Dagsvik, J. and Liu, G. (2009): "A framework for analyzing rank-ordered data with application to automobile demand". Transportation Research Part A: Policy and Practice, 43 (1), January 2009, pp 1-12

DOI: http://dx.doi.org/10.1016/j.tra.2008.06.005

Title: **A framework for analyzing rank-ordered data with application to automobile demand**

Author: John K. Dagsvik
Gang Liu

Publisher: Elsevier
Source: **Transportation Research Part A: Policy and Practice**

Author's websites: **http://ideas.repec.org/e/pda154.html**
http://www.ssb.no/english/research/people/liu/index.html

22 May 2008

A Framework for Analyzing Rank Ordered Data with Application to Automobile Demand

by

John K. Dagsvik and Gang Liu

Research Department, Statistics Norway

Abstract

In this paper we develop a general random utility framework for analyzing data on individuals' rank orderings. Specifically, we show that in the case with 3 alternatives one can express the probability of a particular rank ordering as a simple function of first choice probabilities. This framework is applied to specify and estimate models of household demand for conventional gasoline cars and alternative fuel vehicles in Shanghai based on rank ordered data obtained from a stated preference survey. Subsequently, the framework is extended to allow for random effects in the utility specification to allow for intrapersonal correlation in tastes across stated preference questions. The preferred model is then used to calculate demand probabilities and elasticities and the distribution of willingness-to-pay for alternative fuel vehicles.

Keywords: Random utility models, GEV rank ordered models, Car demand, Alternative fuel vehicles

JEL classification: C25, C33, L92

Address of corresponding author: John K. Dagsvik, Research Department, Statistics Norway, P.O.Box 8131, Dep. N-0033 Oslo, Norway. Email: john.dagsvik@ssb.no

Acknowledgment: The authors would like to thank Olav Bjerkholt, Steinar Strøm and Terje Skjerpen for their valuable comments and Anne Skoglund for the word processing of this paper. The assistance provided by students and staff at the Institute of Population Studies at Fudan University , and Mr. Yuan Zhang (sales manager at a Shanghai car dealership) in designing the questionnaire, is gratefully acknowledged. The authors are particularly grateful to Weizhen Zhu for carrying out the programming and estimation for an earlier version of this paper.

1. Introduction

During the last three decades, there has been rapid development of theoretical and empirical approaches to analyzing individual choice behavior of the demand for differentiated products, such as the choice among brands of cars, types of houses, etc. Important contributions in this area are due to McFadden and collaborators; see for example McFadden (2001). Specifically, random utility models have been applied extensively to analyze urban travel behavior. As regards empirical behavioral analysis, the application of data obtained by means of Stated Preference (SP) type of surveys has become increasingly popular, see for example Louviere et al. (2000), Brownstone et al. (2000), Calfee et al. (2001), and Potoglou and Kanaroglou (2007). Recall that by an SP survey, it is understood that individuals in a sample are exposed to hypothetical choice situations. SP data are useful in situations where market transactions are not available to reveal information about individual preferences. Contrary to the conventional revealed preference method, one important advantage of the SP method is that one can obtain several (hypothetical) choice observations for each respondent.

This paper makes two contributions. The main contribution is to show that for a general additive random utility model the probabilities for individuals rank orderings of alternatives in the case with three alternatives can be expressed as a simple function of first choice probabilities. This implies that if the first choice probabilities can be expressed on closed from, such as in the case of the Generalized Extreme Value (GEV) random utility model, the corresponding ranking probabilities can also be expressed in a simple closed form. Second, we apply this framework to analyze the demand for conventional and alternative fuel cars in the city of Shanghai.

In China, the rapid increase in the demand for private cars is an important and sensitive issue. On the one hand, there is the expressed intention of the Chinese government to use the car industry as an engine to promote industrial and economic growth. On the other hand, one realizes the pressing need to adequately address serious pollution problems owing to car traffic in urban areas. There also appears to be growing awareness within China about the role transportation sources play in increasing greenhouse gas emissions. Finally, there is the concern that an uncontrolled increase in the number of private cars may lead to very serious congestion problems. Traffic problems in a number of large cities in developing countries may serve as a warning of what may happen if the increase in private cars in China is not kept under control.

To the authors' best knowledge, studies on car demand undertaken in China are based on historical aggregated data, and these studies are mostly only loosely founded on microeconomic theory, and do not base their analysis on explicitly formulated behavioral models, see e.g. Guo (2001) and Zhai (2000). In contrast, the empirical analysis conducted in this paper represents a first attempt to undertake a behavioral empirical study of the demand for cars in the city of Shanghai, and it is based on micro data and the theory of discrete choice. The data are obtained from a SP survey collected in Shanghai during the summer of 2001. The survey approach we follow is similar to Dagsvik et al. (2002). Specifically, in our survey each household is presented with 15 choice experiments and is asked to rank-order several hypothetical car alternatives characterized by car-specific attributes (price, size, power, fuel consumption) that vary from one choice experiment to the other. We apply the collected data to estimate several model versions within the framework developed here. In the first model version the preferences are assumed independent across experiments, but allowed to be correlated across alternatives. Subsequently, we introduce random effects in the utility specification to allow for time invariant unobserved population heterogeneity in preferences. The estimation results show that this type of heterogeneity is important. Unfortunately, the sample is rather small, and we have therefore only specified and estimated models with rather limited observed population heterogeneity. This is clearly unsatisfactory, and it is desirable to obtain a larger sample in future research.

The behavioral car demand model estimated in this paper enables the prediction of demand and the computation of demand elasticities with respect to price and other car attributes conditional on car attributes. It also allows us to calculate welfare measures such as willingness-to-pay for alternative fuel vehicles (AFV).

The remainder of the paper is structured as follows. In section 2 the theoretical results about rank order probabilities are obtained. Section 3 describes the survey method and the data. In section 4, the empirical specification of the different models as well as the estimation results are presented. In section 5, we present the results on demand and their elasticities. In section 6, we use the model to calculate willingness-to-pay estimates for alternative fuel vehicles.

2. The relationship between first choice-and rank-ordered probabilities in the case with three alternatives

To analyze data on the rank-ordering of alternatives, a particular methodological framework is required. The development of choice models for rank-ordered data originated with work by Luce (1959), Block and Marschak (1960) and Luce and Suppes (1965), whereas Beggs et al. (1981) represents an early application of such models to SP data with observations on the potential demand for electric vehicles.

In this section we shall derive the probabilities of specified rank-orderings when the set of feasible alternatives contain three elements. Previous models for rank-ordering data are often based on the assumption that the random error terms of the utility function are independent across alternatives. A particularly simple expression for the probability of a specific rank-order follows readily when these error terms are i.i. extreme value distributed (see, for example, Beggs et al., 1981). For general random utility models, however, there does not seem to be a simple closed form expression for the ranking probabilities. In this section we show that in the case with three alternatives one can express the ranking probabilities as a simple function of the first choice probabilities. This is of interest in cases where the first choice probabilities can be expressed on closed form, such as in the case when the random terms of the utilities are multivariate extreme value distributed, (the GEV model) because in this case simple closed form expressions for the first choice probabilities exist. Consequently, within the GEV class the corresponding probabilities for rank orderings follow.

Let U_j denote the utility of alternative j, $j = 1,2,3.$ We assume that $U_j = v_j + \varepsilon_j$, where v_j is a deterministic component and ε_j, $j = 1,2,3,$ are random terms (taste shifters) with joint cumulative distribution function (c.d.f.) that is independent of $\{v_j\}$ and is continuously differentiable. Consider the probability that alternative 2 is ranked on top, alternative 1 is the second and alternative 3 is the third preference. Note that since the statement $\{U_1 > \max(U_2, U_3)\}$ and be expressed as $\{(U_2 < U_1) \cap (U_1 > U_3)\},$ it follows that

$$\{U_1 > \max(U_2, U_3)\} \cup \{U_2 > U_1 > U_3\} \Leftrightarrow$$

$$\{(U_2 < U_1) \cap (U_1 > U_3)\} \cup \{(U_2 > U_1) \cap (U_1 > U_3)\} \Leftrightarrow \{U_1 > U_3\}.$$

From this it follows immediately that

(2.1)
$$P\left(U_1 > \max(U_2, U_3)\right) + P\left(U_2 > U_1 > U_3\right) = P(U_1 > U_3).$$

Consequently, (2.1) implies that

$$P\left(U_2 > U_1 > U_3\right) = P\left(U_1 > U_3\right) - P\left(U_1 = \max_{q \le 3} U_q\right).$$

Similarly, we obtain that

(2.2)
$$P\left(U_j > U_k > U_r\right) = P\left(U_k > U_r\right) - P\left(U_k = \max_{q \le 3} U_q\right)$$

for distinct j, k and r. As mentioned above, equation (2.2) is very useful in the case when the random components of the utility function have a multivariate extreme value distribution (GEV model), because in this case the first choice probabilities have simple closed form. Next, we shall therefore discuss the GEV case in more detail.

Consider the special case where the error terms $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ are multivariate extreme value distributed with joint c.d.f. *F*. Let $P_j(B)$ denote the corresponding choice probability defined by

(2.3) $$P_j(B) \equiv P\left(U_j = \max_{q \in B} U_q\right),$$

where *B* is equal to $\{1,2,3\}$, or a subset of $\{1,2,3\}$, which contains at least two elements. Let

(2.4) $$G(v_1, v_2, v_3) \equiv -\log F(-v_1, -v_2, -v_3).$$

Then, by McFadden (1984),

(2.5) $$P_j(\{1,2,3\}) = \frac{\partial G(v_1, v_2, v_3) / \partial v_j}{G(v_1, v_2, v_3)}$$

for $j \in \{1,2,3\}$, and

(2.6) $$P_j(\{1,2\}) = \frac{\partial G(v_1, v_2, -\infty) / \partial v_j}{G(v_1, v_2, -\infty)}$$

for $j \in \{1,2\}$, and similarly for the choice sets $\{1,3\}$ and $\{2,3\}$.

For the sake of the empirical application below we consider next the structure of the probabilities of rank orderings in the special case within the GEV class where the joint c.d.f. of $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ is a particular nested form

(2.7) $$F(x_1, x_2, x_3) \equiv P(\varepsilon_1 \le x_1, \varepsilon_2 \le x_2, \varepsilon_3 \le x_3) = \exp\left(-e^{-x_1} - \left(e^{-x_2/\theta} + e^{-x_3/\theta}\right)^\theta\right).$$

The parameter θ satisfies $0 < \theta \le 1$ and has the interpretation

(2.8) $$Corr(\varepsilon_2, \varepsilon_3) = 1 - \theta^2,$$

whereas ε_1 is independent of ε_2 and ε_3, see for example McFadden (1984). The special case given in (2.7) is of particular interest in our empirical application below. In this application alternative one is “Not buy”, alternative two is “Buy a gasoline car”, and alternative three is “Buy an alternative fuel vehicle (AFV).” In this case it seems reasonable to assume this tree structure a priori because the two car alternatives may be “similar” in the sense that unobserved attributes such as “car ownership” or “taste for driving” may generate correlations between the error terms of the two car alternatives.

From (2.5) and (2.7) it follows that the first choice probability of choosing alternative one from the choice set {1,2,3} equals

(2.9) $$P_1(\{1,2,3\}) = \frac{e^{v_1}}{e^{v_1} + \left(e^{v_2/\theta} + e^{v_3/\theta}\right)^{\theta}}.$$

Similarly, the first choice probability of choosing alternative j from {1,2,3}, equals

(2.10) $$P_j(\{1,2,3\}) = \frac{\left(e^{v_2/\theta} + e^{v_3/\theta}\right)^{\theta-1} e^{v_j/\theta}}{e^{v_1} + \left(e^{v_2/\theta} + e^{v_3/\theta}\right)^{\theta}}$$

for $j = 2,3$, From (2.6) we obtain that the probability of choosing j from {1,k} equals

(2.11) $$P_j(\{1,k\}) = \frac{e^{v_j}}{e^{v_1} + e^{v_k}},$$

for $j = 1, k, \ k = 2,3,$ and the probability of choosing j from {2,3} equals

(2.12) $$P_j(\{2,3\}) = \frac{e^{v_j/\theta}}{e^{v_2/\theta} + e^{v_3/\theta}},$$

for $j = 2,3$. From (2.2), and (2.10) to (2.12), we obtain that the probability of ranking alternative *j* on top, alternative *k* as second best and alternative one as third best, is given by

(2.13) $$P\left(U_j > U_k > U_1\right) = P_k(\{1,k\}) - P_k(\{1,2,3\}) = \frac{e^{v_k}}{e^{v_k} + e^{v_1}} - \frac{\left(e^{v_2/\theta} + e^{v_3/\theta}\right)^{\theta-1} e^{v_k/\theta}}{e^{v_1} + \left(e^{v_2/\theta} + e^{v_3/\theta}\right)^{\theta}}.$$

Similarly, it follows by symmetry that

(2.14) $$P\left(U_j > U_1 > U_k\right) = P_1(\{1,k\}) - P_1(\{1,2,3\}) = \frac{e^{v_1}}{e^{v_1} + e^{v_k}} - \frac{e^{v_1}}{e^{v_1} + \left(e^{v_2/\theta} + e^{v_3/\theta}\right)^{\theta}}$$

and

(2.15) $$P\left(U_1 > U_j > U_k\right) = P_j(\{j,k\}) - P_j(\{1,2,3\}) = \frac{e^{v_j/\theta}}{e^{v_2/\theta} + e^{v_3/\theta}} - \frac{\left(e^{v_2/\theta} + e^{v_3/\theta}\right)^{\theta-1} e^{v_j/\theta}}{e^{v_1} + \left(e^{v_2/\theta} + e^{v_3/\theta}\right)^{\theta}},$$

for $(j,k)=(2,3),(3,2)$. The formulas (2.13) through (2.15) form the basis for specification of one version of the empirical model below and the corresponding likelihood function.

An alternative to the approach above would for example be to use a mixed logit type of specification, see for example Layton (2000), Calfee et al. (2001), and Srinivasan et al. (2006). In this case simulation techniques are necessary for calculating the ranking probabilities. In contrast, the analysis in this section shows that there is a simple relationship between the ranking probabilities and the first choice probabilities in the case with three alternatives and that one therefore can get closed form expressions for the ranking probabilities in cases where the corresponding first choice probabilities can be expressed on closed form.

3. Survey method and data

Contrary to the conventional revealed preference method, one important advantage of the SP method is that one can obtain several (hypothetical) choice observations for each respondent. For example, it can be utilized to elicit information about respondents' complete rank orderings of a set of alternatives. SP surveys also have the advantage over revealed preference data in that one can design the experiments with independent and widely varying conditions and explanatory variables across respondents as well as across experiments for each individual. A problem with analysis based on revealed preference data to measure preferences over car alternatives is that it is hard to describe and represent the actual choice sets of the consumers since the number of different alternatives available in the market is very large. Furthermore, due to bounded rationality, the consumer may in fact only consider a subset of the alternatives in the market, and make the choice from this (subjective and unobservable) “choice” set. In contrast, alternatives and choice sets in SP surveys can be designed and described in a precise manner by the researcher. Of course, this means that it may be hard, or even impossible to apply models estimated by SP data to simulate aggregate market demand, because the joint distribution of the random components (due to unobservable attributes) in SP data may differ from the corresponding distribution in real markets.[1] For further discussions on this issue, see for example Brownstone et al. (2000), Louviere et al. (2000), Calfee et al. (2001), Potoglou and Kanaroglou (2007).

A survey based on the SP approach was conducted in Shanghai during the summer of 2001. Several concerns lead to the use of the SP method instead of the more conventional revealed preference method. First and foremost, market micro data on car demand were not available. Second,

[1] In analysing work-trip mode choice in Shanghai by using a data set drawn from the same survey as presented in this paper, Liu (2007) found that the variation in preferences due to unobservables obtained from SP data were larger than the corresponding estimates obtained from market revealed preference data.

only a few families in Shanghai actually own cars, although more and more people are planning to buy in the near future. In addition, AFVs have not been commercially available in China, whereas in an SP survey, AFV can be presented as one choice alternative. Finally, the SP method is cost effective as a relatively small sample can provide much information. For instance, if appropriately designed, the researcher can obtain data on individual rankings whereas conventional revealed preference methods only yield data for the most preferred alternative. The disadvantage is that households (represented by a single person in the household denoted the respondent) respond to hypothetical questions that are not directly affected by actual budget constraints and other choice restrictions that may apply in the market. (See the Appendix for the survey questionnaire).

The SP survey is based on a series of 15 experimental settings presented to each respondent. In each experiment, we presented the choice setting to the respondent in the form of a card with three choice alternatives, namely “Not buy”, “Buy a gasoline car”, and “Buy an alternative fuel vehicle (AFV)”. Each car alternative was characterized by given attribute values. First, the respondent was required to choose his or her most preferred alternative. If the respondent chose the alternative “Not buy ”, then the following was asked: “Suppose you are able to buy, which one do you prefer between the remaining two alternatives?” If the respondent’s choice were either a gasoline car or an AFV, then we asked the following: “Suppose the alternative you just chose is not available, which one do you prefer among the remaining two alternatives, namely, either the vehicle not chosen in the first place or ‘Not buy’?” By changing the car attributes with sufficient variation from one experiment to the next, a sample of rank-ordered panel data was obtained. The attributes were selected in an arbitrary way,

[Table 1 here]

To ensure the quality of the survey, we implemented face-to-face interviews where the interviewer was able to control the interview process and explain the context presented in the questionnaire as clearly and realistically as possible. In total, 100 households were selected, among which there were 46 male and 54 female respondents. The same interviewer was used for all interviews. The survey comprises three parts, of which only two are relevant to this paper; namely information about basic household characteristics and car demand information. The empirical income distribution of the households in the sample is presented in Table 1. Additional information about age and household size is given in Table 1A in the Appendix. Table 2 displays summary statistics of the survey results. The designed choice experiments are listed in Table 2A in the Appendix. The Appendix also has the detailed information on the survey design, the questionnaire and the sample selection rules.

[Table 2 here]

After discussing with local car sales companies and salesmen on which car attributes are the most important ones, we decided to choose price, power, fuel consumption and size (in terms of number of seats) as the attributes of the cars to be presented in the survey. The range of the four attributes is listed in Table 3.

[Table 3 here]

4. Empirical specification and estimation results

This section contains details of the empirical specification of the different versions of the model as well as the estimation results.

4.1. Model with fixed coefficients (Model 1)

The alternatives are enumerated as follows; "Not buy" (1), "Buy a gasoline car" (2), "Buy an alternative fuel vehicle (AFV)" (3). Recall that each household in the sample is presented with 15 experiments. Let $U_{hj}(t)$ denote the utility of household *h* of alternative *j* in experiment *t*, $j = 1, 2, 3,$ $t = 1, 2, ..., 15.$ Let $Z_{1j}(t)$ represent "User cost", $Z_{2j}(t)$ "Fuel consumption", $Z_{3j}(t)$ "Size" and $Z_{4j}(t)$ "Power", of alternative *j* in experiment *t*, for $j = 1, 2, 3,$ with $Z_{s1}(t) = 0$ for $s = 1, 2, 3, 4,$ and let y_h be the income of household *h*. Note here that we have replaced car price with the corresponding user cost. Because the maximum regulated lifetime of a car in China is 10 years, we define the user cost per month as

$$Z_{1j}(t) = \frac{w_j(t) \cdot \delta}{12(1+\delta)\left(1-(1+\delta)^{-10}\right)},$$

for $j = 2, 3$, where $w_j(t)$ is the purchase price of car *j* in experiment t, and δ is the annual discount rate, assumed to be 10 percent. When $\delta = 0.10$, is inserted in the above formula we get that $Z_{1j}(t) = 0.148 w_j(t)/12.$ The reason why we have converted the purchase price into its corresponding user cost is because we use user cost to assess whether a specific car is available (affordable) to the household. Specifically, we assume that a car is available to the household if the user cost of the car is less than household income. Otherwise, it is unavailable. The rationale of using user cost requires well-functioning financial markets such that the agents can obtain loans to purchase durables at a given interest rate. In Shanghai, the financial markets function quite well in this respect. However, in the survey questionnaires we presented the respective "original" prices to the individuals in our sample. This implies no inconsistency for the model estimation result because user costs are proportional to prices so that in the model, the proportional factor is absorbed in the price coefficient.

Consistently with Section 2, we assume that

$$U_{hj}(t) = v_{hj}(t) + \varepsilon_{hj}(t) \tag{4.1}$$

where $(\varepsilon_{h1}(t), \varepsilon_{h2}(t), \varepsilon_{h3}(t))$, $t = 1, 2, ..., 15$, $h = 1, 2, ..., N$, are i.i.d. vectors with c.d.f. as in (2.7). The motivation for allowing for correlation between the error terms associated with the two car alternatives is that an unobserved variable, such as the taste for driving-or car ownership, is common to alternatives 2 and 3 but not relevant for alternative 1.

The structural term $v_{hj}(t)$ has the interpretation as the indirect utility of alternative *j*. We assume furthermore that

$$v_{hj}(t) = \mu_j + \gamma_1 \left(y_h - Z_{1j}(t) \right) + \gamma_2 Z_{2j}(t) + \gamma_3 Z_{3j}(t) + \gamma_4 Z_{4j}(t), \tag{4.2}$$

for $j = 2,3$, and $v_{h1}(t) = \mu_1 + y_h \gamma_1$. As mentioned above, we assume that car alternative *j* is only available to household *h* if $y_h > Z_{1j}(t)$ for $j = 2,3$. The parameters μ_2 and μ_3 represent the mean pure taste for conventional and alternative fuel vehicles, respectively, and μ_1 is normalized to zero. We expect the parameter γ_2 to be negative whereas γ_1, γ_3 and γ_4 are expected to be positive. Let

$$Q_{hjk}(t) \equiv P\left(U_{hj}(t) > U_{hk}(t) > U_{hr}(t) \right) \tag{4.3}$$

for distinct $j, k, r \in \{1, 2, 3\}$. Thus, $Q_{hjk}(t)$ is the probability that household *h* shall rank alternative *j* on top and alternative *k* second best in experiment *t*. Let $Y_{hjk}(t)$ be equal to one if household *h* ranks alternative *j* on top and alternative *k* second best in experiment *t,* and zero otherwise. The corresponding likelihood function is given by

$$L \equiv \prod_{h=1}^{N} \prod_{t=1}^{15} \prod_{j,k} \left(Q_{hjk}(t) \right)^{Y_{hjk}(t)} \tag{4.4}$$

where $Q_{hjk}(t)$ is obtained by inserting the expression for $v_{hj}(t)$ in (4.2) into the formula for the ranking probabilities given in (2.13) to (2.15). Maximum likelihood estimates are reported in Table 4.

Altogether, this yields 1,500 observations on first choices and 1500 observations on second choices. However, in the sample, there are 9 observations on first choices, and 67 observations on second choices that are inconsistent with the availability criteria that user cost should not exceed income. These observations are removed. The precise sample selection rules are given in the appendix.
From Table 4, we note that "size" appears to be of no importance to the consumers. Also only the mean taste parameters μ_3 is significantly different from zero. The coefficients associated with most of the remaining variables have the expected sign and are relatively precisely determined. As a measure of goodness-of-fit, we used McFadden's ρ^2. Recall that ρ^2 is defined as

(4.5) $$\rho^2 = 1 - \frac{\log \hat{L}}{\log L_0}$$

where $\hat{L}$ is the estimated likelihood and L_0 is the likelihood for the "reference" case, which corresponds to a completely random ranking. That is, in this reference case the choice probabilities for the first choice are equal to 1/3 and the choice probabilities for the second choice are equal to 1/2. The estimate of θ implies that the correlation between $\varepsilon_2(t)$ and $\varepsilon_3(t)$ is equal to about 0.75, cf. (2.8).

[Table 4 here]

We have also experimented with a more general specification of the functional form of the utility function. Specifically, we have postulated a so-called Box–Cox type of specification given by

(4.2)* $$v_{hj}(t) = \mu_j + \gamma_1 \left(\frac{\left(y_h - Z_{1j}(t) + \beta\right)^{\alpha} - 1}{\alpha} \right) + \gamma_2 Z_{2j}(t) + \gamma_3 Z_{3j}(t) + \gamma_4 Z_{4j}(t)$$

where $\alpha \le 1$ and $\beta \ge 0$. Note that when $\alpha = 1$, the Box–Cox specification reduces to the specification in (4.2), whereas when $\alpha = 0$, the Box–Cox transformation $\left(x^{\alpha} - 1\right)/\alpha$ becomes equal to $\log x$. After some experimentation, we concluded that α values different from 1 appear to yield lower likelihood values than when $\alpha = 1$, as assumed in the analysis in this paper.

4.2. Model with random technology parameters (Model 2)

In this section, we extend the model considered above by allowing the technology parameters $\{\mu_j\}$ to be individual specific and distributed according to the normal distribution. Thus we now assume that

(4.6) $$v_{hj}(t) = \mu_{hj} + \gamma_1(y_h - Z_{1j}(t)) + \gamma_2 Z_{2j}(t) + \gamma_3 Z_{3j}(t) + \gamma_4 Z_{4j}(t)$$

where $\mu_{hj} = \bar{\mu}_j + \sigma_j \eta_{hj}$, and η_{jh}, $j = 1,2,3$, $h = 1,2,...,N$, are i.i.d. standard normally distributed and $\bar{\mu}_j$ and $\sigma_j > 0$, $j = 1,2,3$, are parameters to be estimated. As above, we can normalize so that $\bar{\mu}_1 = 0$. In this model version there are altogether 10 parameters to be estimated, namely θ, $\bar{\mu}_2, \bar{\mu}_3$, σ_1, σ_2, σ_3, γ_1, γ_2, γ_3 and γ_4. Whereas Model 5 in Dagsvik et al. (2002), in addition to random effects, allows for random taste shifters that are serially correlated, our Model 2 assumes serially independent taste shifters, $\{\varepsilon_{hj}(t)\}$.

For notational convenience let $\boldsymbol{\eta}_h = (\eta_{h1}, \eta_{h2}, \eta_{h3})$ and let $Q_{hjk}(t, \boldsymbol{\eta}_h)$ denote the probability obtained from $Q_{hjk}(t)$ by replacing $\bar{\mu}_j$ by η_{hj}, for given $\boldsymbol{\eta}_h$. The corresponding conditional likelihood for household h, given $\boldsymbol{\eta}_h$, equals

(4.7) $$L_h(\boldsymbol{\eta}_h) \equiv \prod_{t=1}^{15} \prod_{j,k} Q_{hjk}(t, \boldsymbol{\eta}_h)^{Y_{hjk}(t)}.$$

The total unconditional likelihood is therefore equal to

(4.8) $$\tilde{L} \equiv \prod_{h=1}^{N} E\, L_h(\boldsymbol{\eta}_h)$$

where the expectation in (4.8) is taken with respect to $\boldsymbol{\eta}_h$. To compute (4.8), the following simulation procedure is practical. Draw M independent vectors η_h^m, $m = 1,2,...,M$, with i.i.d. standard normal components. Then the approximation

(4.9) $$E\, L_h(\eta_h) \approx \frac{1}{M} \sum_{m=1}^{M} L_h(\eta_h^m) \equiv L_h^*$$

is good when M is large, and consequently one can obtain consistent estimates by maximizing

(4.10) $$\log L^* \equiv \sum_{h=1}^{N} \log L_h^*$$

with respect to the unknown parameters. We found that M = 10 000 was sufficient to eliminate variations in estimates due to simulation of the likelihood. Lower values of M seem to introduce additional uncertainty in the estimates.

In Table 5, we report the estimation results based on model specification in (4.6). Recall that the parameter θ is restricted to the interval, $0 < \theta \le 1$. In fact, the estimation procedure yields an estimate that equals the upper boundary, that is $\theta = 1.$ Thus, we conclude that the random effects in Model 2 in fact account for the correlation-across-alternatives effect found for Model 1. It is perhaps surprising that the correlation between the error terms of the utilities of the alternative fuel and the conventional car is zero $(\theta = 1)$ because a reasonable *a priori* conjecture is that the two car alternatives could be close substitutes. Moreover, we see that the parameter associated with the user cost is rather sharply determined, and "Power" and "Fuel consumption" appear to matter, whereas "Size" is barely significant. The alternative specific constant $\bar{\mu}_3$ is found to be significantly different from zero, whereas $\bar{\mu}_2$ is not. Similarly to the results above, this also suggests that Shanghai households are, on average, slightly concerned about the possible environmental benefits of alternative fuel cars. We have used the Likelihood ratio test to test the null hypothesis, $\bar{\mu}_2 = \bar{\mu}_3,$ against the alternative, $\bar{\mu}_2 \neq \bar{\mu}_3.$ Let L^0 denote the likelihood under the null hypothesis, $\bar{\mu}_2 = \bar{\mu}_3.$ It follows that $2(\log L^* - \log L^0) = 9.18$. The corresponding critical value at 5 per cent significance level in the Chi square distribution with one degree of freedom is 3.84. Hence, we conclude that the null hypothesis is clearly rejected. However, the variation in the random effects $\{\eta_{ij}\}$ is substantial. This means that some households may value alternative fuel technology higher than gasoline technology and vice versa. Comparison of the results of Tables 4 and 5 reveals that the estimates of Table 5 are roughly 3.3 times higher than the corresponding estimates in Table 4. The estimate of the standard errors (σ) of the random effect associated with AFV is lower than the estimate of the standard error associated with conventional fuel vehicles. The random effect associated with the "no purchase" decision has very large variance, which may reflect the fact that we have no information about the user cost of other durables and fixed family expenditures such as housing, etc. Compared with Model 1, Model 2 implies a substantial increase in goodness-of-fit; ρ^2 increases from 0.23 to 0.55. In order to check if our model specification should depend on income we estimated the model separately for high- and low-income families. The high-income group consists of those with a monthly income above 5,750 yuan, while low-income families have a monthly income between 1,000 and 5,750 yuan. There are 52 households in the low-income group and 48 in the high-income group. The corresponding estimates are reported in Tables 6 and 7. When comparing Tables 5 to 7, we realize that the estimates are quite similar (when taking into account the standard deviation), apart from the parameter estimates associated with the random effects. For the high-income group $\bar{\mu}_2$ is not significantly different from zero, whereas $\bar{\mu}_3$ is significantly different from zero, in contrast to the low-income households who seem to be indifferent between AFV and conventional cars. This result indicates that high-income households may put greater value on car ownership than low-income households. Also in this case we have used the

Likelihood ratio test to test the null hypothesis, $\bar{\mu}_2 = \bar{\mu}_3$ against the alternative, $\bar{\mu}_2 < \bar{\mu}_3$, for the high income group. The corresponding difference in the Loglikelihood times 2 equals 4.00. Thus, also in this case the null hypothesis is rejected at 5 per cent significance level. The result that high income households appear to be more concerned with environmental issues may be due to the fact that they have higher education than low income households and are therefore better informed.

[Table 5 here]

[Table 6 here]

[Table 7 here]

5. Conditional demand and elasticities

The model estimated above is a disaggregate model that is supposed to capture individual (or household) behavior with respect to demand for cars, conditional on household income and attributes of the vehicles. The estimated model can now be applied to predict the demand for cars and to calculate elasticities for specified levels of the attributes, conditional on attribute values and household income. Recall that our behavioral model is conditional on the choice set of alternatives, which means that after the model have been estimated we may use the model for prediction under alternative choice sets. We shall now use the model in a simulation exercise to assess how the preferences for buying conventional gasoline cars change when attributes change when only the alternatives “Not buy”, and “Buy a gasoline car” are compared. This situation corresponds to the case where AFVs are not available in the market. In Table 8 below, we display the predicted choice probabilities and corresponding elasticities among those households that can afford to buy a car. Here we only assume that conventional gasoline cars are available in the market. The probability of buying a car can in this case be expressed as

(5.1)
$$P_2(\mathbf{Z}_2) \equiv E\left\{\frac{\exp(\mu_{h2} - Z_{12}\gamma_1 + Z_{22}\gamma_2 + Z_{32}\gamma_3 + Z_{42}\gamma_4)}{\exp(\mu_{h1}) + \exp(\mu_{h2} - Z_{12}\gamma_1 + Z_{22}\gamma_2 + Z_{32}\gamma_3 + Z_{42}\gamma_4)}\right\}$$

where the expectation is taken with respect to $(\mu_{h1}, \mu_{h2}) = (\sigma_1\eta_{h1}, \bar{\mu}_2 + \sigma_2\eta_{h2})$, and the attribute vector is equal to $\mathbf{Z}_2' = (Z_{12}, Z_{22}, Z_{32}, Z_{42})'$. Note that the income variable cancels in the choice probabilities. To compute (5.1), we use stochastic simulation similarly to the simulation of the likelihood function in Section 4.2.

[Table 8 here]

From (5.1), it follows that the elasticity with respect to attribute component Z_{s2}, $s = 1,2,3$ and 4, can be expressed as

(5.2)

$$\frac{\partial \log P_2(\mathbf{Z}_2)}{\partial \log Z_{s2}} = Z_{s2}\gamma_s \left[1 - E\left(\frac{\exp(\mu_{h2} - Z_{12}\gamma_1 + Z_{22}\gamma_2 + Z_{32}\gamma_3 + Z_{42}\gamma_4)}{\exp(\mu_{h1}) + \exp(\mu_{h2} - Z_{12}\gamma_1 + Z_{22}\gamma_2 + Z_{32}\gamma_3 + Z_{42}\gamma_4)} \right)^2 \frac{1}{P_2(\mathbf{Z}_2)} \right].$$

From Table 8 we see that when, for example, the price equals 145,000 yuan, power is 120 hp, and fuel consumption is 1.2 (size does not matter) then the probability of buying a car for those who can afford a car is predicted to be 16.8 percent. The corresponding price elasticity in this case is -1.39.

6. The value of alternative fuel vehicles

By means of the estimated model it is possible to assess the value of AFVs as measured in money metric amounts. Specifically, this means how much it is necessary to reduce the user cost for a household so that the utility of conventional cars is equal to the utility of AFV, given that the attributes of both types of cars are the same. Let K_h denote the amount of reduction of user cost for household h, given that the non-pecuniary attributes remain fixed. It follows directly from (4.1) and (4.6) that this amount is determined by

(6.1) $$K_h = \frac{\mu_{h3} - \mu_{h2} + \varepsilon_{h3} - \varepsilon_{h2}}{\gamma_1}.$$

Due to the distributional assumptions of the error terms $\{\varepsilon_{hj}\}$, the difference $\varepsilon_{h3} - \varepsilon_{h2}$ will be logistically distributed. Hence, for all real x,

(6.2) $$P(K_h \le x) = E\left(\frac{1}{1 + \exp(\mu_{h3} - \mu_{h2} - \gamma_1 x)} \right)$$

where the expectation is taken with respect to $\mu_{h3} - \mu_{h2}$. Moreover, from (6.2), the fraction of households with positive compensating amount equals

(6.3) $$P(K_h > 0) = 1 - E\left(\frac{1}{1+\exp(\mu_{h3}-\mu_{h2})}\right) = E\left(\frac{1}{1+\exp(\mu_{h2}-\mu_{h3})}\right).$$

Thus, (6.3) expresses the fraction of households that value AFV higher than conventional fuel cars, *ceteris paribus*. Note that both (6.2) and (6.3) take into account both the random taste shifters and unobserved population heterogeneity in preferences across agents, which is represented by the terms $\mu_{h3} - \mu_{h2}$. It follows furthermore from (6.1) that

(6.4) $$EK_h = \frac{E\mu_{h3} - E\mu_{h2}}{\gamma_1} = \frac{\bar{\mu}_3 - \bar{\mu}_2}{\gamma_1}$$

and

(6.5) $$Var\, K_h = \frac{\pi^2/3 + Var(\mu_{h3} - \mu_{h2})}{\gamma_1^2} = \frac{\pi^2/3 + \sigma_2^2 + \sigma_3^2}{\gamma_1^2}.$$

Although our estimates imply that the mean Compensating Variation of user cost is positive, it is small (331 yuan).[2] However, the individual values of K_h may be both positive and negative. The estimate of the standard deviation, calculated by means of (6.5) is found to be 632 yuan. The fraction of households that value AFV higher than conventional fuel cars (based on (6.3)) is estimated as 0.67.

7. Conclusion

In this paper we have derived a new relationship between the probabilities for rank orderings and first choice probabilities in general random utility models with three alternatives, with particular reference to the General Extreme Value random utility model. We have applied this framework to analyze the demand for conventional and alternative fuel vehicles in the city of Shanghai. Specifically, the model is estimated on data obtained from a stated preference survey conducted in Shanghai in 2001. Subsequently, we extend the model to allow for random effects in the utility function. The estimates of the model version with random effects show that there is considerable unobserved population heterogeneity. Furthermore, when we allow for random effects the correlation between the random taste shifters across alternatives vanishes. We have also estimated the model for high income- and low-income groups separately, and found that the estimates are not very different across the two groups. Specifically, the results indicate that high-income households seem to value AFV higher than conventional cars, in contrast to low-income households who seem to be indifferent between these two

[2] The middle rate of the average exchange rate of RMB yuan against US dollar in 2001 is 1US$=8.28 yuan (*China Statistical Yearbook*, 2002).

types of cars. Due to the limited sample size one must, however, be cautious with interpretation of these results. Measured in terms of McFadden's ρ^2 the fit of the maintained model turns out to be good. This model is used to calculate elasticities and choice probabilities for selected attributes for those who can afford to own a car. We have also discussed and illustrated how choice probabilities can be calculated, and have employed the model to calculate willingness-to-pay estimates. These estimates show that 67 percent of the households in our sample value AFV vehicles higher than conventional fuel vehicles.

Since 2001 when the SP survey was conducted the increase in the number of households that have own car has increased rapidly, mainly because of changes in the income distribution for urban households. For example, the number of private cars in Beijing has more or less doubled in five years. At the same time China has developed production of competitive electric and hybrid alternative fuel vehicles. Due to increased congestion and pollution problems that follow from the increase in car traffic one may suspect that peoples attitude towards AFV may be changing. An interesting topic for future research is to conduct new SP studies with larger sample sizes to assess whether preferences have change, and whether there are differences in preferences across selected population groups and across different cities in China.

References

Beggs, S., Cardell, S., Hausman, J., 1981. Assessing the Potential Demand for Electric Cars. Journal of Econometrics 16, 1–19.

Block, H., Marschak, J., 1960. Random Orderings and Stochastic Theories of Response, In I. Olkin, S. Ghurye, W. Hoeffding, W. Madow and H. Mann (Eds): *Contributions to Probability and Statistics.* Stanford University, Stanford.

Brownstone, D., D. S. Bunch and K. Train (2000): Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-fuel Vehicles. Transportation Research Part B 34, 315-338.

Calfee, J., Winston, C., Stempski, R., 2001. Econometric Issues in Estimating Consumer Preferences from Stated Preference Data: A Case Study of the Value of Automobile Travel Time. Review of Economics and Statistics 83, 699-707.

China Statistical Yearbook, 2002, China Statistics Press.

Dagsvik, J.K., Wennemo, T., Wetterwald, D.G., Aaberge, R., 2002. Potential Demand for Alternative Fuel Vehicles. Transportation Research Part B 36, 361–384.

Guo, K., 2001. Analysis on the Development Conditions of Market, Economic and Technology of China's Auto Industry. Management World 2, 102–111.

Layton, D.F., 2000. Random Coefficient Models for Stated Preference Surveys. Journal of Environmental Economics and Management 40, 21-36.

Liu, G., 2007. A behavioral model of work-trip mode choice in Shanghai. China Economic Review 18, 456-476.

Louviere, J. J., D.A. Hensher and J. D. Swait (2000): *Stated Choice Methods: Analysis and Applications*, Cambridge University Press, Cambridge, UK.

Luce, R.D., 1959. *Individual Choice Behavior*. J. Wiley, New York.

Luce, R.D., Suppes, P., 1965. Preference, Utility and Subjective Probability. In R.D. Luce, R.R. Bush and E. Galanter (Eds): *Handbook of Mathematical Psychology*, Vol. III, Wiley, New York.

McFadden, D., 1984. Econometric Analysis of Qualitative Response Models. In Z. Griliches and M. Intrilligator (eds): *Handbook of Econometrics*, Vol. II, pp. 1396–1456, Elsevier, Amsterdam.

McFadden, D., 2001. Economic Choices. The American Economic Review 91, 351–378.

Potoglou, D. and P. S. Kanaroglou (2007): Household Demand and Willingness to Pay for Clean Vehicles. Transportation Research Part D 12, 264-274.

Srinivasan, S., Bhat, C.R., Holquin-Veras, J., 2006. Empirical Analysis of the Impact of Security Perception on Intercity Mode Choice: A Panel Rank-Ordered Mixed-Logit Model. Transportation Research Record no. 1942, 9-15.

Zhai, F., 2000. Forecast of Auto Demand in 10 Years Ahead in China. In H. Ma and M. Wang (eds): *Research on China's Development*, 121–139, Development Research Centre, State Council, China Development Publishing House.

Appendix

Information about the survey and the questionnaire

The interviewer is supposed to explain to the respondent about the conventional gasoline car and alternative fuel vehicle (AFV) and their relevant attributes as carefully and sufficiently as possible.

There are 15 experiments. In each experiment, the interviewer will present an option card on which there is a choice set of three alternatives, i.e., “Not buy”; “Buy a gasoline car”; “Buy an alternative fuel vehicle (AFV)”. In each experiment, the gasoline car and AFV will have a different attribute combination.

First, among the three alternatives, the respondent is asked which one is the most preferred. If the respondent has chosen the alternative of “Not buy”, then she/he is asked the following question: Suppose you are able to buy, which one do you prefer between the remaining two alternatives?

If the respondent has chosen the alternative of either “Buy a gasoline car” or “Buy an alternative fuel vehicle (AFV)”, then she/he is asked the following question: Suppose the alternative you just chose is not available, which one do you prefer among the remaining two alternatives (that is: either “Buy a gasoline car” or “Buy an alternative vehicle (AFV)”, whichever alternative that was not chosen in your first option, and the option of “Not buy”).

[Table 1A here]

Sample selection rules

The rules for removing some observations are:

1. Remove first (chosen) choices when;

   Income is less than user cost of alternative 2 and alternative 2 was chosen;

   Income is less than user cost of alternative 3 and alternative 3 was chosen.

This implies that 9 observations were removed.

2. Remove second (chosen) choices in cases where;

   Income is less than user cost of alternative 2 and first choice was “buy” and second choice was alternative 2;

   Income is less than user cost of alternative 3 and first choice was “buy” and second choice was alternative 3.

This implies that 3 observations were removed.

3. Remove second (chosen) choices when;

    Income is less than user cost of alternative 2 and first choice was “not buy” and second choice was alternative 2;

    Income is less than user cost of alternative 3 and first choice was “not buy” and second choice was alternative 3.

This implies that 51 observations were removed.

4. Remove observations when;

    Income is less than user cost of alternative 2 and income<user cost of alternative 3 and first choice was “not buy”;

    First choice was “not buy” and income< user cost of alternative 2 and second choice was alternative 3;

    First choice was “not buy” and income <user cost of alternative 3 and second choice was alternative 2.

This implies that 13 observations were removed.

[Table 2A here]

Table 1. Sample income distribution

Income group	Income range (yuan/month)	Frequency
1	< 1,000	0
2	1,000–2,000	5
3	2,000–3,000	7
4	3,000–4,000	12
5	4,000–5,000	14
6	5,000–6,500	14
7	6,500–8,000	14
8	8,000–9,500	13
9	9,500–11,000	9
10	11,000–12,500	4
11	12,500–14,000	3
12	14,000–15500	2
13	> 15,500	3
Sum		100

Table 2. Summary statistics of the survey results

Experiment		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
First choice	# of "Not Buy"	76	73	72	54	54	69	66	56	75	64	73	67	75	65	68
	# of "Buy Gas"	19	15	8	31	5	5	23	31	7	21	8	9	9	26	5
	# of "Buy AFV"	5	12	20	15	41	26	11	13	18	15	19	24	16	9	27
	Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Second choice	# of "Not Buy"	14	15	12	26	18	19	20	22	13	12	16	11	16	18	13
	# of "Buy Gas"	44	43	28	41	32	18	42	39	16	33	19	23	12	41	26
	# of "Buy AFV"	42	42	60	33	50	63	38	39	71	55	65	66	72	41	61
	Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Table 3. Range of car attributes presented in the choice experiments (option cards)

	Price (1,000 yuan)	Power (Horsepower)	Fuel consumption (Liter/100km)	Size (Number of seats)
Range (Gasoline car)	60–200	90–140	6–12	4–7
Range (AFV)	80–250	90–150	2–4	4–7

Table 4. Parameter estimates of Model 1*

Attribute	Parameter	Estimate	Standard error	t-statistic
User cost	γ_1	0.910	0.077	11.77
Fuel consumption	γ_2	-0.510	0.160	-3.18
Size	γ_3	-0.057	0.043	-1.33
Power	γ_4	0.472	0.167	2.82
Mean taste (gasoline car)	μ_2	0.170	0.158	1.08
Mean taste (alternative fuel vehicle)	μ_3	0.492	0.149	3.29
Correlation parameter	θ	0.502	0.028	18.10
Log likelihood		-2,014.74		
McFadden's ρ^2		0.23		
Number of observations (first choices)		1,491		
Number of observations (second choices)		1,433		

* The unit of user cost is 1,000 yuan, fuel consumption is measured in liters per 10 km, and the unit of power is 100 horsepower.

Table 5. Parameter estimates of Model 2*

Attribute	Parameter	Estimate	Standard error	t-value
User cost	γ_1	3.068	0.174	17.63
Fuel consumption	γ_2	-1.661	0.380	-4.37
Size	γ_3	-0.212	0.101	-2.10
Power	γ_4	1.648	0.373	4.42
Mean random effect (gasoline car)	$\bar{\mu}_2$	0.274	0.705	0.39
Standard error of random effect (gasoline car)	σ_2	2.225	0.282	7.89
Mean random effect (alternative fuel vehicle)	$\bar{\mu}_3$	1.288	0.629	2.05
Standard error of random effect (alternative fuel vehicle)	σ_3	0.759	0.371	2.05
Standard error of random effect (Not buy)	σ_1	5.160	0.560	9.21
Correlation parameter	θ	1		
McFadden's ρ^2		0.55		
Log-likelihood		-1,197.18		
Number of observations (first choices)		1,491		
Number of observations (second choices)		1,433		
Number of draws	M	10,000		

* The unit of user cost is 1,000 yuan, fuel consumption is measured in liters per 10 km, and the unit of power is 100 horsepower.

Table 6. Parameter estimates of Model 2. Low-income group*

Attribute	Parameter	Estimate	Standard error	t-value
User cost	γ_1	3.051	0.254	12.01
Fuel consumption	γ_2	-2.037	0.552	-3.69
Size	γ_3	-0.261	0.142	-1.84
Power	γ_4	1.911	0.535	3.57
Mean random effect (gasoline car)	$\bar{\mu}_2$	-0.945	1.024	-0.92
Standard error of random effect (gasoline car)	σ_2	1.930	0.296	6.52
Mean random effect (alternative fuel vehicle)	$\bar{\mu}_3$	0.082	0.960	0.09
Standard error of random effect (alternative fuel vehicle)	σ_3	0.347	0.528	0.66
Standard error of random effect (Not buy)	σ_1	4.702	0.785	5.99
Correlation parameter	θ	1		
McFadden's ρ^2		0.59		
Log-likelihood		-553.00		
Number of observations (first choices)		771		
Number of observations (second choices)		713		
Number of draws	M	10,000		

* The unit of user cost is 1,000 yuan, fuel consumption is measured in liters per 10 km, and the unit of power is 100 horsepower.

Table 7. Parameter estimates of Model 2. High-income group*

Attribute	Parameter	Estimate	Standard error	t-value
User cost	γ_1	3.092	0.239	12.94
Fuel consumption	γ_2	-1.294	0.526	-2.46
Size	γ_3	-0.164	0.142	-1.15
Power	γ_4	1.411	0.523	2.70
Mean random effect (gasoline car)	$\bar{\mu}_2$	1.521	0.875	1.74
Standard error of random effect (gasoline car)	σ_2	2.573	0.376	6.84
Mean random effect (alternative fuel vehicle)	$\bar{\mu}_3$	2.616	0.847	3.09
Standard error of random effect (alternative fuel vehicle)	σ_3	0.971	0.536	1.81
Standard error of random effect (Not buy)	σ_1	5.214	0.773	6.75
Correlation parameter	θ	1		
McFadden's ρ^2		0.51		
Log-likelihood		-637.57		
Number of observations (first choices)		720		
Number of observations (second choices)		720		
Number of draws	M	10,000		

* The unit of user cost is 1,000 yuan, fuel consumption is measured in liters per 10 km, and the unit of power is 100 horsepower.

Table 8. Elasticities and choice probabilities by car attributes given that income is higher than user cost*

Car attributes					Elasticity			Choice probability	
Price	User cost	Power	Fuel	Size	Price	Power	Fuel	Not buy	Buy
60	0.740	1.2	1.2	2	-0.415	0.361	-0.364	0.658	0.342
110	1.356	1.2	1.2	2	-0.924	0.439	-0.442	0.770	0.230
145	1.788	1.2	0.8	2	-1.291	0.466	-0.313	0.800	0.200
145	1.788	1.2	1.2	2	-1.389	0.501	-0.505	0.832	0.168
145	1.788	1.0	1.2	2	-1.423	0.428	-0.517	0.844	0.156
160	1.973	1.2	1.2	2	-1.613	0.527	-0.531	0.854	0.146
180	2.219	1.2	1.2	2	-1.916	0.557	-0.561	0.881	0.119
200	2.466	1.2	1.2	2	-2.259	0.590	-0.595	0.904	0.096
200	2.466	1.0	0.8	2	-2.217	0.483	-0.389	0.895	0.105
200	2.466	1.0	1.2	2	-2.291	0.499	-0.604	0.912	0.088

* The unit of user cost is 1,000 yuan, fuel consumption is measured in liters per 10 km, and the unit of power is 100 horsepower.

Table 1A. Some descriptive statistics of household sample basic information

	Age	Household size (persons)
Mean	33.52	2.62
Standard error	0.96	0.10
Minimum	21	1
Maximum	66	6

Table 2A. Experiment design

Experiment	Type of car	Price (1,000 yuan)	Power (Horse Power)	Fuel consumption (Liters/100 km)	Size (Number of Seats)
1	Gasoline car	145	120	12	6 ~7
	AFV	250	100	2	4 ~5
2	Gasoline car	120	110	8	6 ~7
	AFV	200	150	2	6 ~7
3	Gasoline car	170	125	6	4 ~5
	AFV	155	95	4	4 ~5
4	Gasoline car	60	90	6	6 ~7
	AFV	155	120	2	4 ~5
5	Gasoline car	100	100	10	6 ~7
	AFV	80	95	3	4 ~5
6	Gasoline car	180	140	10	6 ~7
	AFV	125	90	2	6 ~7
7	Gasoline car	105	100	8	6 ~7
	AFV	180	100	2	4 ~5
8	Gasoline car	70	90	7	4 ~5
	AFV	145	120	4	6 ~7
9	Gasoline car	175	125	10	6 ~7
	AFV	155	90	2	6 ~7
10	Gasoline car	100	90	6	4 ~5
	AFV	140	120	3	4 ~5
11	Gasoline car	180	110	6	6 ~7
	AFV	155	95	3	4 ~5
12	Gasoline car	120	90	8	6 ~7
	AFV	130	120	4	4 ~5
13	Gasoline car	200	130	6	6 ~7
	AFV	150	100	4	4 ~5
14	Gasoline car	100	120	10	6 ~7
	AFV	160	100	3	6 ~7
15	Gasoline car	150	120	8	4 ~5
	AFV	130	90	2	6 ~7