Paper:
Hess, S, Fowler, M, Adler, T and Bahreinian, A (2012) *A joint model for vehicle type and fuel type choice: evidence from a cross-nested logit study.*
Transportmetrica, 39 (3). 593 - 625.

http://dx.doi.org/10.1007/s11116-011-9366-5

A joint model for vehicle type and fuel type choice: evidence from a cross-nested logit study

Stephane Hess
Institute for Transport Studies
University of Leeds
s.hess@its.leeds.ac.uk

Mark Fowler
Resource Systems Group
mfowler@rsginc.com

Thomas Adler
Resource Systems Group
tadler@rsginc.com

Aniss Bahreinian
California Energy Commission
ABahrein@energy.state.ca.us

Abstract

In the face of growing concerns about greenhouse gas emissions, there is increasing interest in forecasting the likely demand for alternative fuel vehicles. This paper presents an analysis carried out on stated preference survey data on California consumer responses to a joint vehicle type choice and fuel type choice experiment. Our study recognises the fact that this choice process potentially involves high correlations that an analyst may not be able to adequately represent in the modelled utility components. We further hypothesise that a Cross-Nested Logit structure can capture more of the correlation patterns than the standard Nested Logit model structure in such a multi-dimensional choice process. Our empirical analysis and a brief forecasting exercise produce evidence to support these assertions. The implications of these findings extend beyond the context of the demand for alternative fuel vehicles to the analysis of multi-dimensional choice processes in general. Finally, an extension verifies that further gains can be made by using mixed GEV structures, allowing for random heterogeneity in addition to the flexible correlation structures.

Keywords: cross-nested logit; vehicle type choice; fuel type choice; alternative fuel; stated preference

1. Introduction

Growing environmental concerns and oil price volatility have led to increasing interest in the potential demand for alternative fuel vehicles. Dedicated fuel vehicles such as EV and CNG vehicles use only the alternative fuel, and flex fuel vehicles can mix conventional and alternative fuels in the same fuel tank. There has long been interest in modelling the potential consumer response to the introduction of such vehicles, with examples including Train (1983), Bunch et al. (1993), Train (1993), Golob et al. (1995), Kavalece (1996), Tomkins et al. (1998), Greene (2001), Batley and Toner (2003), Batley et al. (2004), Adler et al. (2004), and Spissu et al. (2009) to name but a few. There is also interest in the long term response to changes in fuel prices, see e.g. the stated adaption work of Erath & Axhausen (2010).

There are a number of ways to use alternative fuels, including a mix of alternative and conventional fuels in flex fuel and hybrid vehicles, as with ethanol and gasoline or electricity and gasoline, as well as the sole use of alternative fuels, in full electric and compressed natural gas vehicles, for instance. The preferences for different types of fuels are difficult to predict, not least because of the strong relationship between fuel type and other attributes such as performance, annual costs and incentives (e.g. tax breaks). At the same time, there is a very strong link between fuel type and vehicle type, with certain types of fuels being more or less appropriate for specific vehicle types.

In this paper, we discuss work based on the 2008-09 California Vehicle Survey (CVS), aimed at providing input data for the California (light-duty) Conventional and Alternative Fuel Response Simulator (CALCARS) model at the California Energy Commission (CEC). The 2008-09 CVS collected data on both stated and revealed preferences of vehicle owners in California, to forecast their vehicle choice and the use of both conventional and alternative fuel vehicles.

In the present paper, we focus on the stated preference survey component of this work, and specifically the estimation of models for the choice of vehicle type and fuel type. The results from this estimation exercise were then later used in a forecasting model which aggregates California households into 364 synthetic households, differentiated by household size, income, number of workers and vehicle ownership and distributes household population among these synthetic households using census data. This model then computes the probability of choosing from 105 vehicle classes (fifteen vehicle types times seven fuel types) for each of these 364 households that plan to buy a vehicle in year *t*. Furthermore, the results are used as inputs into Dynasim, a dynamic transportation fuel simulator that is composed of vehicle and travel models representing different transportation sectors in California, and which is used to forecast transportation fuel consumption in California. Using the results from the present paper as well as corresponding results for multi-car households, Dynasim forecasts demand for light duty vehicles by the household sector. For each of the 364 synthetic households, the vehicle transaction model computes the probability that the household will own more, less, or the same number of vehicles in year *t*, based on changes in each household's attributes between last period (*t-1*) and this year (*t*) as well as the probability of adding a vehicle. The vehicle choice model then determines the household's choice probability for each of the 105 vehicle classes in the model, for each of the 364 synthetic households that the transaction model has projected to be likely to make a vehicle purchase decision in year *t*. These probabilities are then used to compute demand for new vehicles, by vehicle type and fuel type, leading to the total light duty vehicle stock forecast for the household sector. For further details on the overall study, see Fowler & Adler (2009).

The survey used for this study involved the design of a highly complex survey tool, which included seven fuel types, fifteen vehicle types, and up to eleven level-of-service attributes, such as cost, fuel consumption, fuel availability, refuelling time and acceleration. To reduce the survey complexity, each choice experiment made use of only four alternatives, where this included a reference vehicle and three other vehicles assigned on the basis of a weighting approach. An Internet-based survey was used to collect the data, enabling the collection of a very large sample.

In addition to assessment of sensitivities to the various attributes of the vehicles, this study makes a contribution to the state-of-practice in modelling the joint choice of vehicle type and fuel type. In particular, it explores the complex correlation patterns that exist between the different options available to a customer. Earlier empirical work revealed the existence of significant levels of correlation between alternatives sharing the same fuel type as well as between alternatives sharing the same vehicle type. Previous studies, however, would at best accommodate one of these dimensions of correlation, potentially leading to biased model results. This study shows how both dimensions can be accommodated jointly in a cross-nested logit (CNL) framework, and why the CNL model is ideally suited for such multi-dimensional choice processes. Finally, the paper includes a brief discussion of an extension of this specification to a model that incorporates random taste heterogeneity across respondents.

The remainder of this paper is organised as follows. We first describe the survey work carried out for this analysis, followed by methodology discussion, and a presentation of the empirical results and a brief forecasting example. Before proceeding to the conclusions section, we also briefly discuss an extension to a model that additionally allows for random taste heterogeneity.

2. Survey Design

The modelling work presented in this report makes use of the data collected in the 2008-09 California Vehicle Survey (CVS). The survey sampling plan was developed with the goal of obtaining a representative sample of California households across a number of geographic regions and household characteristics, such as household size, the number of vehicles, urban/suburban/rural location, number of workers, and annual income. The CVS sample was stratified by five regions in the State of California, which are defined by county as follows:

- San Francisco region: Alameda, Contra Costa, Marin, Napa, San Mateo, Santa Clara, Solano, Sonoma, and San Francisco Counties.
- Los Angeles region: Los Angeles, Orange, Imperial, Riverside, San Bernardino, and Ventura Counties.
- San Diego region: San Diego County.
- Sacramento region: El Dorado, Placer, Sacramento, Sutter, and Yolo, Yuba Counties.
- Rest of State: All other counties.

Households were sampled proportionally across the five regions of the state and household characteristics using US Census data to determine target proportions.

Respondents were invited to participate in the survey through a random-digit-dial (RDD) approach. The telephone-based household sample was purchased from Survey Sampling Inc. (SSI) of Fairfield, Connecticut using landline telephone banks within the five regions. The sample was screened by SSI to remove non-working and business numbers and then matched for residential addresses. In addition, a subsample of 200 recruited cell phone only households has been included since these households are not included in Random-Digit-Dial telephone samples. These households are largely concentrated among those under 35 years old.

We will now briefly discuss the contents of the actual survey, with further details available in Fowler & Adler (2009, Appendix B). Survey data were collected using a two-phase, multi-method approach. The first phase involved a recruitment survey to collect data on revealed preferences (RP) and identify participants planning to purchase a vehicle to recruit for the stated preference survey. The second phase included the stated preference survey with eight vehicle choice scenarios.

In the RP survey, respondents were asked to indicate the type of vehicle they are most likely to purchase next for their household; including information about the vehicle type, fuel type, expected fuel efficiency, purchase price, vehicle age, and estimated number of miles the vehicle would be driven annually.

After completing the revealed preference (RP) survey over landline and mobile phones, the respondents were given the option of completing the stated preference (SP) survey using either print or online questionnaires. In both cases, data from the RP survey was used to construct a set of eight stated preference scenarios for the SP survey, tailored to the specific individual.

A total of 3,274 respondents completed the Stated Preference survey. Each respondent completed 8 choice situations, for a total of 26,192 choice observations. The response rate for the telephone-based RP survey was 22.1% according SRBI. Of the respondents who completed the RP survey, 49.8 percent went on to complete the SP survey. The resulting sample of respondents was broadly representative of the overall California population in terms of geographic origin, household size, household income, and the number of household vehicles.

Each stated preference scenario presented respondents with four hypothetical vehicles as alternatives. The first vehicle, or the reference vehicle, was presented as the new or used vehicle the respondent planned to purchase next for their household. The attributes that describe the reference vehicle were consistent with what the respondent reported in the RP survey in terms of vehicle type, fuel type and age, with the remaining attributes varying across choice sets. The next three alternatives were presented as vehicles of different sizes, fuel types and ages. The four vehicles in each choice scenario were described by a set of ten to twelve attributes, depending on the fuel type presented. Respondents were asked to select the vehicle they would most prefer to purchase based on the attributes presented in each alternative. The values of each attribute varied according to an experimental design (discussed later), requiring respondents to trade off attributes against each other. Figure 1 presents an example of a stated preference scenario.

The first two attributes for each alternative were vehicle type and fuel type. A total of fifteen vehicle types and seven fuel types were selected for the different stated choice scenarios. The vehicle type for the reference vehicle was fixed to the response given in the RP survey. For the remaining three alternatives, vehicle type was drawn from one of the following fifteen types:

1. Subcompact car
2. Compact car
3. Mid-size car
4. Large car
5. Sport car or “two door high performance subcompact car”
6. Small cross-utility car or “small wagons with flexible seating”
7. Small cross-utility SUV
8. Mid-size cross-utility SUV
9. Compact SUV
10. Mid-size SUV
11. Large SUV
12. Compact van
13. Large van
14. Compact pick-up truck
15. Standard pick-up truck

While it was possible for any vehicle to be used for the three alternate vehicles, the vehicles were selected by using weighted draws based on the respondent’s reference vehicle type. Weighted draws were used because it is expected that respondents will have relatively strong preferences for at least a broad category of vehicles (e.g. small or large), and as a result presenting a respondent with a choice between a subcompact car and a large van makes little sense. In that situation, vehicle type would dominate the choice process and little or no information could be gained for the sensitivities to other attributes. On the other hand, completely restricting the different combinations of vehicle types presented to a respondent did not seem appropriate. As a result, a set of weights were developed for each reference vehicle type. With these weights, all vehicle types have a non-zero probability of being included in a given stated choice scenario, but the probability is higher for those vehicles that are more similar to the reference vehicle type. An especially high weight of approximately 50 percent was used for the reference vehicle type, which ensured that, at least for one pair of alternatives, the relative preference was not influenced by vehicle type (i.e. the vehicle type for one of the hypothetical alternatives was the same as that for the reference alternative). The vehicle types for the three alternative vehicles were drawn without replacement from the list of 15 vehicle types, meaning that, while the reference vehicle was allowed to repeat in one other alternative, allowing respondents to trade off attributes other than vehicle type, no other vehicle types were allowed to repeat across alternatives within a single choice scenario.

For the reference vehicle, fuel type was again fixed to the respondent's RP response. The remaining fuel types were drawn from the following list:

1. Standard Gasoline
2. Flex Fuel/E85
3. Clean Diesel
4. Compressed Natural Gas
5. Hybrid-electric
6. Plug-in Hybrid-electric
7. Full Electric

While the vehicle type selection was done using weighted draws, the fuel types for the three alternate vehicles were selected entirely at random, thus guaranteeing that all possible combinations were represented roughly evenly. As with vehicle types, fuel types were drawn without replacement, meaning that while the reference vehicle fuel type was allowed to repeat in one of the three alternate vehicles, allowing respondents to trade off attributes other than fuel type. No other fuel types were allowed to repeat across alternatives within a single choice scenario.

While values for vehicle type and fuel type were selected using weighted and random draws as described above, the values for the remaining attributes varied according to an orthogonal experimental design. The orthogonal design is described in more detail later on. The actual values for these attributes were dependent on the vehicle and fuel type, and for many of the vehicle attributes, the values presented to a respondent varied around a base value. In the case of purchase price, maintenance cost, miles per gallon equivalent, fuel cost per gallon equivalent, and acceleration, a table with base values was used, representing average values for all vehicles of a particular vehicle type, fuel type and vintage.

The full list of these remaining attributes (i.e. on top of vehicle type and fuel type) included in the survey is as follows:

- The **vehicle age** was automatically set to *new* for plug-in hybrid electric and full electric vehicles, with variations around the reference vehicle age for other vehicles, with levels obtained from the experimental design.
- The **purchase price** of the vehicle varied around a base value. For the reference vehicle, the base value was the response given in the RP survey. For the three remaining alternatives, the base value was dependent on a "list price" determined from the combination of vehicle type, fuel type, and vintage (i.e. age[1]), where this was adjusted by the ratio between the indicated price of the reference vehicle in the RP survey and the list price for that vehicle,

[1] We calculated the base price for used vehicles by taking a new vehicle price for that specific vehicle class and fuel type and depreciating it over the age of the vehicle using depreciation rates provided by the Energy Commission.

thus accounting for the possibility that a respondent was considering a higher than average or lower than average price for the reference vehicle. Variations across choice sets were then based on the experimental design.

- There were six **purchase incentive** levels shown in the survey, with the exception of gasoline-powered vehicles, where no incentives were used. Incentives included carpool lane access, free parking, tax credits, reduced tolls and reduced purchase price. The specific level used in a given choice task was obtained from the experimental design.
- A base **maintenance cost** per mile for each vehicle was assumed based on the vehicle type, fuel type, and vehicle age. The maintenance cost per mile was multiplied by the reported annual VMT to calculate an annual maintenance cost. Variations across choice sets were then based on the experimental design.
- A base value for **miles per gallon equivalent** was assumed based on the vehicle type, vehicle age, and fuel type. Variations across choice sets were then based on the experimental design.
- The **annual fuel cost** was calculated using the fuel cost in gasoline gallon equivalents, which was a design attribute, the vehicle efficiency in miles per gallon equivalent, and the annual miles reported in the RP survey. The variation across choices was based on variations in the fuel cost in gasoline gallon equivalents, as specified in the experimental design.
- The **fuel availability**, **refuelling time**, and **vehicle range** attributes only applied to full electric and compressed natural gas vehicles, where variations across choice sets for applicable vehicles were then based on the experimental design.
- The **acceleration attribute** was presented as the time it takes to accelerate from zero to 60 miles per hour in seconds. The acceleration of each vehicle was assumed to vary based on the vehicle type, fuel type and vehicle age, and varied according to the experimental design.

The experimental design used for this SP survey was based on an underlying orthogonal design. While several types of designs were considered at the outset, including arguably more advanced efficient designs, it was concluded that, given the complexity of the SP scenarios, an orthogonal design was the most appropriate for this particular application. While efficient designs can be preferable in some situations, the generation of an efficient design requires prior parameter values for all coefficients, as well as a priori decisions in relation to model structure and utility specification, including interactions with socio-demographic variables. Such information was not available in a reliable form for the present work. Additionally, vehicle type and fuel type would have to be directly included in the design, leading to the requirement of generating a very large number of different designs for different combinations of vehicle types and fuel types. These design considerations were not necessary with the approach used in this study, where vehicle types and fuel types were added to the design in a second stage, after generating the base design.

This base design is an orthogonal design of 144 rows, split into 18 blocks of 8 choices. Orthogonal blocking was used to avoid any correlation between the attributes and the blocks (e.g. avoiding the situation where one respondent gets all the high price options). The design contains the levels for ten attributes (the attributes other than vehicle type and fuel type) and four alternatives. The vehicle types and fuel types drawn according to the approach described above were used as inputs for calculating the base values for the levels in this underlying design. In the actual survey, each respondent was presented with one block of eight choice situations. Care was taken to ensure that the 18 different blocks were presented the same number of times and that there was no correlation between sample subgroups and blocks. The choice situations presented to the respondent were constructed on the basis of the set of vehicle type/fuel type combinations drawn for that respondent, and the block of 8 choice situations used from the experimental design for that respondent. The order in which the 8 choice situations from a given block were presented to a respondent was randomised across respondents. A summary of the design and the levels used is shown in Table 1, while Table 2 shows the ranges for key attributes in the stated choice scenarios.

While the choice scenarios are non-trivial, we feel that they are in line with the real-world complexity of such choices. To test for any impact of complexity on response quality, we ran a separate model allowing for scale differences across the eight choice tasks. There was no evidence of significant differences, indicating no major issues with respondent fatigue.

3. Modelling methodology

The SP data was used to develop discrete choice models belonging to the family of random utility models. The underlying utility function specification was identical for all estimated models, but we explored the use of different model structures. It is also worth noting that the models used in the present paper exclude respondents from multi-vehicle households. While the original survey collected data from one-vehicle as well as multiple-vehicle households, for the present academic study, only the data for single car households was used, yielding a sample of 7,552 observations from 944 individual respondents. However, this has no implications for the principle of allowing for the full correlation structure at play in this multi-dimensional choice process.

3.1. Utility function specification

An extensive specification search was conducted, leading to the inclusion of the following terms in the final specification of the utility function:

- Constants for the first three SP alternatives, along with constants for fuel type and vehicle type inertia
- Vehicle type specific constants, with subcompact as the reference vehicle
- Fuel type specific constants, with standard gasoline as the reference fuel
- A constants for vehicles aged 1 or 2 years, and a constant for vehicles aged 3 years or more, with new vehicles as the reference age
- Four incentive constants, with *no incentive* as the reference
- Marginal utility coefficients associated with vehicle price ($1000s), annual fuel costs ($1000s), and maintenance costs ($1000s)
- A marginal utility coefficient interacting with vehicle price ($1000s) and the income category, where seven equally sized income groups were used[2], and where a linear relationship was justified on the basis of earlier results using income category specific cost coefficients
- Marginal utility coefficients associated with vehicle attributes of miles per gallon equivalent (MPGE), the natural logarithm of range (miles), acceleration (seconds taken to 60mph)
- Constants associated with the option of plugging in electric vehicles at work and at other locations, and the availability of compressed natural gas at 1 out of 20 stations, with the respective references being home plug-in only, and availability at 1 in 50 stations
- Constants associated with interaction terms for large households and medium-sized vehicles, and large households and large-sized vehicles
- Constants associated with interaction terms for alternative fuel vehicles and medium-sized vehicles, and alternative fuel vehicles and large-sized vehicles

This specification led to the use of 44 individual parameters. The justification for including both annual fuel cost and miles per gallon equivalent is that annual fuel cost is a function not just of the miles per gallon equivalent but also of fuel cost per unit and vehicle miles driven. Additionally, annual fuel cost and miles per gallon are in fact the two variables that appear on new car labels and which each potentially influence buyers' choices in different ways. The actual correlation between

[2] Below $20,000; between $20,000 and $40,000; between $40,000 and $60,000; between $60,000 and $80,000; between $80,000 and $100,000; between $100,000 and $120,000; and above $120,000.

the two attributes in the data was only around -0.3, and the maximum correlation between the resulting parameter estimates was below 0.2, thus not causing any concerns about model stability. Note that efforts to include refuelling time in the models were unsuccessful. In the context of a study aimed primarily at exploring the advantages of different nesting structures, no additional socio-demographic interactions (on top of income and household size) were included, in part as the results were aimed for use in a later forecasting model, where the degree of disaggregation is limited.

3.2. Model structure

Although a large number of attributes are used to describe the various alternatives in the SP survey, two of them stand out as main *product* characteristics, namely the vehicle type and the fuel type. Given the nature of the choice scenarios, there are clear grounds to suspect a heightened degree of correlation between two alternatives sharing the same vehicle type or two alternatives sharing the same fuel type. This is even more so the case for two options that are of the same vehicle type and the same fuel type, but vary along some other dimension. To some extent, these correlations can be explained by the inclusion of vehicle type and fuel type constants, but the degree of variation across respondents in their preference for the different vehicle types and fuel types is potentially so high that a large share of the correlation remains unexplained.

If the effects of this *unobserved* correlation are not accounted for, it is likely to lead to unrepresentative substitution patterns. Indeed, assuming that a respondent is interested in purchasing a compact gasoline car and that for some reason, this vehicle becomes unavailable, he or she is arguably more likely to switch to a differently-sized gasoline vehicle (say a sub-compact), than to a vehicle that is of a different fuel type and a different vehicle type. Additionally, there is the possibility that the respondent may more closely evaluate a switch to a differently fuelled compact vehicle (e.g. a hybrid compact) given the similarity in vehicle type. However, the most basic type of discrete choice model, a Multinomial Logit (MNL) model (cf. McFadden, 1974), is based on the assumption of uncorrelated error terms[3]. As a result, this model cannot represent such substitution patterns, and there will be a proportional shift in probability towards all other vehicle type and fuel type combinations. The main issue to be addressed at the modelling stage is thus the representation of this *unobserved* correlation between two cars sharing the same vehicle type and/or fuel type.

The typical approach for dealing with such an issue is estimating a Nested Logit (NL) model (cf. Daly and Zachary, 1978; McFadden, 1978; Williams, 1977). In this model, the error terms still follow an extreme value distribution, as in the simple MNL model, but the error terms of individual alternatives are no longer independently distributed. Any correlation between the error terms (or unobserved utility components) will lead to heightened substitution patterns between these two alternatives. Each alternative belongs to exactly one nest in a NL model, where a nest groups together alternatives that are closer substitutes for one another, and where single alternative nests are used for any alternatives whose error terms are uncorrelated with those of any other alternatives. For each nest containing at least two alternatives, we estimate an additional model parameter λ, where this parameter is constrained between 0 and 1, with 1 reflecting an absence of

[3] With V_i giving the modelled utility of alternative i out of J alternatives, the MNL probability of choosing alternative i is given by $P_i = \frac{e^{V_i}}{\sum_{j=1}^{J} e^{V_j}}$. Here, V_i is a function of the attributes of alternative i and *estimated* parameters which include the various constants and marginal utility coefficients listed above.

correlation, and where the actual level of correlation between the errors is given by 1- λ^2, so that decreasing values of λ lead to increased correlation[4].

To illustrate the applicability of the NL model to the present context, let us assume that we're in a simplified situation where a respondent has six vehicles to choose from:

A. Compact gasoline car
B. Compact hybrid-electric car
C. Compact gasoline car
D. Compact gasoline SUV
E. Compact flex fuel SUV
F. Compact hybrid-electric SUV

In this scenario, vehicles A, B, and C share the same vehicle type, as do vehicles D, E, and F. Vehicles A, C, and D share the same fuel type, as do vehicles B and F. Finally, vehicles A and C share the same vehicle type and the same fuel type. This provides ample source for correlation between alternatives. Various possible NL structures arise, as illustrated in Figure 2. In the first structure (2a), we use nesting by vehicle type, accounting for the correlation between options that share the same type of vehicle. As an example, if vehicle A was to become unavailable (or less attractive say due to a price increase), a respondent previously interested in this vehicle may be more likely to shift his/her interest to vehicles B or C than to vehicles D, E, or F. The second structure (2b) shows the corresponding two-level NL structure using nesting by fuel type. This structure allows for heightened correlation between vehicles A, C, and D, and between vehicles B and F.

It is a likely that both model structures (i.e. nesting by vehicle type or nesting by fuel type) reveal heightened substitution patterns between nested alternatives, i.e. higher correlation between cars sharing the same vehicle type, or cars sharing the same fuel type. This leads to the need for a structure that can jointly accommodate the two types of correlation. Not doing so is likely to lead to biased results in terms of the type of correlation that is accounted for in the models. One possible solution is the use of a three-level NL structure, such as the one shown in the third structure (2c) in Figure 2, first nesting by vehicle type, and then by fuel type. It can immediately be seen that another option, not shown here, is to nest first by fuel type, and then by vehicle type. The model structure in (2c) in Figure 2 allows for correlation between different compact cars of different fuel types, and for correlation between different compact SUVs of different fuel types. Additionally, it allows for even higher correlation between the two gasoline cars, i.e. options A and B. However, given the ordering of the nesting levels, the model is unable to account for the correlation between two options sharing the same fuel type but being of different vehicle type. As an example, we would expect options B and F to be closer substitutes for each other, but the model treats their errors as completely independent. An analogous problem would arise in a model that nests first by fuel type and then by vehicle type. The issue here is that when using a multi-level NL model for multi-dimensional choice processes, the full correlation can only be accommodated along the highest dimension of nesting in the tree, an issue that was to our knowledge first discussed by Hess & Polak (2006) in the context of air travel behaviour.

The solution put forward by Hess & Polak (2006) is to use a cross-nested logit (CNL) structure. In a CNL model, an alternative is allowed to belong to more than one nest, thus allowing for far greater flexibility in the specification of the correlation structure. This will allow for situations in which there

[4] In a two-level NL model with M different nests, where $j \in S_m$ defines the set of alternatives contained in nest m, the probability of choosing alternative i (where i is contained in nest k) is given by $P_i = \frac{e^{\lambda_k I_k}}{\sum_{m=1}^{M} e^{\lambda_m I_m}} \frac{e^{V_i/\lambda_k}}{\sum_{j \in S_k} e^{V_j/\lambda_k}}$, with $I_k = ln \sum_{j \in S_k} e^{V_j/\lambda_k}$.

is both a correlation between alternatives A and B, and between alternatives A and C, but no correlation between alternatives B and C. The CNL model has its origins in the work of McFadden (1988), while the first use of the term cross-nested logit is usually attributed to Vovsha (1997). Various alternative versions of the CNL model have been proposed by Vovsha & Bekhor (1998), Ben-Akiva & Bierlaire (1999) (further expanded by Bierlaire 2006), Papola (2004), and Wen & Koppelman (2001). The differences between the models arise primarily from the specification of the allocation parameters and the conditions associated with these parameters. The role of the allocation parameters is to explain the membership of an alternative in the different nests of the model. These parameters are required since the model no longer operates under the strict single nest membership condition of the simple NL model[5].

In the present application, using a CNL structure allows the model to jointly accommodate the correlation between alternatives sharing the same vehicle type and the correlation between alternatives sharing the same fuel type. Unlike a three-level NL structure, we do not need to impose an *ordering* condition and consequently can account for the full extent of correlation along both dimensions. Figure 3 shows the application of the CNL model structure in the context of the example used in this study. The model structure uses two separate vehicle type nests and three separate fuel type nests, with each alternative falling into one vehicle type nest and one fuel type nest. In the resulting structure, we have correlation between those alternatives sharing the same vehicle type (i.e. A, B, C, and D, E, F), and between the vehicles sharing the same fuel type (i.e. A, C, D, and B, F), with even higher correlation for those alternatives sharing the same vehicle type as well as the same fuel type (i.e. A and C).

For the joint model of vehicle type and fuel type, the CNL structure allows an alternative to belong to one vehicle type nest and one fuel type nest. For the full set of alternatives used in this application, we thus make use of 15 different vehicle type nests, and 7 different fuel type nests. Clearly, the allocation parameter for a given alternative will take a non-zero value only for one of these vehicle type nests and one of the fuel type nests. In other words, for a subcompact gasoline car, the vehicle type nest allocation parameters will be zero for all vehicle type nests other than the subcompact car nest, and all fuel type nests other than the gasoline car nest. With 105 different vehicle type and fuel type combinations, the model would thus have 210 different allocation parameters, each time with a summation to one constraint between the two non-zero allocation parameters relating to a given alternative. For the final CNL model used in this work, the two non-zero allocation parameters for a given alternative were fixed to a value of 1/2, meaning that an alternative belongs by the same *proportion* to one vehicle type nest and one fuel type nest. While this may be seen as an inferior model specification, an additional model run in which the allocation parameters were freely estimated only led to an improvement in log-likelihood by 30 units, despite the huge increase in the number of parameters (and the associated rise in estimation time and identification issues), and only a small number of parameters were significantly different from ½, while large standard errors were observed for others. We feel that this finding is possibly a reflection that, in the present context, the complexity of the correlation structure is adequately captured by the estimation of nest-specific structural parameters.

---

[5] In the present paper, the general specification also given in Train (2003) is used. Again using different nests, with α_{jm} describing the allocation of alternative j to nest m, we have that

$$P_i = \sum_{m=1}^{M} \left(\frac{\left(\sum_{j \in S_m} \left(\alpha_{jm} e^{Vj} \right)^{1/\lambda_m} \right)^{\lambda_m}}{\sum_{l=1}^{M} \left(\sum_{j \in S_l} \left(\alpha_{jl} e^{Vj} \right)^{1/\lambda_l} \right)^{\lambda_l}} \frac{\left(\alpha_{im} e^{Vi} \right)^{1/\lambda_m}}{\sum_{j=1}^{J} \left(\alpha_{jm} e^{Vj} \right)^{1/\lambda_m}} \right)$$

. Here, the extra summation in comparison with the NL formula ensures that each alternative can potentially belong to each nest. In the present specification, we have two conditions for the allocation parameters, namely $0 \leq \alpha_{jm} \leq 1, \forall j, m$, and $\sum_{m=1}^{M} \alpha_{jm} = 1,\ \forall j$.

As a final point in this discussion, we briefly touch on an alternative way of capturing the correlation structure in the dataset. Generalised extreme value (GEV) models such as NL and CNL are highly flexible structures that additionally have the advantage of a closed form function for the choice probabilities, significantly reducing estimation cost. In recent years however, there has been an increasing focus on the use of mixed multinomial logit (MMNL) models. While the majority of MMNL applications have exploited the structure with a view to accommodating random taste heterogeneity, it is important to recognise that the MMNL model can also be used to allow for correlation between alternatives, with the use of a so called error components logit (ECL) specification (cf. Walker, 2001). In fact, McFadden & Train (2000) discuss how the MMNL model can approximate any other random utility model arbitrarily closely. However, in the present context, we do not see any advantage in replacing our closed form GEV structures by a ECL specification, which, with 15 different vehicle types and 7 different fuel types, would entail the estimation of a model with a very high dimensional integral, leading to impractically high estimation costs, in addition to the need to address the difficult question of model identification (cf. Bowman, 2004; Walker, 2002; Walker et al., 2003). MMNL models obviously have the additional advantage of allowing for random taste heterogeneity alongside any correlation accommodated through error components. However, such taste heterogeneity can also be accommodated in a Mixed GEV context, where a GEV kernel is used to accommodate the correlation structure, with random terms being used only for the unexplained taste heterogeneity (cf. Hess et al., 2005a). This point will be revisited in Section 6.

3.3. Specification of choice set and model estimation

With the above approach of characterising alternatives along two dimensions, we obtain 105 combinations of vehicle types and fuel types. However, with the survey making use of four separate SP alternatives, and each alternative potentially taking on one of those 105 combinations, the model implementation actually made use of four sets of 105 utility functions, i.e. a total of 420 alternatives, of which exactly four were available in each choice situation. Given that only four alternatives are presented in any given choice task, the question arises as to how we can capture the correlation between alternatives sharing the same vehicle type or the same fuel type if these are not routinely presented jointly. The key to understanding this comes in the structure of the NL and CNL models; these models allow for an unobserved component in the utility function that is shared across such alternatives. As long as sufficient cases arise in which they are presented jointly so as to allow for identification, there is no need for this joint presentation to be universal. The number of cases with equal vehicle types or equal fuel types was sufficiently high by design (for each vehicle type and fuel type) so as to allow for these additional parameters to be identified.

All model estimation and forecasting work reported in this paper were carried out using BIOGEME (Bierlaire, 2005), which is easily capable of dealing with such a large CNL structure. To make the results consistent with the above reported specifications, the λ_m parameters reported in the results are obtained as the inverse of the structural parameters μ_m reported by Biogeme, with an appropriate transformation of the standard errors (cf. Hess & Daly, 2009). One further point needs addressing. The data used in this survey contains multiple observations for each respondent, potentially leading to correlations amongst choices for the same respondent. Not accounting for this would lead to a downwards bias in the standard errors for the parameter estimates, and hence overstated measures of confidence. In the present work, we address this issue by using the panel specification of the robust covariance matrix, a measure shown to offer a reliable a stable correction of the standard errors (cf. Daly & Hess, 2010).

4. Empirical results

The estimation results for the different discrete choice models are summarised in Table 3 for the parameters of the utility function and Table 4 for the nesting parameters for the NL and CNL models.

Our first observation is that the NL model using nesting by fuel type gives us an improvement in model fit over the MNL model by 12.43 units in log-likelihood (LL), which is statistically significant, coming at the cost of just six additional parameters. These additional parameters are the nesting parameters where the model would collapse back to a MNL structure if all nesting parameters took on a value of 1. Similarly, the NL model using nesting by vehicle type improves the LL (compared with MNL) by 18.74 units, at the cost of 11 additional parameters. This is again statistically significant, as is the 34.93 unit improvement for the CNL model (compared with MNL), at the cost of 17 additional parameters. Finally, a likelihood ratio test cannot be used in this case to compare the CNL model to the NL given the extra constraint on the allocation parameters in the CNL model, but the adjusted ρ^2 measure shows a small additional improvement. While the improvements in model fit may be regarded as modest (albeit highly significant), this is not uncommon, and significant differences exist between the models in the implied behavioural patterns, as highlighted in Section 5.

Actual estimation results in Table 3, with a few exceptions, show very similar parameter estimates across the four models, with the real differences between the models becoming apparent later on in the forecasting exercise. Going through the various estimates in turn, the values for the three constants suggest some allegiance to the reference alternative, along with a small amount of reading left to right impact (not statistically significant for the third alternative). Further evidence of inertia is given by the two following estimates, showing that respondents are highly likely to choose a vehicle of their initially intended vehicle type and to a slightly lesser extent the same fuel type. Without attempting to read too much into the various vehicle type and fuel type constants, the large negative values for the CNG and full electric vehicles do stand out, suggesting that additional incentives/improvements are required to increase the attractiveness of such vehicle given the low baseline preference. There is clear evidence of decreasing attractiveness with increasing age, while the various incentives have a positive impact on utility, albeit with low overall levels of statistical significance. All different cost components lead to reductions in utility, though the vehicle price sensitivity is reduced as income increases. Better acceleration, longer range (low statistical significance), better fuel efficiency and improved fuel availability (low significance for EV) all have positive impacts on utility, while large households show the expected preference for larger vehicles, and the attractiveness of alternative fuel vehicles reduces with vehicle size, although this effect only attains low levels of statistical significance.

We next turn our attention to the nesting parameters, with estimates shown in Table 4. The base value of 1 equates to an absence of correlation (as in a MNL model) and for this reason, the t-ratios are calculated with respect to a base value of 1 rather than 0. Looking first at the model using nesting by fuel type, we observe that the nesting parameter for full electric vehicles has collapsed to a value of 1, indicating no heightened correlation between different full electric vehicles. In addition, the values for the nesting parameters for three other fuel types, namely Flex Fuel/E85, Clean Diesel, and Compressed Natural Gas, are close to 1 and not significantly different from 1, suggesting that only low levels of correlation arise in these contexts. However, high correlation is observed between different gasoline cars and different hybrid-electric cars, and to a lesser extent plug-in hybrid-electric cars, suggesting in each case the presence of heightened substitution patterns and greater fuel type *allegiance*. In the model using nesting by vehicle type, a number of nesting parameters once again collapse to a value of 1, while high correlation is, for example, observed in the Small cross-utility SUV nest and the Compact pick-up truck nest. Overall, the picture in the CNL model is the combination of the NL results, with the exception that we now observe high correlation in the Flex Fuel/E85 nest, and that the nesting parameter for the Compressed Natural Gas is now even closer to 1 than in the NL model.

As an additional step in the analysis of our results, Table 5 presents implied monetary valuations for a number of key measures, where we show the valuations separately for the three models, as well as for the lowest income group, the middle (and most common) income group, and the highest income group. Using the valuations from the preferred model (CNL), we observe a high willingness to pay for new cars relative to cars aged between 1 and 2 years, and especially cars aged 3 years or more. Working with the middle income group, we see that the valuation of HOV lane use is higher than that for free parking, possibly due to high congestion along with good general availability of parking spaces. Interestingly, and rather irrationally, a reduction in purchase price by $1,000 is valued higher when it appears as an incentive than when it appears as a reduction in the base price. This could possibly be explained on the grounds of respondents *desiring* a more *expensive* car given the added status. The asymmetry applies even more strongly when looking at the value of a tax credit. The sensitivity to fuel cost is much stronger than the sensitivity to vehicle price, while a smaller asymmetry also applies for maintenance cost. In each case, these asymmetries can be explained on the basis of a difference between a regular cost (fuel and maintenance) and a one-off cost (purchase price). We also see high willingness to pay measures for increases in MPGE and range, as well as for increased vehicle performance. Finally, the valuations of greater availability of fuelling options for electric and CNG vehicles show the prevailing concerns about limited current availability. In the CNL model, we see that the valuations in the middle income group are almost sixty percent higher than those in the low income group, while they are just under sixty percent of the valuations in the high income group (a result of the linearity assumption). We also see differences between the three model structures, with variations by up to a third, but on average the values are relatively stable.

5. Forecasting example

A brief forecasting exercise produces a final illustration of the differences between the various models estimated in this paper. Clearly, rescaling of the model outputs and correction of the constants would be required before undertaking any forecasting for the purposes of guiding policy makers (cf. Louviere et al., 2000), but the aim of this example is purely illustrative.

In this forecasting exercise, all 105 combinations of fuel type and vehicle type are available to a single respondent. We further assume that this respondent currently owns a subcompact gasoline vehicle, has an annual mileage of 13,500, and comes from a household falling into the average income category and having four or fewer members. Finally, we assume that this respondent is solely interested in new cars. Our forecasting exercise starts by working out the probabilities for the 105 combinations of vehicle type and fuel type, with the four different models estimated in this paper. Next, we assume that following a government policy intervention, there is a reduction in the cost of Plug-in Hybrid Electric vehicles (fuel type 6) by $4,000, on the condition that they also fall into the Subcompact car, Compact car or Mid-size car categories (vehicle types 1, 2, and 3).

Table 6 presents the changes in probabilities that arise as a result of this change in the attribute for these three vehicles. With $P_{vf,base}$ giving the base probability for a given vehicle type and fuel type combination, and $P_{vf,new}$ giving the corresponding probability after the policy intervention, the value shown in the table is given by $(P_{vf,new}- P_{vf,base})/P_{vf,base}$, i.e. the relative change in probability. Given that we are working with relative rather than absolute changes, the values in the table clearly do not sum to zero.

In the MNL model, we observe an equal increase in the probabilities for the three concerned vehicles, where this is drawn proportionally from all remaining vehicles as a result of the independently distributed errors and the resulting absence of unmodelled correlation in this model. In the model nested by fuel type, we observe a bigger increase in the probability for the three vehicles than in the MNL model, but, with more of the increase in probability drawn from the remaining Plug-in Hybrid Electric vehicles than other fuel types. As a result, the overall increase in the probability for Plug-in Hybrid Electric options, across vehicle type, is far less marked. Both of these effects are consistent with intuition. However, the changes in this model draw proportionally from all other non Plug-in Hybrid Electric Vehicles, independently of the vehicle type, which may again not be completely realistic.

In the model nested by vehicle type, we observe a bigger draw away from those options with the same vehicle type but with different fuel types (i.e. all non "Plug-in Hybrid Electric" subcompact, compact and mid-size cars). However, we also observe that the actual changes are different across the first three vehicle types. This results from different levels of correlation in these nests, with higher correlation leading to higher competition. With the nesting parameter (cf. Table 4) being lowest (and hence the correlation being highest) in the mid-size car nest, followed by the compact car and the subcompact car nests, the changes are also largest in the mid-size car nest, followed by the compact car and the subcompact car nests. This model correctly recognises that a larger share of the draw will come from other options in the same vehicle type categories, such that the increase in the overall share for the three vehicle types is less marked than in the first two models. However, the model doesn't recognise the fact that a bigger draw will also come from other Plug-in Hybrid Electric vehicles. Indeed, for a given vehicle type, all the different fuel types are affected in the same way, including in the case of Plug-in Hybrid Electric. This overstates the overall increase for the probability of that fuel type, especially when compared to the model using nesting by fuel type. Finally, while due to aggregation, the overall changes (i.e. when grossing up across categories) are similar in the CNL and MNL models, the model arguably gives a more intuitively meaningful representation of the changes in probabilities of individual vehicle type and fuel type combinations, with bigger reductions in probabilities for those alternatives that either are of the "Plug-in Hybrid Electric" fuel type or are of the "Subcompact car", "Compact car" or "Mid-size car" vehicle type. In other words, a policy aimed at increasing uptake in small plug-in hybrid electric vehicles will lead to a reduction in the share for all remaining types of car, but especially so for types of plug-in hybrid electric vehicles and other small cars using different types of fuel.

6. Extensions to GEV mixture models

As an additional extension, we estimated four mixture equivalents of the models reported in Table 3, namely a Mixed MNL (i.e. MMNL) model, a Mixed NL model using nesting by fuel type, a Mixed NL model using nesting by vehicle type, and a Mixed CNL model. With the main focus in the present paper being on the representation of the inter-alternative correlation, only a simple specification of the MMNL model was used, with no regards to the issue of distributional assumptions (cf. Hess et al., 2005b; Daly et al., 2010). In particular, we relied on univariate Normal distributions and allowed for random heterogeneity in six marginal utility coefficients, namely the three cost components (fuel cost, maintenance cost, and vehicle price), and three performance indicators (acceleration, range, and miles per gallon equivalent).

The results are summarised in Table 7, where, given the focus of the present paper, only a brief overview is given. These models took significantly longer to estimate, with the Mixed CNL model taking over a week to converge; as such, these highly complex models are even more difficult to use in practice, but provide further insights into behaviour by allowing for random taste heterogeneity on top of the complex correlation structure. The results show that all four models lead to very substantial increases in model fit compared to their fixed coefficient counterparts, as a result of allowing for the random taste heterogeneity. The actual performance of the different models is very similar, but the implied behaviour is very different across models. Firstly, there are differences in the retrieved degree of random taste heterogeneity (expressed in Table 7 in the form of the coefficient of variation). Here, we observe, overall, a drop in the degree of heterogeneity once we accommodate the correlation structure, which would highlight confounding between these two phenomena in the simple Mixed MNL model, in line with the discussions of Hess et al. (2005a). Similarly, there could also be the expectation that the degree of inter-alternative correlation is overstated in the simple GEV models without additional taste heterogeneity, again due to confounding between correlation and taste heterogeneity. With the exception of the nesting parameter for hybrid electric vehicles, this is indeed the case in the models nesting alternatives by fuel type. On the other hand, in the model using nesting by vehicle type, six of the eleven estimated nesting parameters show decreases (i.e. higher correlation), with the remaining five showing increases, where the increases in nesting parameters (and hence decreases in correlation) are typically smaller than the decreases (and hence increases in correlation). In the CNL case, two additional nesting parameters collapse to a value of one when adding in additional taste heterogeneity. Of the remaining fifteen parameters, six show increases (i.e. lower correlation), while seven show decreases (i.e. higher correlation), with two remaining roughly the same. These findings show that allowing for correlation between alternatives in a multi-dimensional choice process remains important even when incorporating additional random taste heterogeneity. They also highlight the possibility of confounding between random taste heterogeneity and unobserved correlation, but show that the direction of this confounding effect is not as clear from the outset as observed in some other studies, e.g. the work of Hess et al. (2005a). As mentioned at the outset, this extension to a Mixed GEV case was primarily illustrative, and further investigation of the use of such structures remains an important avenue for future research.

7. Conclusions

In this paper, we have presented the findings from a modelling analysis of vehicle type and fuel type choices of California consumers, captured through a stated preference survey. The modelling results have revealed the expected negative sensitivities to factors such as vehicle price, operating cost and age, along with positive effects for improved accelerating, range, purchase incentives and fuel availability. In the context of the growing interest in alternative fuel vehicles, the analysis has shown that while, all else being equal, respondents have a higher baseline preference for clean diesel, flex fuel, and hybrid electric vehicles in comparison with standard gasoline vehicles, this is not the case for compressed natural gas and full electric vehicles, where further incentives are required.

From a modelling perspective, the paper has focussed on the representation of the complex correlation structure that arises in the context of this two-dimensional choice process, i.e. the fact that we would expect heightened substitution between two cars that are either the same types of vehicle or use the same type of fuel. Here, our theoretical discussions have shown how the standard approach for dealing with such correlation, namely a Nested Logit model, may not be appropriate, as it is unable to capture the full extent of correlation along both the vehicle type and the fuel type dimensions. We have shown how the more flexible Cross Nested Logit model is able to accommodate the full pattern of correlation, leading to better performance in estimation and more realistic substitution patterns when used in forecasting. As a brief extension we have also illustrated the benefits of additional allowing for random taste heterogeneity within a mixed GEV context, where more in-depth use of such models and study of the confounding between unexplained taste heterogeneity and inter-alternative correlation remains an important avenue for future research.

Although advanced nesting structures have been used in applied studies in the past, going back, for example, to the use of the CNL model for forecasting mode choice in the Tel Aviva by Vovsha (1997), many large scale forecasting systems continue to rely on more basic MNL structures, albeit with attempts to incorporate as much of the inter-alternative correlation as possible in the modelled utility components. The evidence presented here suggests that the use of more advanced nesting structures remains an important avenue for future applied work.

In closing, it should also be mentioned that, in the present paper, we have solely investigated the correlation between alternatives sharing the same vehicle type and the same fuel type. Clearly, there may also be correlations between vehicles of different but similar type, such as compact and sub-compact cars, and incorporating such cross-type correlation is an important area for future work.

Acknowledgements

This paper uses data collected for a project commissioned by the California Energy Commission. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the California Energy Commission. The first author acknowledges the financial support of the Leverhulme Trust in the form of a *Leverhulme Early Career* Fellowship.

References

Adler, T., Wargelin, L., Kostyniuk, L., Kavalec, C. and Occhiuzzo, G. (2004), Experimental Assessment of Incentives for Alternate Fuel Vehicles, presented at the Transportation Research Board Annual Meeting, Washington, D.C. January 2004.

Batley, R. and Toner, J. (2003) Elimination-by-aspects and advanced logit models of stated preferences for alternative-fuel vehicles. Proceedings of the European Transport Conference, Strasbourg, October 2003.

Batley, R., Toner, J. and Knight, M. (2004) A mixed logit model of UK household demand for alternative-fuel vehicles. International Journal of Transport Economics, 31 (1), pp55-77.

Ben-Akiva, M. & Bierlaire, M. (1999), Discrete choice methods and their applications to short term travel decisions, in R. Hall, ed., 'Handbook of Transportation Science', Kluwer Academic Publishers, Dordrecht, The Netherlands, chapter 2, pp. 5–34.

Bierlaire, M. (2006). A theoretical analysis of the cross-nested logit model. Annals of Operations Research, Vol. 144, 287-300.

Bierlaire, M., 2005. An introduction to BIOGEME Version 1.4. biogeme.ep.ch.

Bowman, J. L. (2004), Logit kernel (or mixed logit) models for large multidimensional choice problems: identification and estimation, paper presented at the 83rd Annual Meeting of the Transportation Research Board, Washington, DC.

Bunch, D.S., Bradley, M., Golob, T.F., Kitamura, R. and Occhiuzzo, G.P. (1993), Demand for Clean Fueled Vehicles in California: A Discrete Choice, Stated Preference Survey, Transportation Research A, Vol 27A, pp237-253.

Daly, A. and S. Zachary (1978), Improved multiple choice models, in D.Hensher and M.Dalvi, eds, 'Determinants of Travel Choice', Saxon House, Sussex.

Daly, A.J. & Hess, S. (2010), Simple Approaches for Random Utility Modelling with Panel Data, paper presented at the European Transport Conference, Glasgow, October 2010.

Daly, A., Hess, S. & Train, K. (2009), Assuring Finite Moments For Willingness To Pay In Random Coefficient Models, paper presented at the European Transport Conference, Noordwijkerhout, October 2009.

Erath, A. & Axhausen, K.W. (2010), Long term fuel price elasticity: effects on mobility tool ownership and residential location choice, final report to BFE and BAFU, IVT, ETH Zurich, Zurich.

Fowler, M. & Adler, T. (2009), Transportation Fuel Demand Forecast Household and Commercial Fleet Survey Task 8 Report: Logistic Regression Analysis and Results, report for the California Energy Commission.

Golob, T., Brownstone, D., Bunch, D. and Kitamura, R. (1995), Forecasting Electric Vehicle Ownership and Use in the California South Coast Air Basin, Report to the Southern California Edison Company.

Greene, D. L. (2001), TAFV Alternative Fuels and Vehicles Choice Model Documentation, Oak Ridge National Laboratory.

Hess, S., Bierlaire, M. & Polak, J. W. (2005a), Capturing taste heterogeneity and correlation structure with Mixed GEV models, in R. Scarpa & A. Alberini, eds, 'Applications of Simulation Methods in Environmental and Resource Economics', Springer Publisher, Dordrecht, The Netherlands, chapter 4, pp. 55–76.

Hess, S., Bierlaire, M. & Polak, J.W. (2005b), Estimation of value of travel-time savings using Mixed Logit models, *Transportation Research Part A*, 39(2-3), pp. 221-236.

Hess, S. & Polak, J.W. (2006), Exploring the potential for cross-nesting structures in airport-choice analysis: a case-study of the Greater London area, Transportation Research Part E, 42, pp. 63-81.

Hess, S. & Daly, A. (2009), Calculating errors for measures derived from choice modelling estimates, paper presented at the 88th Annual Meeting of the Transportation Research Board, Washington, D.C.

Kavalec, C. (1996), CALCARS: The California Conventional and Alternative Fuel Response Simulator, A Nested Multinomial Vehicle Choice and Demand Model, Sacramento, CA: California Energy Commission.

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

McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior, in P.Zarembka, ed., 'Frontiers in Econometrics', Academic Press, New York, pp. 105–42.

McFadden, D. (1978), Modeling the choice of residential location, in A.Karlqvist, L.Lundqvist, F.Snickars and J.Weibull, eds, 'Spatial Interaction Theory and Planning Models', North-Holland, Amsterdam, pp. 75–96.

McFadden, D. & Train, K. (2000), 'Mixed MNL Models for discrete response', Journal of Applied Econometrics 15, 447–470.

Ortúzar, J de D., Roncagliolo, D.A., and Velarde, U.C., (1997), Interactions and Independence in Stated Preference Modelling, PTRC Perspectives, London.

Papola, A. (2004). Some developments on the Cross Nested Logit model. Transportation Research Part B, 38, 833-851.

Spissu, E., Pinjari, A., Pendyala, R. and Bhat, C.R. (2009), A Copula-Based Joint Multinomial Discrete-Continuous Model of Vehicle Type Choice and Miles of Travel, forthcoming TRR, Washington DC.

Tompkins, M., Bunch, D., Santini, D., Bradley, M., Vyas, A., and Poyer, D. (1998), Determinants of Alternative Fuel Vehicle Choice in the Continental United States, TRR 1641.

Train, K. (1983), California Personal Vehicle Energy Demand Model, Sacramento: California Energy Commission

Train, K.E. (2003), Discrete choice methods with simulation, Cambridge University Press, Cambridge, MA.

Vovsha, P. (1997), ‘Application of a Cross-Nested Logit model to mode choice in Tel Aviv, Israel, Metropolitan Area’, Transportation Research Record 1607, 6–15.

Vovsha, P. & Bekhor, S. (1998), ‘The Link-Nested Logit model of route choice: overcoming the route overlapping problem’, Transportation Research Record 1645, 133–142.

Walker, J. (2001), Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables, PhD thesis, MIT, Cambridge, MA.

Walker, J. (2002), ‘Mixed logit (or logit kernel) model: Dispelling misconceptions of identification’, Transportation Research Record 1805, 86–98.

Walker, J., Ben-Akiva, M. & Bolduc, D. (2003), Identification of the Logit Kernel (or Mixed Logit) Model, Working Paper, Department of Civil Engineering, MIT, Cambridge, MA.

Wen, C.-H. & Koppelman, F. S. (2001), ‘The Generalized Nested Logit Model’, Transportation Research Part B: Methodological 35(7), 627–641.

Williams, H. (1977), ‘On the formation of travel demand models and economic evaluation measures of user benefits’, Environment and Planning A 9, 285–344.

Table 1: Design of SP experiment

Attribute	Notes	Vehicle A Reference Vehicle		Vehicle B SP Alternative 1		Vehicle C SP Alternative 2		Vehicle D SP Alternative 3	
Vehicle Type	*Reference vehicle fixed to RP vehicle type. Other vehicle types weighted based on reference vehicle*	RP Vehicle Type [fixed]		*RP is only repeated in 1 alternative* Subcompact car Compact car Mid-size car Large car Sport car Small cross-utility car Small cross-utility SUV Mid-size cross-utility SUV Compact SUV Mid-size SUV Large SUV Compact van Large van Compact pick-up Standard pick-up		*RP is only repeated in 1 alternative* Subcompact car Compact car Mid-size car Large car Sport car Small cross-utility car Small cross-utility SUV Mid-size cross-utility SUV Compact SUV Mid-size SUV Large SUV Compact van Large van Compact pick-up Standard pick-up		*RP is only repeated in 1 alternative* Subcompact car Compact car Mid-size car Large car Sport car Small cross-utility car Small cross-utility SUV Mid-size cross-utility SUV Compact SUV Mid-size SUV Large SUV Compact van Large van Compact pick-up Standard pick-up	
Fuel type	*Vehicle A fixed to RP fuel type. RP may be repeated once in vehicles B, C, or D*	RP Fuel Type [fixed]		*RP is only repeated in 1 alternative* Gasoline E85 Diesel CNG HEV PHEV Full EV		*RP is only repeated in 1 alternative* Gasoline E85 Diesel CNG HEV PHEV Full EV		*RP is only repeated in 1 alternative* Gasoline E85 Diesel CNG HEV PHEV Full EV	
Age of Vehicle	*Reference vehicle fixed to RP age*	RP Vehicle Age [fixed]		*PHEV or Full EV* New [fixed]	*All other fuel types* RP age - 3 RP age RP age + 3 *(max 2009)*	*PHEV or Full EV* New [fixed]	*All other fuel types* RP age - 3 RP age RP age + 3 *(max 2009)*	*PHEV or Full EV* New [fixed]	*All other fuel types* RP age - 3 RP age RP age + 3 *(max 2009)*
Base Price (MSRP)	*Base price dependent on vehicle type, age, and fuel type premium*	RP Vehicle Price - 20% RP Vehicle Price - 7% RP Vehicle Price + 7% RP Vehicle Price + 20%		Base price - 20% Base price - 7% Base price + 7% Base price + 20%		Base price - 20% Base price - 7% Base price + 7% Base price + 20%		Base price - 20% Base price - 7% Base price + 7% Base price + 20%	
Purchase Incentive	*Gasoline always sees "none"*	*Gasoline* None [fixed]	*All other fuel types* None HOV access Free parking $1,000 tax credit 50% reduced toll $1000 reduced purch. price	*Gasoline* None [fixed]	*All other fuel types* None HOV access Free parking $1,000 tax credit 50% reduced toll $1000 reduced purch. price	*Gasoline* None [fixed]	*All other fuel types* None HOV access Free parking $1,000 tax credit 50% reduced toll $1000 reduced purch. price	*Gasoline* None [fixed]	*All other fuel types* None HOV access Free parking $1,000 tax credit 50% reduced toll $1000 reduced purch. price
Fuel Cost per Gallon Equivalent *(Not shown)*	*Base cost dependent on fuel type*	Base cost - 15% Base cost - 5% Base cost + 5% Base cost + 15%		Base cost + 15% Base cost + 5% Base cost - 5% Base cost - 15%		Base cost + 15% Base cost + 5% Base cost - 5% Base cost - 15%		Base cost + 15% Base cost + 5% Base cost - 5% Base cost - 15%	
MPG Equivalent	*Base MPGE dependent on vehicle type and fuel type*	Base MPGE + 15% Base MPGE + 5% Base MPGE - 5% Base MPGE - 15%		Base MPGE - 15% Base MPGE - 5% Base MPGE + 5% Base MPGE + 15%		Base MPGE - 15% Base MPGE - 5% Base MPGE + 5% Base MPGE + 15%		Base MPGE - 15% Base MPGE - 5% Base MPGE + 5% Base MPGE + 15%	
Annual Fuel Costs		(Fuel cost per gallon) x (RP VMT) / (MGPE)		(Fuel cost per gallon) x (RP VMT) / (MGPE)		(Fuel cost per gallon) x (RP VMT) / (MGPE)		(Fuel cost per gallon) x (RP VMT) / (MGPE)	
Fuel Availability *(if fuel type is full EV or CNG)*	*Only shown if fuel type is full EV or CNG*	*Full EV* Plug-in only at home Plug-in at work and other loc.	*CNG* 1 in 50 stations 1 in 20 stations	*Full EV* Plug-in only at home Plug-in at work and other loc.	*CNG* 1 in 50 stations 1 in 20 stations	*Full EV* Plug-in only at home Plug-in at work and other loc.	*CNG* 1 in 50 stations 1 in 20 stations	*Full EV* Plug-in only at home Plug-in at work and other loc.	*CNG* 1 in 50 stations 1 in 20 stations
Refueling Time *(if fuel type is full EV or CNG)*	*Only shown if fuel type is full EV or CNG*	*Full EV* 8 Hrs 3 Hrs	*CNG* 10 min (station), 4 hrs (home) 10 min (station), 8 hrs (home)	*Full EV* 8 Hrs 3 Hrs	*CNG* 10 min (station), 4 hrs (home) 10 min (station), 8 hrs (home)	*Full EV* 8 Hrs 3 Hrs	*CNG* 10 min (station), 4 hrs (home) 10 min (station), 8 hrs (home)	*Full EV* 8 Hrs 3 Hrs	*CNG* 10 min (station), 4 hrs (home) 10 min (station), 8 hrs (home)
Range *(if fuel type is full EV or CNG)*	*Only shown if fuel type is full EV or CNG*	*Full EV* 30 miles 40 miles 50 miles 60 miles	*CNG* 150 miles 200 miles 250 miles 300 miles	*Full EV* 30 miles 40 miles 50 miles 60 miles	*CNG* 150 miles 200 miles 250 miles 300 miles	*Full EV* 30 miles 40 miles 50 miles 60 miles	*CNG* 150 miles 200 miles 250 miles 300 miles	*Full EV* 30 miles 40 miles 50 miles 60 miles	*CNG* 150 miles 200 miles 250 miles 300 miles
Maintenance Costs	*Base cost dependent on vehicle type and age*	Base cost - 25% Base cost - 10% Base cost + 10% Base cost + 25%		Base cost - 25% Base cost - 10% Base cost + 10% Base cost + 25%		Base cost - 25% Base cost - 10% Base cost + 10% Base cost + 25%		Base cost - 25% Base cost - 10% Base cost + 10% Base cost + 25%	
Acceleration	*Acceleration dependent on vehicle type and age*	Base acceleration - 2 Base acceleration + 2		Base acceleration - 2 Base acceleration + 2		Base acceleration - 2 Base acceleration + 2		Base acceleration - 2 Base acceleration + 2	

Table 2: Ranges of key attributes in stated choice scenarios

	Age	Price ($)	Fuel cost per gallon ($)	Miles per gallon equivalent (MPGE)	Annual fuel cost ($)	Annual maintenance cost ($)	Time to 60mph (s)
Minimum	0	3,000	1.92	11.00	30.00	20	3.92
Maximum	21	104,600	4.85	135.00	14680.00	3540	14.47
Average	1.24	27,109	3.12	35.24	1478.23	541.50	8.94
Median	0	24,700	3.07	26.00	1140.00	440	8.77

Table 3 Estimation results for different discrete choice models (part 1): main utility parameters

		MNL		NL fuel		NL vehicle		CNL	
	LL	-7681.257		-7668.826		-7662.514		-7646.331	
	par	44		50		55		61	
	adj. rho^2	0.262		0.263		0.263		0.264	
		est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
Constants	Vehicle A constant	0.849	12.53	0.51	6.5	0.664	9.19	0.409	5.03
	Vehicle B constant	0.157	3.26	0.154	3.39	0.147	3.38	0.143	3.41
	Vehicle C constant	0.0403	0.87	0.0451	1.04	0.0319	0.77	0.0361	0.84
	Vehicle Type Inertia	0.988	16.78	0.933	15.8	1.18	16.47	1.09	15.6
	Fuel Type Inertia	0.209	2.26	0.585	5.11	0.177	2.23	0.465	3.92
Vehicle type	Subcompact car	0	-	0	-	0	-	0	-
	Compact car	-0.143	-1.15	-0.144	-1.21	-0.109	-0.87	-0.111	-1
	Mid-size car	0.246	1.86	0.23	1.84	0.247	1.86	0.238	2.1
	Large car	-0.128	-0.68	-0.097	-0.56	-0.213	-1.15	-0.162	-0.9
	Sport car	0.0289	0.19	0.0632	0.43	-0.0243	-0.16	0.00135	0.01
	Small cross-utility car	0.424	3.3	0.393	3.24	0.41	3.21	0.396	3.36
	Small cross-utility SUV	0.29	1.73	0.302	1.96	0.258	1.6	0.274	1.71
	Mid-size cross-utility SUV	0.141	0.73	0.153	0.85	0.0793	0.43	0.0988	0.55
	Compact SUV	0.433	2.48	0.422	2.6	0.369	2.19	0.389	2.28
	Mid-size SUV	0.319	1.65	0.306	1.7	0.241	1.29	0.262	1.72
	Large SUV	0.443	1.86	0.424	1.93	0.343	1.39	0.363	1.71
	Compact van	0.0405	0.19	0.0463	0.24	-0.0232	-0.11	-0.00283	0.01
	Large van	-0.591	-2.04	-0.538	-2	-0.618	-2.25	-0.551	-1.99
	Compact pick-up truck	0.0277	0.16	0.0377	0.23	0.00316	0.02	0.0291	0.18
	Standard pick-up truck	-0.0608	-0.33	-0.0364	-0.22	-0.159	-0.87	-0.0999	-0.56
Fuel type	Standard Gasoline	0	-	0	-	0	-	0	-
	Flex Fuel/E85	0.304	2.81	0.297	2.77	0.271	2.9	0.314	3.3
	Clean Diesel	0.301	1.99	0.241	1.61	0.268	2.01	0.214	1.6
	Compressed Natural Gas	-2.15	-2.15	-2.2	-2.2	-1.7	-1.96	-1.9	-2.1
	Hybrid-electric	0.197	1.77	0.176	1.59	0.188	1.98	0.171	1.8
	Plug-in Hybrid-electric	0.538	5.57	0.509	5.32	0.467	5.54	0.445	5.49
	Full Electric	-2.86	-3.67	-3	-3.85	-2.4	-3.53	-2.64	-3.71
Age	1 or 2 years old	-0.209	-2.27	-0.193	-2.26	-0.187	-2.33	-0.174	-2.29
	3 or more years old	-0.426	-5.67	-0.397	-5.7	-0.404	-6.08	-0.384	-6.29
Incentives	HOV lane use	0.0606	0.98	0.0699	1.16	0.0424	0.78	0.0472	0.85

	Free parking	0.0446	0.74	0.0443	0.76	0.0362	0.68	0.0306	0.57
	$1,000 tax credit	0.185	3.1	0.181	3.14	0.16	3.03	0.157	3.05
	$1,000 reduced purchase	0.0706	1.15	0.0687	1.16	0.0543	0.98	0.0561	1.01
Costs	Vehicle price	-0.0744	-8.98	-0.0702	-8.88	-0.0689	-8.99	-0.0653	-8.98
	Vehicle price * income cat	0.00684	3.94	0.00663	4.11	0.00629	4.04	0.00618	4.21
	Fuel costs	-0.158	-3.03	-0.158	-3.12	-0.136	-2.84	-0.133	-2.73
	Maintenance costs	-0.0579	-3.24	-0.0546	-3.28	-0.0525	-3.34	-0.0481	-3.19
Performance, efficiency, and fuel availability	MPGE	0.0241	7.51	0.025	7.86	0.0212	7.18	0.0224	7.37
	Range (log transform)	0.285	1.55	0.292	1.59	0.213	1.34	0.242	1.46
	Acceleration (seconds to	-0.0406	-5.36	-0.0393	-5.57	-0.0375	-5.64	-0.0357	-5.54
	Plug-in at work and other	0.123	0.9	0.119	0.88	0.123	1.06	0.118	0.97
	1 in 20 Stations (CNG)	0.329	2.27	0.324	2.24	0.306	2.36	0.31	2.34
Other	Large HH - Medium	0.395	2.54	0.401	2.73	0.384	2.49	0.36	2.45
	Large HH - Large vehicles	0.728	3.1	0.693	3.16	0.72	3.07	0.659	2.94
	Alt Fuel - Medium vehicles	0.0884	0.8	0.0633	0.61	0.0853	0.86	0.0564	0.51
	Alt Fuel - Large vehicles	-0.137	-0.77	-0.154	-0.93	-0.154	-0.98	-0.161	-0.99

Table 4: Estimation results for different discrete choice models (part 2): structural parameters

		NL (fuel type)		NL (vehicle type)		CNL	
		est.	t-rat. (1)	est.	t-rat. (1)	est.	t-rat. (1)
Fuel type	Standard Gasoline	0.68	-5.46	-	-	0.4	-3.02
	Flex Fuel/E85	0.76	-0.50	-	-	0.08	-76.52
	Clean Diesel	0.82	-0.99	-	-	0.8	-0.64
	Compressed Natural Gas	0.9	-0.16	-	-	0.97	-0.03
	Hybrid-electric	0.56	-2.82	-	-	0.46	-2.54
	Plug-in Hybrid-electric	0.74	-3.33	-	-	0.6	-2.72
	Full Electric	1	-	-	-	1	-
Vehicle type	Subcompact car	-	-	0.9	-0.74	0.81	-0.82
	Compact car	-	-	0.72	-4.34	0.53	-4.89
	Mid-size car	-	-	0.71	-4.59	0.48	-3.99
	Large car	-	-	1	-	1	-
	Sport car	-	-	1	-	1	-
	Small cross-utility car	-	-	0.77	-1.81	0.65	-1.65
	Small cross-utility SUV	-	-	0.61	-3.95	0.37	-5.54
	Mid-size cross-utility SUV	-	-	0.75	-1.42	0.66	-0.77
	Compact SUV	-	-	0.68	-1.24	0.15	-6.70
	Mid-size SUV	-	-	0.77	-1.98	0.6	-1.67
	Large SUV	-	-	0.76	-0.37	0.4	-0.96
	Compact van	-	-	1	-	1	-
	Large van	-	-	0.63	-2.26	0.47	-2.26
	Compact pick-up truck	-	-	0.71	-1.34	0.39	-2.03
	Standard pick-up truck	-	-	1	-	1	-

Table 5: Implied monetary valuations

	MNL			NL fuel			NL vehicle			CNL		
	income < $20K	income between $60K and $80K	income above $120K	income < $20K	income between $60K and $80K	income above $120K	income < $20K	income between $60K and $80K	income above $120K	income < $20K	income between $60K and $80K	income above $120K
Price reduction needed to accept a 1 or 2 year old car	$3,094	$4,791	$7,881	$3,036	$4,781	$8,113	$2,987	$4,606	$7,519	$2,943	$4,641	$7,895
Price reduction needed to accept a car older than 3 years	$6,306	$9,766	$16,063	$6,245	$9,835	$16,688	$6,453	$9,952	$16,244	$6,495	$10,243	$17,423
Implied value of HOV lane use	$897	$1,389	$2,285	$1,100	$1,732	$2,938	$677	$1,044	$1,705	$798	$1,259	$2,142
Implied value of free parking	$660	$1,022	$1,682	$697	$1,097	$1,862	$578	$892	$1,456	$518	$816	$1,388
Implied value of $1,000 tax credit	$2,738	$4,241	$6,976	$2,847	$4,484	$7,608	$2,556	$3,941	$6,433	$2,656	$4,188	$7,123
Implied value of $1,000 reduced purchase price	$1,045	$1,619	$2,662	$1,081	$1,702	$2,888	$867	$1,338	$2,183	$949	$1,496	$2,545
Fuel cost sensitivity relative to vehicle price sensitivity	2.34	3.62	5.96	2.49	3.91	6.64	2.17	3.35	5.47	2.25	3.55	6.03
Maintenance cost sensitivity relative to vehicle price sensitivity	0.86	1.33	2.18	0.86	1.35	2.30	0.84	1.29	2.11	0.81	1.28	2.18
Implied value of one mile increase in MPGE	$357	$552	$909	$393	$619	$1,051	$339	$522	$852	$379	$597	$1,016
Implied value of one mile increase in range (base of 134 miles)	$31	$49	$80	$34	$54	$92	$25	$39	$64	$31	$48	$82
Implied value of second reduction in time to 60mph	$601	$931	$1,531	$618	$974	$1,652	$599	$924	$1,508	$604	$952	$1,620
Implied value of ability to plug-in at work and other locations	$1,821	$2,820	$4,638	$1,872	$2,948	$5,002	$1,965	$3,030	$4,946	$1,996	$3,148	$5,354
Implied value of CNG availability at 1 in 20 stations vs 1 in 50	$4,870	$7,542	$12,406	$5,097	$8,027	$13,619	$4,887	$7,538	$12,304	$5,244	$8,269	$14,065

Table 6: Forecasting example

MNL

Vehicle type	Fuel type 1	2	3	4	5	6	7	Total
1	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	17.55%	-1.28%	**9.88%**
2	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	17.55%	-1.28%	**9.88%**
3	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	17.55%	-1.28%	**9.88%**
4	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
5	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
6	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
7	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
8	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
9	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
10	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
11	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
12	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
13	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
14	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
15	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	-1.28%	**-8.94%**
Total	**-19.16%**	**-19.16%**	**-19.16%**	**-19.16%**	**-19.16%**	**37.31%**	**-19.16%**	

NL (fuel type)

Vehicle type	Fuel type 1	2	3	4	5	6	7	Total
1	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	19.92%	-1.23%	**12.54%**
2	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	19.92%	-1.23%	**12.54%**
3	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	19.92%	-1.23%	**12.54%**
4	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
5	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
6	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
7	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
8	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
9	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
10	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
11	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
12	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
13	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
14	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
15	-1.23%	-1.23%	-1.23%	-1.23%	-1.23%	-3.79%	-1.23%	**-11.18%**
Total	**-18.46%**	**-18.46%**	**-18.46%**	**-18.46%**	**-18.46%**	**14.26%**	**-18.46%**	

NL (vehicle type)

Vehicle type	Fuel type 1	2	3	4	5	6	7	Total
1	-1.69%	-1.69%	-1.69%	-1.69%	-1.69%	17.80%	-1.69%	**7.67%**
2	-2.40%	-2.40%	-2.40%	-2.40%	-2.40%	22.27%	-2.40%	**7.88%**
3	-2.64%	-2.64%	-2.64%	-2.64%	-2.64%	22.23%	-2.64%	**6.42%**
4	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
5	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
6	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
7	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
8	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
9	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
10	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
11	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
12	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
13	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
14	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
15	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	-1.32%	**-9.27%**
Total	**-22.61%**	**-22.61%**	**-22.61%**	**-22.61%**	**-22.61%**	**46.42%**	**-22.61%**	

CNL

Vehicle type	Fuel type 1	2	3	4	5	6	7	Total
1	-1.52%	-1.48%	-1.54%	-1.45%	-1.53%	19.05%	-1.46%	**10.08%**
2	-2.70%	-3.17%	-2.25%	-1.44%	-2.56%	24.78%	-1.52%	**11.13%**
3	-3.57%	-4.46%	-2.74%	-1.69%	-3.38%	26.15%	-1.57%	**8.73%**
4	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-2.23%	-1.20%	**-9.43%**
5	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-2.64%	-1.20%	**-9.84%**
6	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-3.54%	-1.20%	**-10.74%**
7	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-3.57%	-1.20%	**-10.77%**
8	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-2.77%	-1.20%	**-9.97%**
9	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-5.71%	-1.20%	**-12.91%**
10	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-2.98%	-1.20%	**-10.18%**
11	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-3.05%	-1.20%	**-10.25%**
12	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-2.23%	-1.20%	**-9.43%**
13	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-2.78%	-1.20%	**-9.98%**
14	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-3.56%	-1.20%	**-10.76%**
15	-1.20%	-1.20%	-1.20%	-1.20%	-1.20%	-2.11%	-1.20%	**-9.31%**
Total	**-22.19%**	**-23.52%**	**-20.94%**	**-18.98%**	**-21.87%**	**32.83%**	**-18.96%**	

Table 7: Summary of results for Mixed GEV models

	Mixed MNL	Mixed NL (fuel)	Mixed NL (vehicle)	Mixed CNL
Model fit for base model	0.262	0.263	0.263	0.264
Model fit for mixture model	0.314	0.314	0.316	0.316

Coefficient of variation	**Mixed MNL**	**Mixed NL (fuel)**	**Mixed NL (vehicle)**	**Mixed CNL**
Acceleration	2.68	2.54	2.46	1.97
Range (log transform)	12.18	9.39	9.75	5.45
Miles per gallon equivalent	6.88	6.92	6.64	7.49
Fuel cost	1.82	1.91	2.08	1.65
Maintenance cost	1.50	0.14	0.44	1.25
Price	1.00	1.02	1.00	1.08

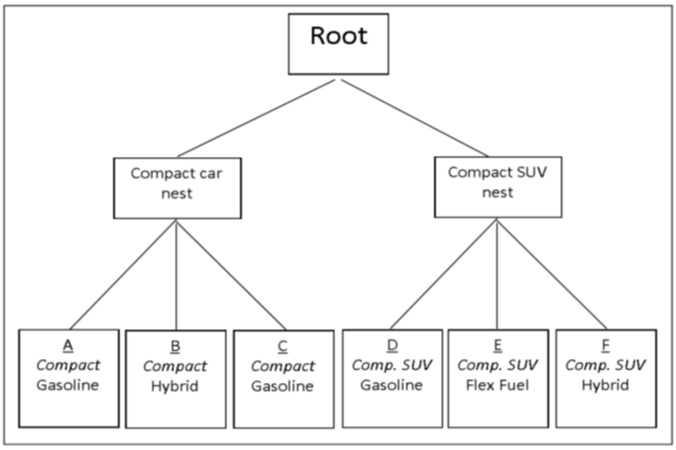
	NL (fuel)		**NL (vehicle)**		**CNL**	
Nesting parameter	**no mixture**	**mixture**	**no mixture**	**mixture**	**no mixture**	**mixture**
Standard Gasoline	0.68	0.81	-	-	0.4	0.59
Flex Fuel/E85	0.76	0.88	-	-	0.08	0.08
Clean Diesel	0.82	1	-	-	0.8	1
Compressed Natural Gas	0.9	1	-	-	0.97	1
Hybrid-electric	0.56	0.45	-	-	0.46	0.33
Plug-in Hybrid-electric	0.74	0.88	-	-	0.6	0.79
Full Electric	1	1	-	-	1	1
Subcompact car	-	-	0.9	0.63	0.81	0.21
Compact car	-	-	0.72	0.57	0.53	0.23
Mid-size car	-	-	0.71	0.59	0.48	0.20
Large car	-	-	1	1	1	1
Sport car	-	-	1	1	1	1
Small cross-utility car	-	-	0.77	0.79	0.65	0.68
Small cross-utility SUV	-	-	0.61	0.69	0.37	0.27
Mid-size cross-utility SUV	-	-	0.75	0.65	0.66	0.34
Compact SUV	-	-	0.68	0.48	0.15	0.16
Mid-size SUV	-	-	0.77	0.67	0.6	0.20
Large SUV	-	-	0.76	0.81	0.4	0.45
Compact van	-	-	1	1	1	1
Large van	-	-	0.63	0.88	0.47	0.87
Compact pick-up truck	-	-	0.71	0.76	0.39	0.46
Standard pick-up truck	-	-	1	1	1	1

2008 California Vehicle Survey

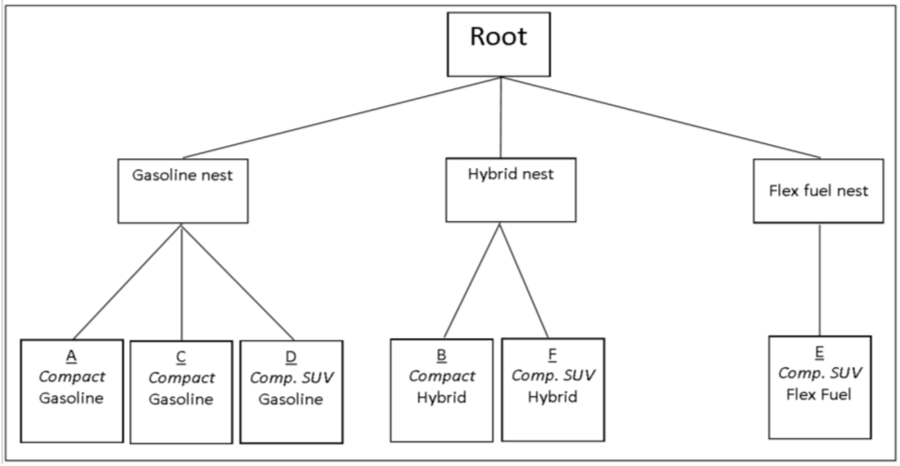
If the following vehicle options were available to you, which would you choose?
Please carefully examine all the attributes of each vehicle and then select the one you will most likely purchase by filling in the circle below your choice.

Vehicle Choice 1	Vehicle A	Vehicle B	Vehicle C	Vehicle D
Vehicle type	Midsize car	Compact SUV	Midsize car	Compact van
Fuel type	Gasoline	Natural Gas (NGV)	Plug-in Hybrid (PHEV)	Clean Diesel
Age of vehicle	New (2009)	New (2009)	New (2009)	New (2009)
Purchase price	$29,400	$36,600	$31,100	$20,900
Incentive	--	--	$1,000 tax credit	--
MPG or equivalent	29 MPG	15 MPG	60 MPG	31 MPG
Fuel cost per year	$1,090	$1,950	$780	$1,170
Fuel availability		1 in 50 stations		
Refueling time		10 Minutes at station, 4 hours at home		
Driving range		300 Miles		
Maintenance cost per year	$460	$370	$350	$550
Acceleration (0-60 mph)	10.2 seconds	11 seconds	8 seconds	11.8 seconds
Select One:	○	○	○	○

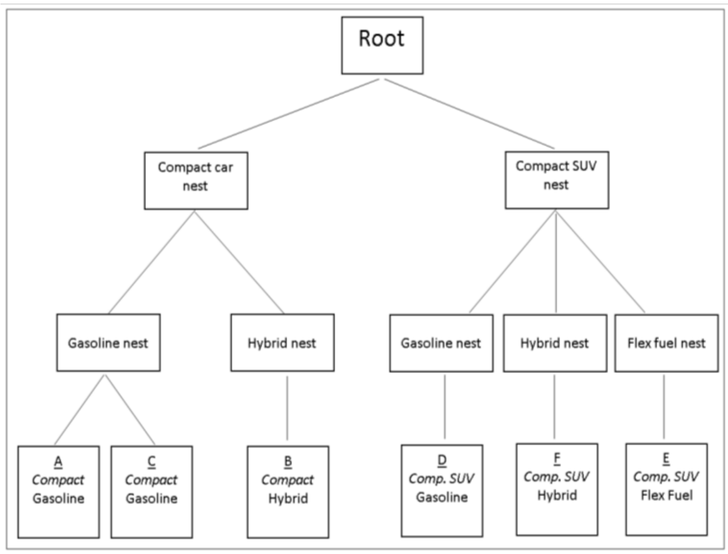
Figure 1: Example Stated Preference scenario



(2a): nesting by vehicle type



(2b): nesting by fuel type



(2c): vehicle type nested above fuel type

Figure 2: Different possible NL structures

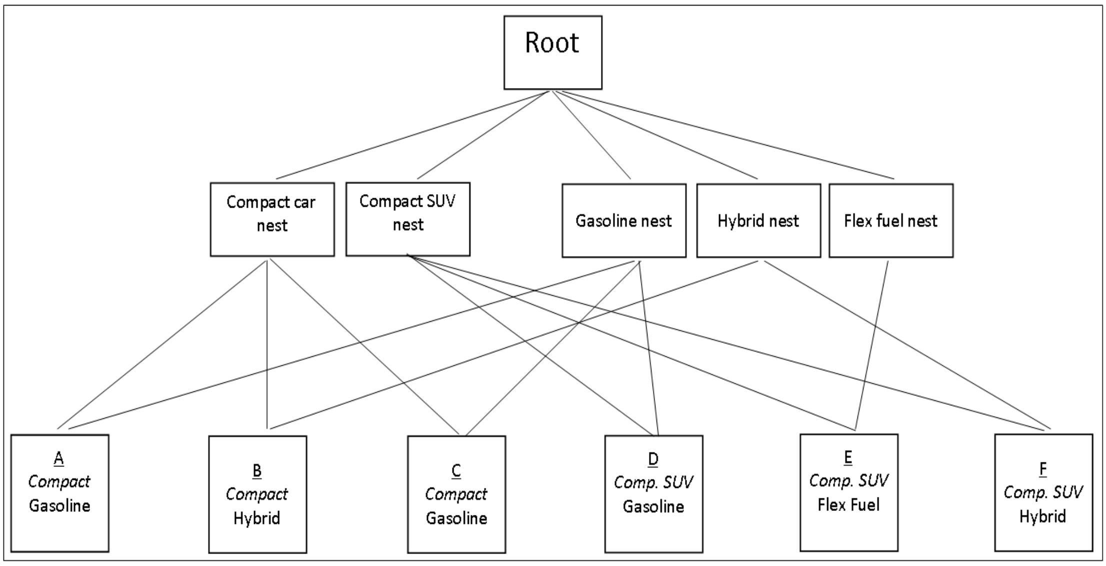


Figure 3: CNL structure