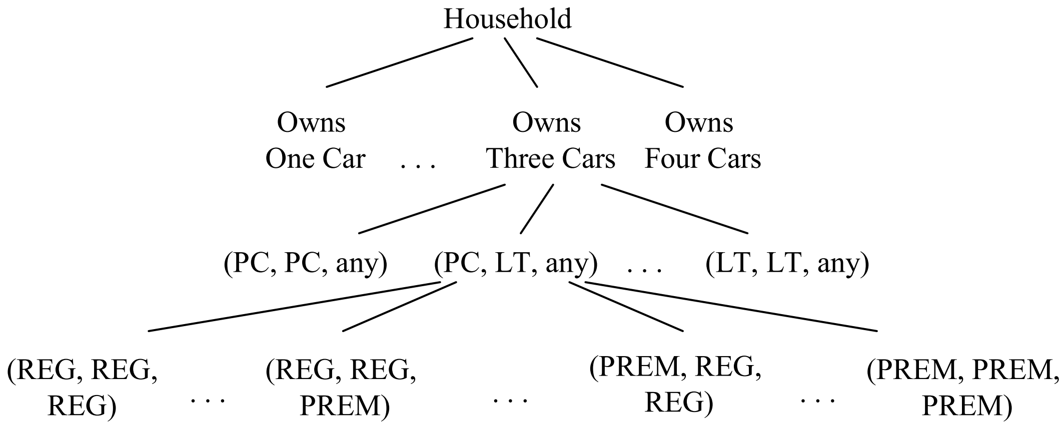

Figure 1: Vehicle Stock and Corresponding Fuel Types Choice

Thus, the discrete choice probability does not depend on the parameters of the indirect utility function μ, ρ, η and γ_1 to γ_N.

I assume that the error term ε_{sf} is distributed according to the generalized extreme value distribution and that the discrete choice for the vehicle stock and the corresponding gasoline types has the nested structure as depicted in Figure 1. With this assumption I relax the IIA assumption which is common to multinomial logit models (for a more detailed discussion see later in this section). The assumptions with respect to the correlations of unobserved factors across alternatives, as they are implied by this nested logit model of Figure 1 deserve some comment. Though most of the time vehicles are bought because consumers care about driving them, some types of consumers may have immediate preferences for just owning cars. Since such preferences are not observable, I assume that households who possess the same number of cars share this general "warm glow" for vehicles and I nest households with same vehicle numbers together.

Many specific factors that relate to the choice of vehicle type, such as the household's driving style, her preference for safety in traffic, the topographical conditions of her residence, the driving distances to working place etc. are not covered by the data either and thus remain unobserved. Consequently, I nest together alternatives which consist of the same vehicle categories. Gasoline types are not nested together because relevant unobserved factors which may be similar for certain gasoline types do not seem plausible when only considering differentiated unleaded gasoline. The main motivation for this assumption is that differentiated unleaded gasolines do not systematically differ in their enviromental performance. This is the only possible unobserved factor with respect to gasoline when vehicle characteristics can be controlled for.

As indicated in Figure 1, I have used a categorization of the vehicle stock that only takes the types and ranking of the first two cars into account. The categories (*passenger car, light truck, passenger car*) and (*passenger car, light truck, light truck*), for instance, are treated as identical alternatives for vehicle stock choice. This aggregation over categories allows to keep the number of observations at each decision node at the last stage in the nested logit model sufficiently high.[9] In contrast to vehicle stock choice, no aggregation was made, neither over the fuel types nor over the rankings in the entering position of gasoline types into f. Thus, the fuel vectors (*regular, regular, premium*) and (*premium, regular, regular*), for instance, are treated as different choice categories even though in both cases regular gasoline is chosen twice and premium gasoline is chosen once.

The probability of choosing alternative (s, f) can be written as $P_{sf} = P_s P_{f|s}$, i.e as the product of the marginal probability P_s of choosing vehicle stock s and the conditional probability $P_{f|s}$ of choosing fuel vector f, given vehicle stock s.[10] As shown in McFadden (1978, 1981), the conditional and marginal probabilities have the following form under the nested logit assumption:

$$\begin{aligned} P_{f|s} &= \frac{\exp\left(\left(\beta' b_{sf} - \gamma' \ln p_f\right)/\lambda\right)}{\sum_{f'=1}^{F} \exp\left(\left(\beta' b_{sf'} - \gamma' \ln p_{f'}\right)/\lambda\right)} \\ P_s &= \frac{\exp\left(\alpha' a_s + \lambda I_s\right)}{\sum_{s'=1}^{S} \exp\left(\alpha' a_{s'} + \lambda I_{s'}\right)} \end{aligned}$$

where $I_s \equiv \ln\left(\sum_{f'} \exp\left(\left(\beta' b_{sf'} - \gamma' \ln p_{f'}\right)/\lambda\right)\right)$ stands for the inclusive value term. The coefficient of the inclusive value, λ, measures the degree of substitutability across alternatives. This coefficient has to lie within the unit interval, for otherwise the nested logit model would not be consistent with the random

[9] In the estimation of fuel types choices, I used type dummies for the third and fourth cars to control for this aggregation.

[10] Since in our notation the index s includes both the number of cars chosen as well as the car types, in order to fit into the nested structure of Figure 1, P_s must itself be written as the product of the marginal probability of owning a number of N cars and the conditional probability of choosing a vector of vehicle types given the number of cars owned. For the sake of a simplified presentation, I retain this notation and abstract form the explicit modelling of vehicle number choice.

utility approach, since with $\lambda > 1$, substitution across nests exceeds substitution within nests.

There are at least two advantages arising from imposing the nested structure on fuel types selection processes. First, in the proposed modelling a multitude of combinations of vehicle stock holdings and fuel vector choices is generated which makes the number of alternatives included in the choice set very large. The nesting according to Figure 1 helps to considerably reduce the size of the choice set at each node of the tree.

Moreover, choosing the nested logit model as the estimation strategy relaxes the "independence of irrelevant alternatives (IIA)" assumption that is common to multinomial logit models and which implies implausible substitution patterns (for a detailed discussion, see Berry, 1994; Berry et al., 1995; Goldberg, 1995). The introduction of more flexible correlation patterns in the unobserved factors through the assumption of a generalized extreme value distribution of ε_{sf} should thus lead to less restrictive substitution patterns over alternatives in the whole model. Further, I include interaction terms between household/vehicle specific attributes and gasoline specific type dummies in the specification of the model so that even within a nest more plausible substitution patterns should result from estimation.[11]

A few remarks should be made with respect to endogeneity. Throughout the paper, I assume that the explanatory variables of the nested logit model are identified. However, this assumption may seem problematic when taking into account the fact that supply of gasoline is provided by oligopolistic firms who may follow a price-setting behavior in a differentiated-goods industry. As long as unobserved product characteristics are present in the error term ε_{sf}, a simultaneity problem in form of endogenous prices may arise as has been discussed in numerous applications of NEIO to differentiated goods industries (see e.g. Berry, Levinsohn and Pakes, 1995; Nevo, 2000).

The main justification of the assumption that gasoline prices are not correlated with the error term is the use of micro-level data, and, in particular, the availability of data concerning the vehicle attributes. It is natural to assume that fuel type specific attributes are only valuable in combination with vehicle specific attributes. By letting these vehicle attributes being interacted with gasoline type specific dummies the specification of the model should allow for a sufficient control of the unobserved characteristics.

[11] With the inclusion of interaction terms, the model described here becomes similar to the random coefficients model which is a generalization of the nested logit model (see Cardell, 1997; Berry and Pakes, 2001).

5.2 Gasoline Demand Equation

The suggested specification of the indirect utility function produces tractable demand models with constant price and income elasticities. As shown in Appendix A predicted gasoline demand, conditional on the fact that alternative (s, f) is chosen is given by

$$E\left[\ln X_g^* | \varepsilon \in A_{sf}\right] = \theta_1 + \theta_2 \ln p_g + \sum_{g'} \theta_{g'} \ln p_{g'} + \eta \ln y + (\rho - 1) \ln D_{sf} \quad (7)$$

where g and g' occur in (s, f) and

$$D_{sf} = \sum_s \exp\left(\alpha' a_s\right) \left[\sum_f \exp\left(\left(\beta' b_{sf} - \gamma' \ln p_f\right)/\lambda\right)\right]^\lambda.$$

That is, the model predicts gasoline demand for every household, conditioned on the vehicle stock choice, as a function of gasoline prices, income, the observed characteristics of the alternatives, and of two sets of parameters: the parameters to be estimated in the demand regression (the θ's, η and ρ) and the parameters being estimated in a previous step in the nested logit model (α, β,γ, and λ). D_{sf} can be interpreted as a discrete choice term and is an observed explanatory varying over households and its vehicle stock attributes. Thus, introducing $\ln D_{sf}$ into the model allows for controlling for the multiple-discreteness part in the multiple-discrete/continuous choice feature in gasoline demand.

Defining the prediction error as

$$\nu = X_g^* - E\left[\ln X_g^* | \varepsilon \in A_{sf}\right], \quad (8)$$

where X_g^* is the actual demand for gasoline type g, one gets a regression model with the moment condition $E\left[\nu | a_s, b_{sf}, p_f, y, \theta, \eta, \rho\right] = 0$ at the true parameter values.

5.3 Estimation

In the fuel types choice model, the coefficients α, β, λ, and γ_1 to γ_N were estimated by sequential maximum-likelihood which decomposes estimation into three stages according to Figure 1. This estimation procedure was chosen because the size of the choice set and the large number of parameters involved in the specification of the model made the use of the full information maximum-likelihood method infeasible. To get consistent estimates of the covariance matrix, standard errors were corrected by applying the recursive formulas derived in McFadden (1981).

In the analysis of the continuous choices, the estimated coefficients of the nested logit model were inserted into the regression equation that was then fitted by OLS. This procedure gives estimates of the parameters θ_1, θ_2, the θ_g's, η and ρ which are consistent, however the covariance matrix of these estimates is incorrect. For this reason I modified the estimation of the covariance matrix by taking into account that in the moment conditions some of the variables are estimated parameters from the previous stage. The application of this sequential estimator yields consistent standard errors.

6 Results and Interpretation

This section summarizes the results from the estimation of gasoline demand. First, the results of the discrete choice part are presented, followed by a presentation of the results of the continuous choice part.

6.1 Results of the Vehicle Stock / Fuel Vector Choice Model

According to Figure 1, I sequentially estimated the choice of the vehicle stock and its corresponding fuel vector. Four groups of explanatory variables have been used to describe households' discrete choice behavior: household characteristics, vehicle attributes for each vehicle, mobility variables and fuel specific attributes, namely gasoline prices and fuel type dummies. The estimation results for the nested logit model are provided in Appendix C and its interpretation is presented seperately to each stage of estimation below.

6.1.1 Choice of Fuel Vector

The explanatory variables included at this stage of estimation are vehicle attributes and gasoline prices. To limit the number of parameters to be estimated, only income is used as explanatory household characteritics. Vehicle attributes and income enter into the estimation in interaction with dummy variables for the gasoline vector choices, otherwise they would fall out of the probability in the conditional logit model. At all fourteen nodes in the last stage of the nested logit model, virtually the same set of variables are used. Two potential problems arise with the estimation of the gasoline vector. First, as also reported in Goldberg (1995), some vehicle attributes are highly correlated. Engine size and cylinders, for instance, have a correlation coefficient of 0.91. The value for the correlation of fuel efficiency and engine size or cylinders ranges between

-0.78 and -0.71. To avoid the adverse consequences of multicollinearity, either cylinders or engine size were involved into the estimation. The second problem is associated with the large number of fuel vector choices when the household's stock of cars is large. A household owning three cars, for instance, and having the choice of three gasoline types for each car faces a potential choice set of 27 alternatives. Further, since most of the explanatory variables are vehicle and household attributes, they consequently do not vary across the chosen alternatives and thus come along in interaction with the choice types in the estimation model. This feature of the model, however, makes the set of potential explanatory variables immensely big.[12] To keep the number of explanatory variables at relatively modest levels, the five to seven most frequently chosen alternatives were used as dummy variables in interaction with vehicle and household attributes. Since the number of explanatory variables still was remarkably high, further reduction was achieved by selecting out groups of explanatory variables using Akaike's Information Criterion AIC (see Amemiya, 1985). These models with reduced numbers of covariates were then reestimated including type specific dummies for the omitted alternatives and/or type specific dummies for the used alternatives. It turned out that the signs and precision of the parameters of the reduced models did not change. Further, the type specific dummies were partly highly imprecise and were therefore omitted in the model.

The results are reported in Tables C1 to C6 in Appendix C.[13] With a few exceptions, gasoline prices are highly imprecise and sometimes have a positive sign. I also estimated models omitting vehicle characteristics and using instead only income and alternative household attributes. In these models, price effects were significantly negative, supporting the idea that controlling for the car park's attributes, prices do not matter in the choice of gasoline type. The coefficients for income are more precisely estimated and have most of the time the expected sign for high quality gasoline. However, precision decreases when the vehicle stock consists of light trucks. The fact that the income coefficient is as high for regular gasoline as for premium gasoline in few cases seems to be somewhat puzzling. The vehicle attributes parameter estimates conform to expectations: new cars, luxury cars and sport utility vehicles are more likely to be driven by premium gasoline and less likely to be driven by regular gasoline, and vice versa for minivans and fuel efficient cars. No clear picture emerges as to cylinder and

[12] With four vehicle attributes and income as explanatory variables, for instance, the potential number of explanatory variables is 130.

[13] The results for the remaining eight models corresponding to the nodes for households with three and four cars, respectively are not reported for the sake of saving space. These models have comparable results and are available from the author upon request.

engine size.

6.1.2 Choice of Vehicle Category

The parameter estimates of the following stage are used to compute the inclusive value term. For the sake of a simpler interpretation of the results preference was given to the current specification with only one value for the inclusive value term. Besides the inclusive value, the variables that enter at this stage are household specific attributes (age, income, familiy size, marital status, residence), mobility attributes (whether the vehicle is used for commuting) and vehicle attributes, taking again into account the possibility of multicollinearity between engine size and cylinders. One crucial determinant in vehicle category choice, however, is not considered in our model because of lack of data. The holding of a stock of cars is associated with vehicle specific costs other than fuel costs. Such costs (e. g. vehicle insurance premiums or costs for vehicle inspection) are expected to affect the choice of vehicle category since these costs differ between and across passenger cars and light trucks.

As shown in Tables C7 to C8c, the inclusive values lie within the unit interval.[14] Further, they are quite low for the nodes regarding two to four cars.[15] That is, the suggested nested structure of Figure 1 seems to be consistent with the random utility approach. The estimated coefficient on the inclusive value term in Table C7, at the node for households having only one car, was negative (-0.98) when passenger car was the normalized choice category and thus not consistent with the the random utility approach. Therefore I selected the alternative choice, namely light trucks, as the normalized choice category which yielded a positive sign. In general, the estimated coefficients have the expected signs.[16] Households living in cities or suburbs are less likely to choose light trucks. The probability to have a light truck increases with familiy size for households having more than two cars and with households with married couples. The income coefficient is positive and highly significant for households with one or two cars, but becomes less precise for households owning more cars.

[14] The standard errors of the estimates in this and the first nest of the nested logit model and in the regression models which are presented later are not yet corrected according to McFadden's formula, and according to the sequential estimator, respectively.

[15] The coefficients for the inclusive value for the nodes regarding three cars are equal 0.004 and regarding four cars equal 0.176. The results on the vehicle types choices for three and four cars are not reported to save space. They are, however, briefly mentioned in the text. The corresponding tables are available from the author upon request.

[16] Note that the coefficients in Tables C8a to C8c are simultanously estimated in the same multinomial logit model. They are presented in seperated tables for the sake of presentation. The same holds for Tables C9a to C9c in the next subsection.

As to vehicle characteristics, only the coefficients for engine size/cylinders are regularly insignificant and do not show a clear picture. The same holds for commuting.

6.1.3 Choice of Number of Vehicles

All variables at this stage of estimation besides the inclusive value are household attributes. The estimation results are reported in Tables C9a to C9c. The signs and levels of the coefficients mostly correspond to intuition. The dominant determinants of the choice for vehicle number are income and family size. The estimated coefficients of both of these variables are quite high and highly significant in all three choice alternatives of the model. Employment status of the second household head and lifecycle variables (marital status, age of children in the household) also play a crucial role. In the former group of explanatory variables, the coefficients are particularly precise in the choice between one and two cars and between two and four cars. Somewhat surprisingly, the probability of having one car declines faster for part time employed second household heads than for their full time employed counterparts. Further, households having children below sixteen tend to have fewer cars. This result is most striking for the choice of one car where the signs are positive though imprecise. Having children between sixteen and seventeen also tends to lower the probability of having more cars. The exception is the choice group of having three cars where the sign is positive and significant. The reason for this may be that teenagers at this age get their own cars but still stay in their parents' home. Residence and region variables mostly have the expected signs but are not significant throughout the entire choice sets.

Generally, the parameter coefficients have the expected signs and the estimates of the inclusive value coefficients are consistent with the random utility paradigm. The latter indicate that the above nesting conforms with the data. Further specification tests, however, are necessary, particularly by comparing the previous model with alternative nesting structures. A plausible alternative would be the addition of a preceding node where the discrete choice is between owning vehicles versus not owning vehicles.

6.2 Results from the Gasoline Demand Estimation

Estimated coefficients from the previous nested logit model are used to generate $\ln D_{sf}$ in the gasoline demand equation(7). This explanatory variable contains the information on the bundle of attributes of the vehicle stock and its correponding fuel vector chosen by the households to satisfy their mobility needs.

My chief concern in this paper is to provide results that directly can be compared with those of traditional gasoline demand studies using aggregate time-series data. One important characteristic of those studies is that overall demand for gasoline is estimated without differentiating across various types of this product. Thus, for the sake of comparison, I estimate a modified version of (7), where for every household the demand for each gasoline type selected is summed up to get her aggregate gasoline consumption. Similarly, to get a corresponding counterpart for prices, the aggregate gasoline price P_{aggr} is used as an explanatory variable (see section 2). Remember that the aggregate gasoline price varies over households.

Reintroducing the household index, the modified aggregate gasoline demand equation can then be written as

$$X^h_{aggr} = \tilde{\theta}_1 + \tilde{\theta}_2 \ln P^h_{aggr} + \eta \ln y^h + (\rho - 1) \ln D^h_{sf} + \tilde{\nu}^h,$$

where $X^h_{aggr} = \sum_g X^{h^*}_g$, $\tilde{\nu}^h$ is the new error term, and the greek letters are the parameters to be estimated.

The main advantage in this comparison exercise is that the effect of the households' previous choice of vehicle stock and fuel types on price and income elasticities can be controlled for through D_{sf}. Further, because of the use of micro-level data, household and vehicle characteristics can additionally be involved in the demand estimation.

Seven different versions of the above model are estimated whose results are reported in Table 3. These versions differ according to what further groups of explanatory variables are involved in the gasoline demand equation besides those indicated in the original model. The additional variables are household attributes and/or averages of vehicle characteristics in a household's vehicle stock. To be able to select among the different versions of the model, Akaike's Information Criterion (AIC) was calculated for each regression equation. A lower AIC may be interpreted as an indicator that the respective model fits the data better.

Before commenting the results of the estimates, I should make a note on the price and income elasticities that the model implies. The regression equation has the form of a constant elasticity demand. However, gasoline prices and income are also included in the discrete choice term $\ln D_{sf}$. It can be shown that the derivative of $\ln D_{sf}$ with respect to gasoline prices depends on the estimated fuel price coefficients of the previous nested logit stage. Since these coefficients were highly insignificant, $\partial \ln D_{sf} / \partial \ln p_g$ can be set equal to zero and the coefficient

$\tilde{\theta}_2$ in the above regression equation thus adequately represents the gasoline price elasticity.

Income was treated as a categorical variable in the vehicle and fuel type choice estimation, and as a continuous variable in the indirect utility function (6) to get tractable demand models. Thus, $\ln D_{sf}$ is by construction not differentiable in income. Still, in the interpretation of the results in this subsection I will not put much weight on the income coefficient because the estimated parameters for the categorical income variable in $\ln D_{sf}$ are significant and may still have an influence on the income elasticity.

Column (1) in Table 3 presents the estimation results of the simplest gasoline demand model where only price and income are considered. The coefficients have the expected signs and are highly significant. Column (2) additionally includes the disrete choice term $\ln D_{sf}$. By comparing the estimates of models (1) and (2), one can determine the effect of holding a specific vehicle stock on gasoline consumption. Taking $\ln D_{sf}$ into account leads to a decrease of the price elasticity (in absolute terms) of about 20%. This result conforms to the intuition that, controlling for her vehicle stock, a household is less flexible in adjusting its driving behavior to price changes. This finding remains robust when household and/or vehicle characteristics are involved in the estimation. In these cases the drop in price elasticity even amounts to about 25%. Further, precision of the price coefficient decreases in all models when the explanatory variables consist of $\ln D_{sf}$ whereas the coefficient for $\ln D_{sf}$ itself is mostly significant.

The remaining models in Table 3 all include $\ln D_{sf}$. Model (3) additionally includes household attributes, model (4) is the same as (3) but additionally has residential variables as regressors. Generally, in model (3) the coefficients have the expected signs. Price elasticity is substantially lower than in models (1) and (2) and it is highly insignificant. This result is mainly due to the inclusion of the vehicle number (without this variable price elasticity equals -0.378 and the standard error is 0.183). When residence variables (city, suburb, rural) are taken into account, price elasticity becomes even less precise and changes sign (see model (4)). Thus, by controlling for the number of vehicles owned and residence, price effects do not seem to matter at all in gasoline demand.

In model (5), model (2) is extended by average attributes of the vehicle stock. Price elasticity has a negative sign and is less insignificant than in the models with household attributes. More interestingly, model (5) gives evidence for the existence of the rebound effect which states that fuel efficient cars reduce environmental pollution, but are driven more often, thus, countervailing the former positive effects. In this model, a one percent increase in fuel efficiency

Table 3: OLS Estimates, Dependent Variable is LNTOTGALS

Variable	Estimate (1)	Estimate (2)	Estimate (3)	Estimate (4)	Estimate (5)	Estimate (6)	Estimate (7)
$\ln P_{aggr}$	-0.463	-0.36	-0.085	0.044	-0.174	0.064	0.161
	(0.212)	(0.195)	(0.163)	(0.162)	(0.182)	(0.144)	(0.144)
$\ln y$	0.307	0.202	0.055	0.056	0.175	0.01	0.012
	(0.02)	(0.019)	(0.017	(0.017)	(0.018)	(0.016)	(0.016)
$\ln D_{sf}$		-0.008	-0.001	-0.001	-0.007	-0.0002	-0.0002
		(0.0004)	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Household Attr.			X	X		X	X
Residence				X			X
LNAVGMPG					1.063	0.85	0.848
					(0.096)	(0.076)	(0.075)
LNAVGENG					0.964	0.835	0.822
					(0.056)	(0.044)	(0.044)
LNVAGE					0.078	-0.003	-0.005
					(0.026)	(0.021)	(0.021)
AVGCHNG					-0.097	-0.498	-0.488
					(0.037)	(0.032)	(0.032)
AVGCOMMUT					0.263	0.125	0.128
					(0.033)	(0.03)	(0.03)
CONS	6.819	6.803	6.65	5.847	-5.151	-2.654	-3.2
	(1.503)	(1.38)	(1.183)	(1.178)	(1.554)	(1.26)	(1.255)
Number of Obs.	2342	2287	2287	2287	2287	2287	2287
AIC	0.57	0.47	0.332	0.328	0.41	0.266	0.265

Note: Household attributes include the following variables: LNAGE, LNFAMSIZE,LNVEHNR, BLACK, HISP, UNEMPLY, COLLEGE, MALE, and dummies for life cyle variables. Residence includes CITY, SUBURB, RURAL.

rises driven miles by about 1%.

The last two models are combinations of the previously discussed specifications. These models generally reproduce the aforementioned results. However, the coefficients for income and for the discrete choice term become insignificant.

In general, those specifications which additionally include household and/or vehicle attributes have the lowest AIC. Thus, models (3) to (7) seem to describe the data better than specifications (1) and (2). The most relevant models in the sense of AIC are the last two specifications in which household as well as vehicle characteristics are included and price effects are virtually non-existent.

7 Conclusion

In this paper a discrete/continuous model that integrates both multiple fuel types choices and gasoline consumption is developed, taking into account the characterstics of the household and the vehicle stock hold. The proposed model permits the estimation of gasoline demand at the household-level. The main goal is to study the effect fo this kind of modelling on the results concerning the price elasticity of gasoline demand. Compared to earlier estimates from time-series studies, we find that gasoline demand is less elastic. By including additional household and/or vehicle stock characteristics, price elasticity in absolute terms decreases further and becomes very insignificant. It even has a positive sign when household residential variables and number of vehicles owned are accounted for.

Further, fuel prices do not matter in households' choices for fuel types either. This finding suggests that consumers are guided in their gasoline type selection by the characteristics of their vehicle stocks. Vehicle age and the category to which the vehicle belongs play a significant role whether a high or low quality gasoline is chosen. Vehicle characterisitcs also matter in gasoline demand. Of particular interest is fuel efficiency which has an elasticity of roughly 0.8 to 1.0. This finding may explain why gasoline consumption grew steadily in the last few decades despite a continuous improvement in average fuel efficiency.

As it stands, the suggested model in the paper has to be further examined for robustness. First, I did not explicitely model an outside alternative. This would require to specify an alternative nesting structure in gasoline consumption with an additional nest in which households may have to choose between owning and not owning a vehicle stock. Second, the data employed in this paper lack of vehicle stock holding costs such as vehicle insurance premiums, vehicle inspection fees etc. Such costs are expected to affect the choice of vehicle categories. However, there is no reason to expect that his would make demand more elastic with respect to price.

An interesting extension in the estimation may be the analysis of externalities between the different cars in a household's vehicle stock. It is reasonable to expect that in consumers' preferences attributes of the individual cars are somehow related. Further, such externalities may vary over households' characteristics and affect gasoline demand. Such spillovers within the vehicle stock can be captured by introducing interaction terms between the attributes of the individual vehicles.

Finally, as already suggested by Hendel (1999), there are many commodities and services that have a multiple-discrete feature. For example, the proposed framework of this model can be applied to goods such as cereals, electric household appliances, holidays during a given year, or the composition of the airline fleet. Beside demand analysis, one may suggest alternative fields of application in economics, e.g. in labor market economics, the analysis of labor supply behavior of part-time workers who may choose several employers or in the context of international trade theory the choice of locations by multinational firms.

8 Appendices

A Derivation of the Conditional Gasoline Demand Function

Denote by i those cars in the vehicle stock which are driven by gasoline type g, and by j those cars which are driven by any other gasoline type $g' \neq g$. Application of Roy's Identity to the indirect utility function (6) yields the predicted gasoline demand which depends on the random tastes

$$X_g^* = \kappa_1 + \kappa_2 \ln p_g + \sum_{g'} \kappa_{g'} \ln p_{g'} + \eta \ln y + (\rho - 1) \ln \phi\left(\cdot, \varepsilon_{sf}\right), \qquad (9)$$

where $\kappa_1 = \ln \mu + \ln \sum_i \gamma_i$, $\kappa_2 = (1-\rho) \sum_i \gamma_i - 1$, $\kappa_{g'} = (1-\rho) \sum_{j,g'} \gamma_j$ for each g'. Note that the coefficients $\tilde{\theta}$ vary over the number of vehicles the household owns. According to Hanemann (1984, p. 550 and 556) and Johnson et al. (1995, p. 12)

$$E\left[\varepsilon_{sf} | \varepsilon \in A_{sf}\right] = \ln \psi_{sf} + 0.57722 \qquad (10)$$

where $\psi_{sf} = \exp\left(\zeta_{sf}\right) Q\left(e^{\zeta_{11}}, ..., e^{\zeta_{SF}}\right)$ and $\zeta_{sf} = \alpha' a_s + \beta' b_{sf} - \sum_n \gamma_n \ln p_{g_n}$. Q is a nonnegative linear homogenous function. Deviating from the presentation of the econometric model of section 5 by explicitely considering the additional choice of vehicle numbers at the first stage, the specification of $Q(\cdot)$ for a three-level nested logit model, as shown in Amemiya (1985), is given by

$$Q(\cdot) = \sum_n \exp\left(\delta' Z_n\right) \left\{ \sum_s \exp\left(\alpha' a_{ns}/\lambda_1\right) \left[\sum_f \exp\left(\left(\beta' b_{nsf} - \sum_n \gamma_n \ln p_{g_n} \right) / \lambda_2 \right) \right]^{\lambda_2/\lambda_1} \right\}^{\lambda_1}, \tag{11}$$

where Z_n is the vector of explanatory variables in the first stage of the nested logit model, δ the corresponding paramter vector to be estimated, and λ_1 and λ_2 are the inclusive value coefficients of stage one and stage two, respectively. Inserting (11) and (10) into (A) yields the predicted gasoline demand (7), conditional on the vehicle stock chosen, where $\theta_1 = \kappa_1 + (\rho - 1)\, 0.57722, \theta_2 = -1$ and $\theta_{g'} = 0$.

B Variable Definitions and Summary Stastics

B.1 Variable Definitions

B.1.1 Household Attributes

AGE	Age of Household Head
INC	Categorical Income Variable, 25 Categories
y	Continuous Income Variable
VEHNR	Number of Vehicles Owned
FAMSIZE	Familiy Size
MARRIED	1 if Married
Life Cycle Variables:	
KIDLESS7	1 if Oldest Child Less Than 7 Years
KIDS7TO15	1 if Oldest Child 7–15 Years
KIDS16TO17	1 if Oldest Child 16–17 Years
NOCHILDLESS35	1 if Two Adults, Age of Head Less Than 35 Years
NOCHILD35TO59	1 if Two Adults, Age of Head 35–59 Years
NOCHILD60OVER	1 if Two Adults, Age of Head 60 Years and Over
SINGLELESS35	1 if One Adult, Age of Head Less Than 35 Years
SINGLE35TO59	1 if One Adult, Age of Head 35–59 Years
SINGLEOVER60	1 if One Adult, Age of Head 60 Years and Over
UNEMPLY	1 if Unemployed
PARTTIMEHH1	1 if Household Head Works Part-Time
FULLTIMEHH1	1 if Household Head Works Full-Time
PARTTIMEHH2	1 if Second Household Head Works Part-Time
FULLTIMEHH2	1 if Second Household Head Works Full-Time
COLLEGE	1 if Household Head Went to College
MALE	1 if Household Head is Male
HISP	1 if Household Head is Hispanic
BLACK	1 if Household Head is Black

CITY	1 if Household lives in a City
SUBURB	1 if Household lives in a Suburb
RURAL	1 if Household lives in the Country
NE	1 if North East
MW	1 if Mid West
WE	1 if West

B.1.2 Vehicle Attributes

VAGE_n	Vehicle Age of n'th car, n=1,...,4
CYL_n	Number of Cylinders of n'th Car
ENG_n	Engine Size of n'th car
MPG_n	Fuel Efficiency, Miles-Per-Gallon of n'th car
AUTOMAT_n	1 if n'th Vehicle has Automatic Transmission
COMMUT_n	1 if n'th Vehicle is Used for Commuting
Vehicle Categories:	
LUX_n	1 if n'th Vehicle Belongs to Luxury Category
MVAN_n	1 if n'th Vehicle is a Mini Van
LVAN_N	1 if n'th Vehicle is a Large Van
PICKUP_n	1 if n'th Vehicle is a Pickup
SPORT_n	1 if n'th Vehicle is a Sport Utility
AVGVAGE	Average Age of Household's Vehicle Stock
AVGENG	Average Engine Size of Household's Vehicle Stock
AVGMPG	Average Fuel Efficiency of Household's Vehicle Stock
AVGCHNG	Percentage of Vehicles in Household's Vehicle Stock Changed During 1994
AVGCOMMUT	Percentage of Vehicles in Household's Vehicle Stock Used for Commuting

B.1.3 Fuel Attributes

TOTGALS	Total Gallons of Gasoline Consumed
PGAS	Differentiated Gasoline Price; in Tenth of Cents
if North East, PGAS =	1120 (if Reg); 1247 (if Mid); 1337 (if Prem)
if Mid West, PGAS =	1075 (if Reg); 1156 (if Mid); 1243 (if Prem)
if West, PGAS =	1195 (if Reg); 1326 (if Mid); 1398 (if Prem)
if South, PGAS =	1062 (if Reg); 1178 (if Mid); 1254 (if Prem)
P_{aggr}	Aggregated Gasoline Price, Calculated for Each Household as a Weighted Mean of Gasoline Types Consumed; in Tenth of Cents
Fuel Types Dummies:	
(i) When Houshold Owns One Car	
MID	1 if Gasoline is Midgrade
PREM	1 if Gasoline is Premium
(ii) When Household Owns Two Cars	
RR	1 if Fuel Vector is (Regular, Regular)
RM	1 if Fuel Vector is (Regular, Midgrade)
RP	1 if Fuel Vector is (Regular, Premium)
...	...
MM	1 if Fuel Vector is (Midgrade, Midgrade)
...	...
PP	1 if Fuel Vector is (Premium, Premium)

B.2 Summary Statistics

B.2.1 Household Attributes

Variable	Mean	Std. Dev.
AGE	47.604	17.341
y	31724	21863
VEHNR	1.956	0.894
FAMSIZE	2.647	1.368
MARRIED	0.613	0.487
UNEMPLY	0.048	0.214
PARTTIMEHH1	0.089	0.285
FULLTIMEHH1	0.558	0.497
PARTTIMEHH2	0.096	0.295
FULLTIMEHH2	0.399	0.489
COLLEGE	0.294	0.456
MALE	0.556	0.497
HISP	0.074	0.262
BLACK	0.096	0.294
CITY	0.435	0.496
SUBURB	0.22	0.414
RURAL	0.18	0.384
NE	0.182	0.386
MW	0.253	0.435
WE	0.193	0.395

Variable	Mean	Std. Dev.
KIDLESS7	0.117	0.322
KIDS7TO15	0.193	0.395
KIDS16TO17	0.068	0.252
NOCHILDLESS35	0.087	0.282
NOCHILD35TO59	0.169	0.376
NOCHILDOVER60	0.167	0.373
SINGLELESS35	0.038	0.19
SINGLE35TO59	0.073	0.261
SINGLEOVER60	0.086	0.281

B.2.2 Vehicle Attributes

Variable	Mean	Std. Dev.
AVGVAGE	8.613	4.519
AVGENG	195.46	72.086
AVGMPG	207.398	46.106
AVGCHNG	0.239	0.358
AVGCOMMUT	0.581	0.419

B.2.3 Fuel Variables

Variable	Mean	Std. Dev.
TOTGALS	990.045	636.555
P_{aggr}	1151.52	87.068

C Results of the Nested Logit Estimation

Table C1: Gasoline Choice With One Car (PC)

Variable	Estimate	Standard Error
LNPGAS_1	0.218	4.466
INCMID	0.033	0.019
INCPREM	0.022	0.016
VAGEMID	-0.051	0.028
VAGEPREM	-0.027	0.023
ENGMID	-0.003	0.002
ENGPREM	-0.0001	0.002
MPGMID	-0.005	0.002
MPGPREM	-0.006	0.002
LUXMID	0.744	0.504
LUXPREM	1.076	0.394
Number of Observations	2085	
Log Likelihood	-590.5	

Note: Vehicle owned is a passenger car. Choice alternatives are regular, midgrade and premium. Regular is the normalized choice category.

Table C2: Gasoline Choice With One Car (LT)

Variable	Estimate	Standard Error
LNPGAS_1	0.036	6.674
INCMID	-0.001	0.048
INCPREM	0.008	0.041
VAGEMID	-0.104	0.055
VAGEPREM	-0.074	0.043
CYLMID	0.038	0.155
CYLPREM	-0.068	0.149
MPGMID	-0.007	0.005
MVANPREM	-2.509	1.055
Number of Observations	396	
Log Likelihood	-98.3	

Note: Vehicle owned is a light truck. Choice alternatives are regular, midgrade and premium. Regular is the normalized choice category.

Table C3: Gasoline Choice With Two Cars (PC, PC)

Variable	Estimate	Standard Error
LNPGAS_1	0.305	2.122
LNPGAS_2	-7.575	1.930
INCRR	0.018	0.019
INCRP	0.045	0.040
INCMM	0.045	0.013
INCPR	0.034	0.032
INCPP	0.078	0.017
VAGERR_1	0.025	0.023
CYLMM_1	-0.004	0.062
CYLRP_2	0.144	0.096
MPGRR_1	0.004	0.002
MPGRR_2	0.004	0.002
MPGRP_2	-0.003	0.003
MPGPR_2	-0.004	0.003
LUXPR_1	1.381	0.480
LUXRP_2	0.857	0.633
Number of Observations	3843	
Log Likelihood	-633.5	

Note: Vehicles owned are two passenger cars. The nine choice alternatives are (regular, regular), (regular, midgrade), ..., (premium, midgrade) and (premium, premium).

Table C4: Gasoline Choice with Two Cars (PC, LT)

Variable	Estimate	Standard Error
LNPGAS_1	0.336	3.023
LNPGAS_2	-7.623	2.774
INCRR	0.018	0.026
INCRP	0.003	0.057
INCMM	0.019	0.022
INCPR	-0.005	0.026
INCPP	-0.057	0.042
VAGERR_1	0.099	0.032
VAGEPR_1	0.044	0.050
ENGPP_1	0.004	0.003
MPGRR_1	0.006	0.002
MPGRP_1	0.004	0.005
MPGPP_2	0.009	0.004
LUXRR_1	-0.513	0.484
LVANRR_2	0.7565	0.681
SPORTMM_2	1.567	0.576
SPORTPP_2	1.242	0.568
Number of Observations	2232	
Log Likelihood	-336.8	

Note: The first vehicle owned is a passenger car, the second vehicle owned is a light truck. The nine choice alternatives are (regular, regular), (regular, midgrade), ..., (premium, midgrade) and (premium, premium).

Table C5: Gasoline Choice With Two Cars (LT, PC)

Variable	Estimate	Standard Error
LNPGAS_1	-4.417	2.686
LNPGAS_2	-2.799	2.733
INCRR	0.024	0.026
INCRP	0.042	0.027
INCMM	0.040	0.017
INCPR	-0.002	0.034
INCPP	0.044	0.029
VAGERP_1	0.067	0.043
VAGERP_2	-0.079	0.056
VAGEPR_2	0.022	0.048
VAGEPP_2	0.047	0.041
MPGRR_2	0.007	0.002
MVANRP_1	-1.189	1.060
MVANPR_1	0.879	0.661
LVANPP_1	1.174	0.764
SPORTPP_1	-1.205	0.836
LUXRR_2	-0.464	0.846
LUXRP_2	0.977	0.888
LUXPP_2	2.011	0.828
Number of Observations	1845	
Log Likelihood	-314.5	

Note: The first vehicle owned is a light truck, the second vehicle owned is a passenger car. The nine choice alternatives are (regular, regular), (regular, midgrade), ..., (premium, midgrade) and (premium, premium).

Table C6: Gasoline Choice With Two Cars (LT, LT)

Variable	Estimate	Standard Error
LNPGAS_1	-4.797	5.263
LNPGAS_2	-3.737	4.431
INCRR	0.084	0.047
INCRM	0.019	0.039
INCMM	0.205	0.101
INCPR	-0.051	0.063
INCPP	0.093	0.039
VAGERR_1	0.129	0.071
VAGEMM_1	0.031	0.147
VAGEPR_2	0.103	0.065
CYLRR_1	-0.033	0.167
CYLMM_1	0.276	0.388
CYLMM_2	-0.824	0.386
LVANRR_1	1.647	1.260
LVANPR_2	2.979	1.216
SPORTRM_2	0.697	0.944
Number of Observations	711	
Log Likelihood	-106.2	

Note: Vehicles owned are two light trucks. The nine choice alternatives are (regular, regular), (regular, midgrade), ..., (premium, midgrade) and (premium, premium).

Table C7: Vehicle Category Choice With One Car

Variable	Estimate	Standard Error
INCL2	0.981	0.795
AGE	0.029	0.008
INC	-0.044	0.022
FAMSIZE	-0.004	0.099
MARRIED	-0.459	0.286
CITY	0.367	0.319
SUBURB	0.104	0.389
RURAL	-0.454	0.389
MPG_1	0.034	0.004
ENG_1	0.004	0.002
AUTOMAT_1	2.148	0.311
COMMUT_1	-0.203	0.266
CONS	-8.621	1.480
Number of Observations	819	
Log Likelihood	-257.4	

Note: The choice alternatives are passenger car and light truck. Light truck is the normalized choice category.

Table C8a: Vehicle Category Choice With Two Cars (PC, LT)

Variable	Estimate	Standard Error
INCL2	0.014	0.211
AGE	-0.012	0.007
INC	0.031	0.018
FAMSIZE	-0.035	0.092
MARRIED	0.819	0.245
CITY	-0.252	0.285
SUBURB	-0.346	0.321
RURAL	0.257	0.347
MPG_1	0.004	0.003
MPG_2	-0.034	0.003
CYL_1	-0.054	0.098
CYL_2	-0.194	0.091
AUTOMAT_1	0.307	0.300
AUTOMAT_2	-2.209	0.257
COMMUT_1	0.153	0.231
COMMUT_2	-0.061	0.227
CONS	8.184	10.559
Number of Obs.	957	
Log Likelihood	-792.3	

Note: The table presents estimated coefficients for the outcome (passenger car, light truck). The four choice alternatives are (passenger car, passenger car), (passenger car, light truck), (light truck, passenger car) and (light truck, light truck). The normalized choice category is (passenger car, passenger).

Table C8b: Vehicle Category Choice With Two Cars (LT, PC)

Variable	Estimate	Standard Error
INCL2	0.014	0.211
AGE	-0.028	0.009
INC	0.091	0.023
FAMSIZE	-0.037	0.098
MARRIED	1.006	0.283
CITY	-0.436	0.317
SUBURB	-0.247	0.345
RURAL	0.909	0.365
MPG_1	-0.043	0.004
MPG_2	0.005	0.003
CYL_1	-0.450	0.107
CYL_2	0.016	0.099
AUTOMAT_1	-1.982	0.297
AUTOMAT_2	0.476	0.299
COMMUT_1	-0.207	0.246
COMMUT_2	-0.019	0.250
CONS	10.244	10.639
Number of Obs.	957	
Log Likelihood	-792.3	

Note: The table presents estimated coefficients for the outcome (light truck, passenger car). The four choice alternatives are (passenger car, passenger car), (passenger car, light truck), (light truck, passenger car) and (light truck, light truck). The normalized choice category is (passenger car, passenger).

Table C8c: Vehicle Category Choice With Two Cars (LT, LT)

Variable	Estimate	Standard Error
INCL2	0.014	0.211
AGE	-0.039	0.013
INC	0.116	0.031
FAMSIZE	0.237	0.129
MARRIED	0.808	0.407
CITY	-0.063	0.451
SUBURB	0.031	0.492
RURAL	1.008	0.507
MPG_1	-0.041	0.005
MPG_2	-0.033	0.005
CYL_1	-0.528	0.149
CYL_2	-0.059	0.136
AUTOMAT_1	-1.702	0.400
AUTOMAT_2	-1.383	0.381
COMMUT_1	-0.681	0.339
COMMUT_2	0.262	0.344
CONS	17.725	10.726
Number of Obs.	957	
Log Likelihood	-792.3	

Note: The table presents estimated coefficients for the outcome (light truck, light truck). The four choice alternatives are (passenger car, passenger car), (passenger car, light truck), (light truck, passenger car) and (light truck, light truck). The normalized choice category is (passenger car, passenger).

Table C9a: Choice of Vehicle Number (One Car)

Variable	Estimate	Standard Error
INCL1	0.682	0.045
AGE	0.0012	0.004
INC	-0.059	0.012
FAMSIZE	-0.285	0.083
MARRIED	-0.793	0.144
KIDSLESS7	0.066	0.241
KIDS7TO15	0.122	0.231
KIDS16TO17	0.473	0.305
PARTTIMEHH1	-0.159	0.218
FULLTIMEHH1	-0.149	0.158
PARTTIMEHH2	-0.915	0.226
FULLTIMEHH2	-0.606	0.149
COLLEGE	0.017	0.137
MALE	0.010	0.126
HISP	0.146	0.218
BLACK	0.183	0.198
CITY	0.309	0.168
SUBURB	0.0992	0.192
RURAL	-0.254	0.199
NE	0.253	0.165
MW	-0.105	0.152
WE	0.003	0.165
CONS	0.846	0.370
Number of Observations	2346	
Log Likelihood	-2207.2	

Note: The table presents estimated coefficients for the outcome (one car). The four choice alternatives are one car, two cars, three cars, four cars. The normalized choice category is (two cars).

Table C9b: Choice of Vehicle Number (Three Cars)

Variable	Estimate	Standard Error
INCL1	0.682	0.045
AGE	0.002	0.006
INC	0.046	0.015
FAMSIZE	0.212	0.079
MARRIED	0.371	0.182
KIDSLESS7	-0.950	0.255
KIDS7TO15	-0.901	0.243
KIDS16TO17	0.478	0.283
PARTTIMEHH1	0.209	0.267
FULLTIMEHH1	0.324	0.187
PARTTIMEHH2	0.312	0.221
FULLTIMEHH2	0.345	0.166
COLLEGE	-0.064	0.148
MALE	0.307	0.148
HISP	-0.012	0.251
BLACK	-0.674	0.286
CITY	-0.280	0.193
SUBURB	0.152	0.202
RURAL	-0.032	0.208
NE	-0.504	0.191
MW	0.165	0.161
WE	-0.355	0.192
CONS	-3.862	0.467
Number of Observations	2346	
Log Likelihood	-2207.2	

Note: The table presents estimated coefficients for the outcome (three cars). The four choice alternatives are one car, two cars, three cars, four cars. The normalized choice category is (two cars).

Table C9c: Choice of Vehicle Number (Four Cars)

Variable	Estimate	Standard Error
INCL1	0.682	0.045
AGE	0.012	0.009
INC	0.075	0.024
FAMSIZE	0.517	0.102
MARRIED	0.366	0.286
KIDSLESS7	-1.373	0.367
KIDS7TO15	-1.444	0.342
KIDS16TO17	0.045	0.379
PARTTIMEHH1	0.442	0.354
FULLTIMEHH1	0.088	0.265
PARTTIMEHH2	1.128	0.353
FULLTIMEHH2	1.424	0.284
COLLEGE	0.148	0.209
MALE	0.148	0.219
HISP	0.251	0.555
BLACK	0.286	0.388
CITY	0.193	0.296
SUBURB	0.202	0.313
RURAL	0.208	0.310
NE	0.191	0.260
MW	0.161	0.231
WE	-0.822	0.311
CONS	-7.468	0.757
Number of Observations	2346	
Log Likelihood	-2207.2	

Note: The table presents estimated coefficients for the outcome (four cars). The four choice alternatives are one car, two cars, three cars, four cars. The normalized choice category is (two cars).

References

AITCHESON, J., AND BROWN, J. A. C. The Lognormal distribution, Cambridge, UK: Cambridge University Press, 1957.

AMEMIYA, T. Advanced Econometrics, Cambridge, MA: Harvard University Press, 1985.

ARCHIBALD, R., AND GILLINGHAM, R. (1980). "An Analysis of the Short-Run Consumer Demand for Gasoline Using Household Survey Data," Review of Economics and Statistics, 62, 622-628.

ARCHIBALD, R., AND GILLINGHAM, R. (1981). "A Decomposition of the Price and Income Elasticities of the Consumer Demand for Gasoline," Southern Economic Journal, 47, 1021-1031.

BERRY, S. T. (1994). "Estimating Discrete-Choice Models of Product Differentiation," RAND Journal of Economics, 25, 242-262.

BERRY, S. T., LEVINSOHN, J., AND PAKES, A. (1995). "Automobile Prices in Market Equilibrium," Econometrica, 63, 841-890.

BERRY, S. T., AND PAKES, A. (2001). "Additional Information for: "Comments on 'Alternative Models of Demand for Automobiles' by Charlotte Wojcik," Economics Letters, 74, 43-51.

CARDELL, N. S. (1997). "Variance Components Structures for the Extreme-Value and Logistic Distributions with Applications to Models of Heterogeneity," Econometric Theory, 13, 185-213.

CARDON, J. H., AND HENDEL, I. (2001). "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey," RAND Journal of Economics, 32, 408-427.

DAHL, C., AND STERNER, T. (1991). "Analysing Gasoline Demand Elasticities: A Survey," Energy Economics, 13, 203-210.

DEATON, A. "Demand Analysis," in: Handbook of Econometrics – Volume III, ed. by Z. Griliches and M. D. Intriligator, Amsterdam; New York; Oxford; Tokyo: North-Holland, 1986.

DUBIN, J. A., AND MCFADDEN, D. L. (1984). "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," Econometrica, 52, 345-362.

ENERGY INFORMATION ADMINISTRATION (EIA) (1997) Household Vehicles Energy Consumption 1994, Washington: Energy Information Administration, DOE/EIA-0464(94).

ESPEY, M. (1998). "Gasoline Demand Revisited: An International Meta-Analysis of Elasticities," Energy Economics, 20, 273-295.

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

HANEMANN, M. W. (1984). "Discrete/Continuous Models of Consumer Demand," Econometrica, 52, 541-561.

HAUSMAN, J. A., AND NEWEY, W. K. (1995). "Nonparametric Estimation of Exact Consumer Surplus and Deadweight Loss," Econometrica, 63, 1445-1476.

HENDEL, I. (1999). "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns," Review of Economic Studies, 66, 423-446.

JOHNSON, N. L., KOTZ, S., AND BALAKRISHNAN, N. Continuous Univariate Distributions – Volume 2, New York; Chichester; Brisbane; Toronto; Singapore: John Wiley & Sons, 2nd Edition, 1995.

JORGENSON, D. W., SLEZNICK, D. T., AND STOKER, T. M. (1988). "Two-Stage Budgeting and Exact Aggregation," Journal of Business and Economic Statistics, 6, 313-325.

MAS-COLLEL, A., WHINSTON, M. D., AND GREEN, J. R. Microeconomic Theory, Oxford: Oxford University Press, 1995

MCFADDEN, D. L."Modelling the Choice of Residential Location," in: Spatial Interaction Theory and Residential Location, ed. by A. Karlquist et al., Amsterdam; New York: North Holland Pub. Co., 1978.

MCFADDEN, D. L."Econometric Models of Probabilistic Choice," in: Structural Analysis of Discrete Data with Econometric Applications, ed. by C. Manski and D. L. McFadden, Cambridge, MA: MIT Press, 1981.

NEVO, A. (2000). "Mergers with Differentiated Products: The Case of Ready-to-Eat Cereal Industry," RAND Journal of Economics, 31, 395-421.

SCHMALENSEE, R., AND STOKER, T. M. (1999). "Household Gasoline Demand in the United States," Econometrica, 67, 645-662.

TRAIN, K. E., MCFADDEN, D., AND BEN-AKIVA, M. (1987). "The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices," RAND Journal of Economics, 18, 109-123.