Consumer Vehicle Choice Model Documentation

Consumer Vehicle Choice Model Documentation

Assessment and Standards Division
Office of Transportation and Air Quality
U.S. Environmental Protection Agency

Prepared for EPA by
Oak Ridge National Laboratory
EPA Contract No. DE-AC05-00OR22725

TABLE OF CONTENTS

B.3 Sensitivity Analysis

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENT

1. INTRODUCTION

1.1 PROJECT OVERVIEW

In response to the Fuel Economy and Greenhouse Gas (GHG) emissions standards, automobile manufacturers will need to adopt new technologies to improve the fuel economy of their vehicles and to reduce the overall GHG emissions of their fleets. The U.S. Environmental Protection Agency (EPA) has developed the Optimization Model for reducing GHGs from Automobiles (OMEGA) to estimate the costs and benefits of meeting GHG emission standards through different technology packages. However, the model does not simulate the impact that increased technology costs will have on vehicle sales or on consumer surplus. As the model documentation states, "While OMEGA incorporates functions which generally minimize the cost of meeting a specified carbon dioxide (CO_2) target, it is not an economic simulation model which adjusts vehicle sales in response to the cost of the technology added to each vehicle."

Changes in the mix of vehicles sold, caused by the costs and benefits of added fuel economy technologies, could make it easier or more difficult for manufacturers to meet fuel economy and emissions standards, and impacts on consumer surplus could raise the costs or augment the benefits of the standards. Because the OMEGA model does not presently estimate such impacts, the EPA is investigating the feasibility of developing an adjunct to the OMEGA model to make such estimates. This project is an effort to develop and test a candidate model. The project statement of work spells out the key functional requirements for the new model.

> "ORNL shall develop a Nested Multinomial Logit (NMNL) or other appropriate model capable of estimating the consumer surplus impacts and the sales mix effects of GHG emission standards. The model will use output from the EPA's Optimization Model for reducing Emissions of Greenhouse gases from Automobiles (OMEGA), including changes in retail price equivalents, changes in fuel economy, and changes in emissions, to estimate these impacts. …The model will accept approximately 60 vehicle types, with the flexibility to function with fewer or more vehicle types, and will use a 15 year planning horizon, matching the OMEGA parameters. It will be calibrated to baseline sales projection data provided by the EPA and will include a buy/no-buy option to simulate the possibility that consumers will choose to keep their old vehicle or to buy a used vehicle. The first version of the model must be completed by the spring of 2011. Additional versions may be created in the future, pending further discussion and negotiation between the consultant and the EPA."

Briefly, given changes in each vehicle's price and fuel economy, the model

    (1)   calculates impacts of standards on vehicle sales mix, and
    (2)   calculates cost of standards in terms of consumer surplus.

The initial version of the model, at least, is not intended to project market trends due to other factors, although this might be a fruitful area for future research and development. The goal of this project is to create a simple model to test the concept of incorporating market share and consumer surplus changes to the OMEGA model and to produce a working initial model.

A research team at Oak Ridge National Laboratory (ORNL) has designed and implemented a Consumer Vehicle Choice Model (CVCM) for the project based on NMNL theory with a representative consumer. This document will detail CVCM design principles, model equations, parameter calibration, implementation and user guide. Specifically, Section 1.2 further explains CVCM functionality and intended usage and summarizes possible sources of prediction errors. Section 2 reviews relevant new vehicle type choice models in the literature and compares their merits and limitations. Then Section 3 describes NMNL equations and model calibration procedure. Finally, Section 4 gives instructions to use the model.

1.2   MODEL USAGE AND RESULTS INTERPRETATION

1.2.1   Model Functionality and Usage

The CVCM is intended to perform specific functions as an adjunct to the EPA's OMEGA model. As such, it has been designed to use the same theoretical basis and premises as the OMEGA model. Specifically, it has been designed to self-calibrate to the baseline vehicle sales distribution used by OMEGA and, given estimates for each individual vehicle of (1) changes in vehicle fuel economy, and (2) changes in vehicle prices, it:

1. - calculates impacts of those changes on vehicle sales and the distribution of vehicle sales and the resulting impact on manufacturers' abilities to meet fuel economy standards and,
2. - calculates changes in consumers' surplus as a consequence of the changes in fuel economy and vehicle purchase cost.

The CVCM is not intended to be a tool for forecasting the future vehicle fleet. There is no doubt that, over time periods longer than a few years, vehicle designs will come and go, new vehicle models will be introduced and others retired, new manufacturers will enter the U.S. market, existing manufacturers will exit, and there will be mergers and divestitures. However, predicting such events is outside the scope of the CVCM. It is also likely that over future time periods manufacturers will introduce new types of vehicles: plug-in hybrid, battery electric, hydrogen fuel cell vehicles and perhaps vehicles that are not foreseen at the present time. The CVCM was not designed to predict consumers' acceptance of these advanced technology vehicles. This capacity was left for future research and development.

The CVCM was developed to test the concept of predicting the differential sales impacts of fuel economy changes together with price changes brought about by fuel economy standards. It is intended to produce credible estimates of such changes to determine whether they may have important implications for manufacturers' abilities to meet the standards and for consumer well-being. Given the EPA's need for periodic and timely analyses to support its responsibilities for GHG emissions and fuel economy rulemakings, the CVCM should be capable of being readily calibrated to new data sets and updated with new sales and fuel economy data.

Given the intended purpose and functions of the model, it is most appropriately used for estimating *changes* in the following variables relative to the baseline values:

1. Market-wide consumers' surplus, total sales, total gross revenue, and fleet average miles per gallon (MPG) and GHG emissions,
2. Sales, average MPG and GHG emissions by manufacturer, and
3. Sales by market segment.

The CVCM models vehicle type choice at the most complete level of detail possible, corresponding to the level of detail at which fuel economy measurements are made by the EPA. Given that the price sensitivity of consumers' choices is greatest at the lowest[1] level of the NMNL nest, i.e., when vehicles are the closest substitutes, modeling at the greatest feasible level of detail can capture the full range of sales mix shifts. If vehicle type choice were modeled at a more aggregate level, the modeling process may be open to the questions about whether it misses important sales mix changes that would have been evident had the model operated at a greater level of detail. However, the CVCM prediction at the lowest level (i.e., make, model, engine and transmission configuration) is most sensitive to changes in input data and model parameters. Reporting results at this level may imply a higher degree of precision than is appropriate. Thus, we recommend reporting CVCM predictions at more aggregate levels.

The model provides highly detailed results. For reasons discussed in Section 1.2.2, the model's predictions are unlikely to be as precise as is suggested from the model output. The detail is provided for situations where the CVCM would be used iteratively with OMEGA, where the detail may provide advantages for model convergence. On the other hand, when final results are presented for consideration, false precision should be avoided. The sensitivity analyses we have done (see Appendix B) suggest that outputs should be presented to no more than three digits and perhaps only two in the case of consumers' surplus impacts and impacts on total vehicle sales.

1.2.2 Prediction Errors

The aggregate, or representative consumer, NMNL model makes simplifying assumptions about consumer behavior. Since consumer behavior is complex, we have focused the modeling initially on the decisions by consumers to trade-off fuel savings for higher vehicle prices, holding all other vehicle attributes constant. A change in a particular vehicle's fuel economy is translated into a change in price equivalent (present value dollars) based on a model or theory of how consumers value fuel economy. The change in the present value of future fuel costs perceived by consumers is added to the estimated change in the vehicle's price Δp. A price sensitivity parameter, B, translates the resulting net change in present value into a change in a utility index that determines a vehicle's market share. The change in utility for the ith vehicle in nest j, U_{ij}, is the following, where PV represents whatever function is chosen to transform a change in fuel economy to a change in the present value of fuel savings considered in the vehicle purchase decision.

$$\Delta U_{ij} = B_j \left(\Delta p_i - PV \left\{ fuel\ savings \right\} \right) \qquad (1)$$

---

[1] In this document, the higher/lower levels are referred to by their relative positions in the nested tree/nesting structure (Figure 1 in chap. 3) which has a buy/no-buy decision at the top/highest level and vehicle configurations (combinations of make, model, engine, and transmission) at the bottom/lowest level. It implies that the lower levels in the tree are more disaggregated.

The NMNL model is a tool for estimating changes in market shares as a function of changes in the present value (in dollars) of vehicles. If the changes are small relative to the prices of the vehicles and if the price sensitivity parameters are reasonably accurate, the NMNL model should give reasonably accurate predictions.[2] Prediction errors arise from incorrectly estimating changes in the utility index, caused either by errors in the estimation of the role of a change in fuel economy or inaccurate specification of the price sensitivity parameter, B. Such errors have a specific functional form in logit models.

For illustrative purposes, consider a simple multinomial logit (MNL) model (a derivation of the NMNL model that begins with a specification of the simple MNL model can be found in section 2.2.1 below).[3] The derivative of a vehicle's market share S_i with respect to a change in its utility index is the following:

$$\frac{dS_i}{dU_i} = \frac{d}{dU_i} \frac{e^{U_i}}{\sum_{j=1}^{N} e^{U_j}} = S_i - (S_i)^2 \tag{2}$$

Since, in general, S_i (which is between 0 and 1) will be approximately two orders of magnitude larger than $(S_i)^2$, the change in market share dS_i is approximately the change in utility weighted by the vehicle's market share, $S_i dU_i$. As changes in utility are propagated up the nesting structure (as inclusive values, or expected utility changes) this simple relationship applies at each step. Since a shock (error) in the utility index of vehicle i is a change in its utility, the impact of errors in the utility index on the predicted share is proportional to the market share of vehicle i.

Prediction errors will be negatively correlated between alternatives within a nest. At the lowest level of nesting, shocks to the utilities of individual vehicles are independent and identically distributed, in theory. However, the errors in one vehicle's utility index induce a change in the predicted shares of other vehicles that are negatively related to changes for the initial vehicle. The error term of a utility function directly induces a change in utility so its impact can be described by the derivative of the share of vehicle i with respect to a change in (shock to) the utility of vehicle j.

$$\frac{\partial S_i}{\partial U_j} = \frac{\partial}{\partial U_j} \frac{e^{U_i}}{\sum_{j=1}^{N} e^{U_j}} = e^{U_i} (-1) \left(\sum_{j=1}^{N} e^{U_j} \right)^{-2} e^{U_j} = -S_i S_j \tag{3}$$

Thus, a shock to the utility index of vehicle j induces a negative error in the prediction of the share of vehicle i that is proportional to the product of their shares, and prediction errors within a nest are negatively correlated. Because of this, errors in utility functions within a nest will tend to cancel, and the sum of the shares within a nest (i.e., the share of that nest) will have a smaller relative error than the relative errors of the individual vehicles within the nest.

---

[2] There is reason to expect the changes in dollar value to be small relative to the vehicle's price in that they will, in general, be comprised of an increase in price (>0) minus a present value of future fuel savings (also >0).
[3] Each lowest level nest of a NMNL model is a simple multinomial logit model.

In reality, prediction errors can arise from a number of simplifications in the CVCM, errors in parameters and errors in input data.

1. Non-optimizing consumer behavior
2. Aggregate NMNL model applied to heterogeneous consumers
3. Errors in NMNL model structure
4. Errors in NMNL parameters
5. Omitted variables (including manufacturer pricing decisions)
6. Inaccuracies in baseline sales data
7. Inaccuracies in OMEGA model predictions
8. Unanticipated technological innovations over time
9. Changes in consumers' behavior over time

There is substantial evidence that consumers' decision-making in markets for energy efficiency, and in particular fuel economy, may not correspond to the classical rational economic model (e.g., Jaffe and Stavins, 1994; Greene, 2009). A review of the econometric evidence found contradictory and inconclusive results (Greene, 2010). If, as Greene (2011) proposes, consumers' decisions about fuel economy are best described by prospect theory of behavioral economics, then the theory of utility maximization that underlies random utility models like the NMNL model would not be the most appropriate context for evaluating consumers' surplus impacts. Further theoretical and empirical research is needed to better understand how consumers' value fuel economy and how fuel economy and emissions standards affect consumers' surplus.

The CVCM is a market or representative consumer NMNL. It does not explicitly represent differences in consumers' preferences. The only recognition of differences in consumers' tastes is in the logit formulation itself which assumes that each individual perceives a different value for each vehicle (e.g., Train, 1993 ch. 2). However, this representation of heterogeneity is very limited and, in particular, does not allow for different price sensitivities. The population of consumers is undoubtedly heterogeneous but it is not known how important that heterogeneity is to the intended purpose of the CVCM. If further research and development is undertaken, investigating the importance of consumer heterogeneity should be given a high priority. Explicit heterogeneity was not incorporated in the CVCM in order to keep the model and its calibration simple.

The nesting structure used in the CVCM is similar to nesting structures used in empirically estimated models and in constructed models such as Bunch et al. (2011) and NERA (2009). Grouping vehicles by size, functionality and price is intuitive and consistent with the theoretical requirement that vehicles in a nest be similar with respect to unobserved attributes, i.e., be close substitutes. However, there is no guarantee that the nesting structure chosen is the best possible nesting structure.

Price sensitivities and alternative-specific constants are the two classes of parameters of the CVCM. Price sensitivities are the most important because the constants are computed so that the model exactly predicts the baseline market shares, given the assumed price sensitivities. The price sensitivities have been chosen to be consistent with the estimates in the published literature

and to conform to the theoretical requirement that price sensitivities increase in absolute value as one moves down along the nesting tree. However, the price sensitivity parameters have not been estimated to be consistent with a specific data set and it is always possible that an additional empirical analysis could yield insights missing from the existing literature.

Numerous possible explanatory variables have been excluded from the CVCM. Indeed, the only variables included are the changes in price and fuel economy supplied by the OMEGA model. Other variables are implicitly held constant in that they are included in the baseline constant terms. Including factors such as income and demographic variables may be desirable in a model to be used for estimation over an extended time period. However, this would require predicting values for those variables over the same time period. A potentially important endogenous variable in the OMEGA/CVCM system might be internal pricing decisions by manufacturers to meet especially stringent (strongly binding) fuel economy and emissions standards. This is beyond the scope of the current CVCM project, however.

To the extent that the baseline sales data, including the definitions of individual vehicles, differ from the actual market data, errors could be induced in the CVCM estimates. OMEGA is itself a model and thus its estimates undoubtedly contain some differences from what will occur, and these will also affect the accuracy of CVCM estimates.

Over extended time periods, automotive technology will change, and may change in ways that cannot be foreseen at the present time. Furthermore, consumers' preferences may also change in unpredictable ways. The 2002 National Research Council report on the CAFE standards and potential for fuel economy improvement did not foresee a successful market for hybrid vehicles (NRC, 2002). The emergence of minivans, SUVs, crossovers, the near disappearance of station wagons, and more, could not have been predicted with any certainty a long time period (e.g. 15 years) in advance. Assessment of technological innovation and trends in consumers' preferences is beyond the state-of-the-art in economic modeling and is probably best handled by scenario analysis.

The CVCM was designed to estimate the impacts of changes in vehicle prices and fuel economy provided by the OMEGA model on consumers' surplus and changes in vehicle sales that could impact manufacturers' abilities to meet fuel economy and GHG emissions standards. It was developed as a first test of the potential for such estimations to contribute to improved rule making. The goal was to develop a simple model that could be readily calibrated and operated in conjunction with the OMEGA model, and that had a sound theoretical and empirical basis.

2. LITERATURE REVIEW ON NEW VEHICLE TYPE CHOICE MODELING

The impacts of changes in vehicle prices and fuel economy on vehicle sales and consumer surplus can be estimated by means of systems of demand equations and discrete choice models, which are reviewed in this section. The emphasis is on two types of discrete choice models – NMNL and Mixed Logit (ML) models.

2.1 AGGREGATE DEMAND MODELS

Automobile demand by type of vehicle can be represented by a system of linear or non-linear demand equations. Kleit (2004, 2002a & b, 1990) created a market segment vehicle demand model that he used to evaluate the costs and benefits of CAFE standards. Kleit divided the market into eleven vehicle classes and four manufacturers. Demand functions, in the 2002b paper at least, were specified as simple linear functions of vehicle price (i.e., $Q = a + bP$). These equations can be calibrated given an initial set of prices and quantities and own price elasticities. Kleit estimated own price elasticities and cross price elasticities by exercising a proprietary model developed by GM (Table 1). Own price elasticities for the eleven vehicle classes ranged from -1.5 for large trucks to -4.5 for large cars. The own price elasticity for luxury cars is -1.7, less than those of standard cars (-2.8 to -4.5) but of the same order of magnitude. In general, cross price elasticities are small relative to own price elasticities. Cross elasticities are not symmetric because classes with high sales volumes have a greater effect on classes with low sales volumes than vice versa. However, there are sets of classes for which cross price elasticities are substantial, indicating that the vehicle types are relatively close substitutes. The groupings in Kleit's model suggest that standard cars are relatively close substitutes, with small cars being better substitutes for medium cars than for large cars. Small and Large SUVs are relatively good substitutes, as are small and large (pickup) trucks. Cars and pickup trucks are not close substitutes, and the only vehicle that is even a weak substitute for a full size van is a minivan. In the discussion of discrete choice models below, such grouping become "nests".

Table 1 Demand Elasticities of Kleit's Vehicle Class Demand Model (Kleit, 2004)

		1	2	3	4	5	6	7	8	9	10	11
1	Small Car	**-2.808**	0.423	0.063	0.018	0.000	0.036	0.027	0.009	0.009	0.009	0.000
2	Medium Car	0.684	**-3.528**	1.107	0.027	0.018	0.018	0.018	0.036	0.045	0.054	0.009
3	Large Car	0.270	1.926	**-4.500**	0.027	0.216	0.009	0.054	0.018	0.063	0.054	0.009
4	Sport Car	0.549	0.423	0.324	**-2.250**	0.009	0.090	0.198	0.045	0.108	0.018	0.000
5	Luxury Car	0.045	0.405	1.062	0.009	**-1.737**	0.000	0.027	0.045	0.189	0.072	0.009
6	Small Truck	0.162	0.099	0.000	0.009	0.000	**-2.988**	0.702	0.045	0.054	0.009	0.009
7	Large Truck	0.063	0.072	0.018	0.009	0.000	0.234	**-1.548**	0.027	0.090	0.018	0.036
8	Small SUV	0.216	0.279	0.099	0.027	0.009	0.090	0.351	**-3.645**	0.747	0.108	0.072
9	Large SUV	0.117	0.243	0.171	0.018	0.018	0.054	0.387	0.414	**-2.043**	0.234	0.108
10	Minivan	0.081	0.171	0.063	0.000	0.009	0.009	0.045	0.027	0.135	**-2.286**	0.180
11	Van	0.027	0.036	0.009	0.009	0.000	0.009	0.054	0.036	0.072	0.387	**-2.385**

Automobile supply was represented by assuming a short-run price elasticity of supply of +2 and a long-run elasticity of +4.

Bordley[4] (1993) estimated own and cross price elasticities for 200 passenger car nameplates using aggregate time-series sales data by market segment plus survey data on the first and second choices of consumers who had just purchased a new car. The aggregate sales data allowed estimation of own price elasticities for seven passenger car market segments and an overall price elasticity of automobile demand. The survey data were used to estimate "diversion fractions" quantifying the propensity of purchasers of one nameplate to buy any of the others given an increase in its price. Bordley estimated an own price elasticity for passenger car purchases versus all other commodities of -1.0. Car segment elasticities ranged from -1.7 for small cars to -3.4 for sporty cars (Table 2). Elasticities for individual nameplates ranged from -1.7 to -8.2; mean values within segments ranged from -2.4 to -4.7.

Table 2 Market Segment and Nameplate Own Price Elasticities Estimated by Bordley (1993)

Car Class	Class Elasticity	Car Nameplate Elasticities		
		Minimum	Average	Maximum
Economy	-1.9	-3.3/-3.4	-4.7	-8.2/-8.1
Small	-1.7	-1.9/-1.7	-2.4	-3.1/-3.4
Compact	-2.0	-2.1/-2.2	-3.1	-4.9/-4.7
Midsize	-2.3	-2.3/-2.6	-3.3	-4.6/-4.2
Large	-3.0	-3.1/-3.5	-3.8	-4.3/-4.0
Luxury	-2.4	-3.2/-3.4	-3.7	-5.3/-4.5
Sporty	-3.4	-2.6/-3.4	-4.2	-6.5/-5.3

Bordley's method could be used to calibrate a system of linear nameplate demand equations, as was done by Kleit (1990). More complex systems including cross price elasticities can also be calibrated, as Bordley (1993) points out, but does not explicitly describe calibration of such a model.

Austin and Dinan (2005) used the own- and cross-price elasticity matrix developed by Kleit (2002a) to estimate the impacts of changes in vehicle prices due to fuel economy standards on consumers' demand for 10 vehicle classes. Consumer demand for a class is a linear function of the difference between vehicle price and the value of future fuel savings induced by the standards. For manufacturer i, demand for its vehicle classes is given by the following matrix equation,

$$q_i = A_i(p_i - c_i) \tag{4}$$

in which q_i is the vector of quantities for each of the 10 vehicle classes, p_i is the vector of prices, c_i is the vector of present value of fuel economy improvements and A_i is a matrix of own- and cross-price elasticities. Austin and Dinan (2005) do not provide the numerical values for the

---

[4] Bordley was employed by General Motors Research Laboratory at the time he conducted and published his study. Thus, there may be a relationship between Kleit's elasticity estimates, which are based on a GM model, and Bordley's.

elasticities they used in their model, nor does Kleit (2002a), apparently because the model is proprietary.

2.2 DISCRETE CHOICE OR RANDOM UTILITY MODELS

Discrete choice models, sometimes referred to as random utility (RU) models, are by far the most common methodology used to mathematically model automobile demand. Baltas and Doyle (2001) succinctly summarize the methodology.

> "In RU models, preferences for such discrete alternatives are determined by the realization of latent indices of attractiveness, called product utilities. Utility maximization is the objective of the decision process and leads to observed choice in the sense that the consumer chooses the alternative for which the utility is maximal. Individual preferences depend on characteristics of the alternatives and the tastes of the consumer….The analyst cannot observe all the factors affecting preferences and the latter are treated as random variables." (Baltas and Doyle, 2001, pp. 115).

Since the early applications of random utility models in the 1970s (McFadden, 1973), formulations of RU models have proliferated. Baltas and Doyle (2001) identified fourteen different methods which they grouped into three fundamentally different approaches depending on the nature of the random utility:

- Unobserved product heterogeneity,
- Taste Variation (consumer heterogeneity), and
- Choice Set Heterogeneity.

Nearly all applications of random utility models to automobile choice fall into the first two groups because the availability of different types of automobiles is rarely a significant issue. Randomness in the simple multinomial logit model derives primarily from unobserved attributes. Its error term may also include unobserved variations in taste but the representation of these variations is limited and simplistic. The same applies to NMNL Models although their ability to represent randomness in unobserved attributes and tastes is much more complex. In these models, heterogeneity in consumers' preferences is commonly represented by explicit functional relationships between product attributes and consumer characteristics. MNL models allow variations in consumers' preferences to be represented by random coefficients, whose distributions can be inferred either from survey or market shares data.

Which methodology is best for a given application depends not only on the richness of the modeling approach but on the objectives of the exercise, as well as practical constraints, including data and resource availability. Baltas and Doyle sum up the dilemma well.

> "Finally, a general concern relates to overall model practicality. As our discussion illustrates, recent developments have increased model complexity and made estimation, interpretation, and forecasting less straightforward. Some specifications are still rather impractical. The issue can be viewed as the common dilemma between simplicity and flexibility. There is no universal answer to this question as it depends on one's rate of exchange between the two criteria." (Baltas and Doyle, 2001, p. 123).

2.2.1 Multinomial Logit

The first application of a multinomial logit model to automobile choice appears to be the seminal paper by Lave and Train (1979) which estimated a multinomial logit model of consumers' choices among 10 vehicle classes using what was then a new method for analyzing qualitative choice behavior (McFadden, 1973). The probability of an individual consumer choosing a vehicle class was assumed to be a function of a vector of vehicle attributes and household attributes. The model formulation allowed for interaction of household and vehicle variables in a linear "representative" utility function. Let X_{ijk} be the kth variable, for the jth vehicle class and the ith consumer. The representative utility function is defined as,

$$V_{ij} = \sum_{k=1}^{K} \beta_k X_{ijk} + \varepsilon_{ij} \tag{5}$$

in which the β_ks are fixed coefficients and the ε_{ijk}s are independent, identically distributed random variables that have extreme value distributions. The probability that consumer i will purchase a vehicle from class j is a multinomial logit function of the representative utilities of all classes.

$$P_{ij} = \frac{e^{V_j}}{\sum_{l=1}^{L} e^{V_l}} \tag{6}$$

In the Lave and Train model, vehicle price was represented by price divided by household income and the same variable squared. The results implied both that sensitivity to price decreased with increasing vehicle price and that price sensitivity decreased with increasing income. The model was calibrated to survey data from 541 households collected in seven U.S. cities in 1976.

McCarthy and Tay (1998) estimated a MNL model of consumers' choices among 68 makes and models. Their objective was to test whether buyers of domestic, European and Japanese manufactured vehicles valued vehicle attributes in the same way. Their analysis rejected the hypothesis that vehicle attributes are similarly valued regardless of country of origin. They also noted certain "anomalies" in their coefficient estimates. For example, faster acceleration decreased the probability of choosing American and Japanese vehicles, while operating costs were an insignificant variable for makes and models of Japanese manufacture. Similar results have been observed in other studies and may point to an inherent difficulty in estimating random utility models. A key assumption of such models is that the unobserved attributes are uncorrelated with the observed attributes. If they are not, then biased estimates can result. Given the strong correlations among many observed attributes (e.g., size, price, horsepower, fuel economy, weight, interior volume, number of seats, etc.), the assumption that unobserved attributes are uncorrelated with observed attributes seems unlikely. In addition, the problem of defining and obtaining measures of precisely the right attributes that determine consumers' choices has also been a persistent issue for random utility models. Is acceleration best measured by the ratio of horsepower to weight, by 0-60 mph time, or by the various measures the industry uses to capture the experiences of launch from a stop, intermediate speed range acceleration,

passing acceleration, and responsiveness? Inaccuracies in defining and measuring attributes lead to errors in observed variables. Correlated omitted variables, errors in observed variables and correlated observed variables makes statistical inferences challenging indeed.

Lave and Train (1979) noted two key limitations of the MNL model. First is the so-called Independence of Irrelevant Alternatives (IIA) property, which makes the ratio of the probabilities of choice of any two alternatives independent of the presence or attributes of any other alternatives. A related property is that all alternatives are assumed to have the same probability distribution of unobserved utility (i.e., ε has the same distribution for every alternative) and that these distributions are independent. These properties severely restrict the patterns of substitution the model can represent. For example, apart from the measured utility component, Lave and Train's MNL model implies that a two-seater sports car is just as good a substitute for a luxury sedan as it is for a sporty subcompact. Note that the measured utility component in the Lave-Train model directly accounted for factors such as the number of seats, household size and acceleration performance. Unobserved factors might be such things as styling or image.

Second, because automobile attributes do not vary across the population of consumers, it is not possible to estimate a MNL model that includes vehicle attributes and a vehicle specific constant. In a model estimated using household data, attributes can only be entered when interacted with some household characteristic. On one hand, this allows attribute values to vary across individuals. On the other, it imposes specific functional relationships on how attribute values vary that may not be supported by any theory. Thus, heterogeneity of consumers' preferences is an inherent property of MNL models estimated using household data but is restricted to specific functional relationships chosen by the researcher.

2.2.2 Probit and Nested Multinomial Logit

The shortcomings of the simple MNL model, especially its IIA property, led researchers to explore alternative formulations that allowed greater flexibility in patterns of substitution among vehicles and representations of heterogeneous consumer preferences. The probit model was derived by relaxing the assumptions of independent, identical error distributions (see, e.g., Train, 1993). Instead the error terms in a probit model are assumed to be jointly normally distributed. Instead of leading to a simple, closed form equation for the choice probabilities (like equation (6) for the MNL model), the probit model requires numerical integration of a series of integrals. The probit model's inherent complexity, combined with the ability of a variant of the MNL model to overcome most of its limitations, is responsible for the very infrequent use of probit models in modeling automobile choice.

The NMNL Model, a special case of the Generalized Extreme Value (GEV) model, is based on the premise that the full choice set can be portioned into subsets (nests) within which the IIA property is appropriate but across which it is not. Put another way, within a nest all vehicles are assumed to be equal substitutes, conditional on their observed utility. Formally, within a subset alternatives error terms are independent and identically distributed. Across subsets, they are not.

Building on the notation of equations (5) and (6), the probability that a consumer will choose a specific make, model, engine and transmission configuration, *m*, given that the consumer will choose a vehicle in nest (class) *j*, is a simple MNL probability.

$$P_{im|j} = \frac{e^{U_{ijm}}}{\sum\limits_{l=1}^{L} e^{U_{ijl}}} \tag{7}$$

The probability that consumer i will choose class j is a function of the utility of attributes common to class j, V_j, as well as a function of the composite utility of all vehicles within class j, I_{ij}.

$$P_{ij} = \frac{e^{V_{ij} + \lambda_j I_{ij}}}{\sum\limits_{l=1}^{L} e^{V_{il} + \lambda_l I_{il}}} \tag{8}$$

The term I_{ij} is the "inclusive value" or expected value of the utility of vehicles in set j. It is defined by equation (9).

$$I_{ij} = ln \left(\sum\limits_{m=1}^{M} e^{\sum\limits_{k=1}^{K} \beta_{jk} X_{ijmk}} \right) \tag{9}$$

In equation (9) each nest has a different set of coefficients that map vehicle attributes into the utility index. In particular for this model, these coefficients differ across nests. This allows different degrees of substitutability for the choices within different nests. The unconditional probability of consumer i choosing vehicle m in class j is the following.

$$P_{ijm} = P_{im|j} P_{ij} \tag{10}$$

Another feature of the NMNL model that helps overcome the limitations of the MNL model is the ability to define any number of levels of nesting. A key advantage of this is that the top nest can represent the choice to buy or not to buy a new automobile. Thus, an NMNL market model can predict the impacts of changes in vehicle attributes and other factors on total vehicle sales as well as the type of vehicles purchased.

The flexibility and mathematical simplicity of the NMNL model have made it the most widely used tool for modeling automobile choices. Goldberg (1995, 1998) estimated NMNL models of automobile choice in order to evaluate the impacts of fuel economy standards. In the 1995 study, her nests comprised (1) small cars including subcompacts and compacts, (2) luxury automobiles including sports cars, and (3) all other vehicles. A likelihood ratio test was used to test (and reject) the hypothesis that coefficient values within the three nests were equal. While such tests can be used to reject a nesting structure, there is no accepted methodology for identifying a correct nesting structure. Goldberg's 1998 study used nine vehicle classes, within which consumers could choose between a foreign or domestic car. This structure was chosen to allow exploration of differential impacts of standards on foreign and domestic manufacturers.

Stated preference survey data were used by Brownstone et al. (1996) to estimate a NMNL model of consumers' choices among conventional and alternative fuel vehicles. The 1993 California

survey asked households to choose among hypothetical vehicles that included alternative fuel vehicles. Stated preference methods were necessitated by the fact that very few households purchase or have any experience with vehicles powered by non-petroleum fuels.

Most often, NMNL models are calibrated via statistical inference based on the vehicle choices of individual consumers or households. However, MNL and NMNL models can also be interpreted as representing the choice probabilities of a representative consumer or a population of consumers with diverse tastes (Anderson et al., 1988). In this interpretation, the random error term (ε) represents not only unobserved attributes but also unobserved variations in tastes and errors in perception and optimization by consumers (Madalla, 1992, p. 60). Several modelers have used NMNL models in this way to represent aggregate market behavior.

Greene (1994) constructed a NMNL choice model for predicting market shares of alternative fuel vehicles. Rather than estimating a model based on stated preference survey data, Greene followed a methodology invented by Donndenlinger and Cook (1997) to infer the values of automobile attributes. The model coefficients were constructed by postulating how vehicle attributes such as range or fuel economy would be valued by consumers, deriving a coefficient that translates unit changes in each variable to a present dollar value and applying a multiplier to transform that coefficient into one that translates unit changes into the utility index. This multiplier is referred to as generalized cost coefficient in the remainder of this document. Greene reasoned that since the overwhelming majority of consumers had no first-hand experience with alternative fuel vehicles (e.g., battery electric vehicles, compressed natural gas vehicles, etc.) stated preference surveys data would likely be misleading. The model did not include a buy/no-buy decision. The first level nest included eight alternative fuel technologies. The second level nest comprised the choice of fuel for bi-fuel or flex-fuel vehicles. A similar model also constructed by Greene (2001) contained Conventional Internal Combustion Engine (ICE) vehicles, Dedicated Alternative Fuel vehicles (CNG and LPG), Hydrogen Fuel Cell vehicles and Battery Electric vehicles in the first level nest and subcategories of these vehicle technologies in the second. E.g., ICE was divided into conventional liquid fuel vehicles, hybrid vehicles and gaseous-fueled vehicles. Within the conventional liquid fuel nest, consumers chose among gasoline, diesel, ethanol FFVs and methanol FFVs. Within the FFV nests, consumers chose fuel types, e.g., gasoline or E85. To estimate price coefficients for the nests in his model, Greene (2001) relied on existing studies and the theoretical requirement that sensitivity to price must increase from the top nest to the bottom (from vehicle technology choice to fuel choice). Since the overall price elasticity of automobile demand is generally believed to be approximately -1.0 (Kleit, 1990; McCarthy, 1996; Bordley, 1993) and the choice of fuel is highly but not infinitely elastic (approximately -10 or more: Greene, 1998, p. 228), this bounds the range of price sensitivity for nests in between. Although this range is an order of magnitude, with three nests between the top and bottom choices it provides useful information that can be used in conjunction with estimates from published studies to greatly reduce the uncertainty about coefficient values.

Greene et al. (2005) and Greene (2009) calibrated constructed NMNL models to the market shares of over 800 carline/engine/transmission configurations. The data sets included every vehicle in the National Highway Traffic Safety Administration's (NHTSA) model year 2000 and 2005 fuel economy data sets, respectively, except those with annual sales below 25 units per year.

Generalized cost coefficients were chosen based on the published literature and the relative value rule for nests described above. Vehicle-specific constants were used to insure the model exactly fit the base year data. The ability to calibrate the model to fit any given year's sales data is an advantage for use in policy analysis where the correspondence of model estimates to real world experience is of value.

Harrison et al. (2008), like Greene et al. (2005), used a constructed NMNL model to evaluate the benefits and costs of the 2011-2015 CAFE standards. The authors assumed a plausible nesting structure based on their judgments about the substitutability of different types of vehicles. The guiding principle is that vehicles within a nest are closer substitutes for one another than they are for vehicles in other nests.[5] Consumers are assumed to decide to buy or not to buy at the top nest, then choose among three car classes: passenger cars, pickup truck/full-size van, or SUV/minivans. The next level contains 14 vehicle classes based on size and price. Within these subclasses are non-intersecting subsets of over 200 vehicle models. Like Greene et al. (2005), Harrison et al. made use of the NMNL requirement that price sensitivity (price coefficients) must decrease in absolute value (increase in value) as one moves up the nesting tree. They began with a price elasticity of -1.0 for the buy/no-buy decision, and then assumed the ratios of parameters at each level in order to calculate price coefficients for each lower nest. Harrison et al. also calculated a constant term for each model, as Greene et al. did, but then regressed those constant terms against other vehicle attributes in an effort to infer the value of those attributes.

2.2.3 Mixed Logit Model (MLM)

The MLM adds to the NMNL a greater capability to include heterogeneous consumer tastes. The utility of vehicle *m* to consumer *i* is given by equation (11).

$$U_{im} = \delta_m + \sum_{k=1}^{K} \beta_k X_{imk} + \sum_{h=1}^{H} \mu_{ih} z_{imh} + \varepsilon_{im} \tag{11}$$

In equation (11), δ_m represents the average utility (intercept term) of vehicle m, the X_{imk} are vehicle attributes interacted with consumer characteristics, the β_k are mean coefficient values for these variables, the μ_{ih} are individual specific random coefficients reflecting deviations of individual tastes from those β_k for which tastes vary, and the z_{imh} are vehicle attributes interacted with consumer characteristics for which tastes vary (Train and Winston, 2007). Assuming that the ε_{im} are independent and identically distributed and have an extreme value distribution, the probability that consumer *i* chooses vehicle *m* is given by the mixed logit model (the integral sign represents many integrals over the many probability distributions of the random variables).

$$P_{im} = \int \frac{e^{\delta_m + \sum_{k=1}^{K} \beta_k X_{imk} + \sum_{h=1}^{H} \mu_{ih} z_{imh}}}{\sum_l e^{\delta_l + \sum_{k=1}^{K} \beta_k X_{ilk} + \sum_{h=1}^{H} \mu_{ih} z_{ilh}}} \tag{12}$$

---

[5] More accurately, the vehicles are more similar with respect to their unobserved attributes. Vehicles may differ greatly with respect to the measured attributes that enter the utility index function yet still belong in the same nest.

Train and Winston (2007) estimated a mixed logit model of vehicle choice using a random sample of 458 U.S. consumers who had just purchased a new model year 2000 vehicle in order to investigate reasons for the declining market shares of U.S. auto manufacturers. Each consumer's choice set consisted of 200 makes and models. There is no closed form solution for estimating the parameters of the MLM. Instead, simulation was used to approximate the integrals for choice probabilities and the resulting log likelihood function.

The parameters of the MLM are functions of consumer attributes and random variables. For example, the price coefficient in the Train and Winston model is,

$$\beta = -0.073 - \frac{-1.60}{Y_i} + \frac{0.86\nu}{Y_i} \tag{13}$$

The variable ν is a standard normal random variable. This adds richness to the model by representing varying tastes across the population. There is even some small probability of finding a consumer who prefers higher prices ($\beta > 0$). On the other hand, the functional form is, to a degree, chosen a priori by the researcher, and both estimating the model and predicting with it are substantially more complicated but still quite feasible. Both require simulations (perhaps only a few hundred) and both require information about the distributions of consumer characteristics (available from national surveys).

A comparison of MLM and NMNL models was made by Brownstone et al. (2000), combining stated and revealed preference survey data for California households. The authors observed that the MLM improves the fit of model to data, and indicated substantial heterogeneity of preferences across the population. They also noted that revealed preference (RP) data are essential for obtaining realistic predictions of consumers' choices of vehicle types. However, they also commented on the difficulty of statistical inference using RP data.

> "RP data appear to be critical for obtaining realistic body-type choice and scaling information, but they are plagued by multicollinearity and difficulties with measuring vehicle attributes. SP data are critical for obtaining information about attributes not available in the marketplace, but pure SP models with these data give implausible forecasts." (Brownstone et al., 2000)

Bento et al. (2005, 2009) estimated a mixed logit model of vehicle choice and a paired model of vehicle use using data from the 2001 National Household Travel Survey. They divided vehicles into 10 vehicle classes, 5 age categories and 7 manufacturers. The paired models not only estimate new vehicle choices but vehicle use, as well as aging and scrappage. Jacobsen (2010) used the model to assess the impacts of CAFE standards on manufacturers but did not include in his model the option they have to use technology to improve the fuel economy of vehicles at increased cost. The mean price elasticity of new vehicle demand was estimated to be -2.0, substantially more than the unit elasticity found in models cited above.

Cambridge Econometrics (2008) estimated a mixed logit model of vehicle choice in the UK based on a survey of households who had purchased a new or less than 1-year-old vehicle during the years 2004 to 2007. Households identified the manufacturer, model and engine size of their

vehicle, which the researchers matched to a separate data base of vehicle attributes. The survey asked what attributes consumers considered important to the purchase of a vehicle. Respondents cited many difficult to measure factors, such as reliability, safety, comfort, warranty and security. Estimated mean price elasticities by vehicle class ranged from -0.96 for multi-passenger vehicles, to -3.51 for luxury vehicles. Relatively elastic market segments included Minicars (-2.46), Upper Medium cars (-2.81) and Executive cars (-3.24). Less price elastic segments were Superminicars (-1.15), Lower Medium cars (-1.15), Sports cars (-1.79) and 4X4s (-1.75). The observed patterns of own- and cross- price elasticities led the researchers to comment on the importance of models that allow flexibility in substitution patterns.

> "We observe substitution patterns that represent a significant departure from proportional substitution, i.e. there is a higher level of substitution between similar models of cars." (Cambridge Econometrics, 2008, p. vii)

Mixed logit models can also be estimated using aggregate market shares, as first shown by Boyd and Mellman (1980) and Cardell and Dunbar (1980) and later in a seminal paper by Berry, Levinsohn and Pakes (BLP) (1995). BLP provided a practical method of estimating a mixed logit model from aggregate sales data. Prices are endogenous in the BLP model, an issue they addressed by means of instrumental variables comprised of the attributes of other vehicles. Estimates relying on instrumental variables in this context can be unreliable, as Knittel and Metaxoglou (2008) demonstrated using BLP's data. Noting that the objective function in the BLP model is highly nonlinear and thus prone to multiple local optima, they tested 10 different optimization algorithms, using 50 different starting values for each. Their results call for caution both in interpreting parameter estimates from BLP-type models and in their use for forecasting.

> "We find that convergence may occur at a number of local extreme, at saddles and in regions of the objective function where first-order conditions are not satisfied. We find own- and cross-price elasticity estimates that differ by a factor of over 100 depending on the set of candidate parameter estimates." (Knittel and Metaxoglou, 2008)

On the other hand, other researchers, using variants of the BLP model and different estimation procedures, have obtained more stable results.

Moon, Shum and Weidner (2010) extend the BLP method by adding interactive fixed effects to the unobserved product characteristics. The specification multiplicatively combines time-specific fixed effects with vehicle-specific fixed effects. The consumer's utility function is,

$$U_{ijt} = \sum_{k=1}^{K} \alpha_{ik} X_{kjt} + \sum_{r=1}^{R} \lambda_{rj} f_{rt} + \delta_{jt} + \varepsilon_{ijt} \qquad (14)$$

in which α's are coefficients measuring the marginal value of each of the K vehicle attributes X, whose mean value also includes the R interactive fixed effects of product j, plus δ_{jt}, represented by the third hand side term. The difference between this formulation and that of BLP is the specific structure imposed on the distribution of product-specific tastes. The final term, ε_{ijt}, is the individual, product and time specific utility component. Note that if there are on the order of 10^3

vehicles and just a few time periods, this model has thousands of parameters. In addition, projecting taste heterogeneity into the future requires specifying future values for f_{rt} and δ_{jt} or assuming they remain constant. If these are assumed to be constant at the values of a given year or at average values, the heterogeneity of tastes is limited to product-specific heterogeneity.

An advantage of the Moon et al. approach is that it explicitly represents some endogenous factors by means of interactive fixed effects and thereby reduces the need for instrumental variables, in particular, to represent price endogeneity. The authors find that, given their formulation, coefficient estimates produced by methods that assume prices are exogenous versus endogenous differ little.

The Moon et al. method also produces price elasticities that are much higher in absolute value than those obtained by the standard BLP model estimation methods. This is apparently due to the inclusion of the fixed effect variables. They applied the method to the same data used by BLP (1995). Own and cross price elasticities were estimated for 23 vehicle classes. Using their interactive fixed effect formulation and assuming prices to be endogenous produced own price elasticities ranging from -7.0 for Cadillacs to -36.5 for large Mercurys. Twenty of the 23 estimated elasticities were more price elastic than -25.0. Omitting the interactive fixed effects produced own price elasticity estimates ranging from -7.8 (again, for Cadillac) to -17.6 for a "remainder of the market" category. This time, 10 of the 23 elasticity estimates were more elastic than -15.0.

In a study for the UK Department of Transport, the Economics for the Environment Consultancy (EFTEC) estimated a MLM model of consumers' choices of automobiles in the UK (EFTEC, 2008). The researchers estimated their model using the method of BLP and data on new car market shares for 2,190 different vehicle types registered by private households in 11 regions of the UK. They note that their choice set is considerably larger than that of any previous study. The ability to calibrate a model to such a large choice set is a consequence of the BLP estimation procedure. Vehicles were nested into 9 classes based on size, body style and price. Estimated median price elasticities ranged from -1.3, for vehicles in the SUV class with a range from -2.4 (90th percentile) to -1.0 (10th percentile), to -5.4 for vehicles in the small-to-medium size family car segment with a range from -7.1 (90th) to -4.5 (10th). Sports cars also had relatively low price elasticities and subcompact and mini car choices were relatively price elastic.

A number of recent studies have employed forms of the Mixed Logit model to estimate the relative effects of vehicle price and fuel economy or fuel costs on vehicle choice (e.g., Allcott and Wozny, 2009; Klier and Linn, 2008; Gramlich, 2008; Sawhill, 2008). These and other related studies were reviewed by Greene (2010). All used extensive, detailed data bases on vehicle purchases in the United States but reached very different conclusions about how consumers trade off vehicle price and fuel economy. Some of the differences can be attributed to how consumers form expectations about future fuel prices, although most models assumed static expectations based on the observation that fuel prices appear to follow a random walk.

Aggregate, mixed logit type models can be used to predict market shares and estimate changes in consumer surplus. For example, Greene and Liu (1988) used both a random coefficient MNL model and Lave and Train's (1979) model to estimate the impacts of changes in vehicle

attributes related to fuel economy on the consumer surplus associated with automobiles sold in the United States between 1978 and 1985. The random coefficient model utilized Monte Carlo simulation to execute repeated draws from the vector distribution of random coefficients. Greene and Liu found that the estimated mean consumer surplus values were highly sensitive to the mean values of attributes but they did not test sensitivity of consumer surplus estimates to the variance of attribute values.

2.3 SUMMARY OBSERVATIONS

All three categories of models (aggregate demand models, NMNL, and MLM) can be used to estimate changes in market shares and consumer surplus due to increases in vehicle prices and fuel economy. Aggregate demand models, like those developed by Kleit (2002a) or Austin and Dinan (2005), could, in principle, produce estimates for 60 or even 800 vehicle types. Given own- and cross-price elasticities, calibration of such models to sales data would be straightforward. Estimating the price elasticity matrix, however, is a major challenge. An 800 by 800 matrix would require 640,000 elasticity estimates and even a 60 by 60 matrix would need 3,600 elasticity values. Bordley's (1993) method offers a potential solution to this problem but it requires rarely available data on consumers' first and second choices. Perhaps this is why it appears not to have been used in subsequent studies.

The ability of mixed logit models to represent consumer heterogeneity also comes at the price of greater information requirements for model calibration and simulation. Mixed logit models require specification of not only the central tendencies of key parameters but also their variance, and possibly their correlations. Running a mixed logit model requires repeated randomized draws from the distributions of parameters. Fortunately, software is available for performing the necessary simulations. Calibration and updating of MLMs requires considerable effort. Survey based estimation methods require extensive, detailed survey data. Aggregate methods have more modest data requirements but the validity of the estimates by the most prevalent algorithms has been called into question by recent research (Knittel and Metaxoglou, 2008). In either case, there is presently no evidence that MLMs produce more accurate predictions than other methods. Should the EPA determine that vehicle choice modeling can make an important contribution to its regulatory analyses, it may be worthwhile to determine whether the potential benefits of using mixed logit models to represent consumer heterogeneity are worth the extra complexity and data requirements of the mixed logit model.

NMNL models have been constructed, calibrated and used in policy analyses of fuel economy issues by Greene et al. (2005), Harrison et al. (2008) and Bunch et al. (2011). All three applications modeled vehicle choices at a fine level of detail, ranging from 200 makes and models to over 800 make/model/engine/transmission combinations. This high level of detail was considered necessary to adequately represent the changes in market shares that might result from fuel economy and emissions standards or fiscal policies. Given that the price sensitivity of consumers' choices is greatest at the lowest level of the NMNL nest, i.e. when vehicles are the closest substitutes, modeling at the greatest feasible level of detail should produce a model with the potential to measure the full impacts of price and fuel economy changes on fleet average fuel economy and consumer surplus.

Given a nesting structure and corresponding price coefficients, NMNL models can be quickly and precisely calibrated to historical or projected sales data using closed form equations. NMNL models are capable of accommodating the introduction, termination, or modification of product lines. They are not capable, however, of predicting when product lines will be introduced or terminated. NMNL models that must be calibrated to sales data are also not able to predict the sales of newly introduced vehicles, since there is no vehicle-specific constant term available for new products. This is a general limitation of models that include fixed effects to accurately predict sales shares and applies to Mixed Logit Models and other formulations, as well.

For the purpose of developing an initial model to test the value of making such estimates, the NMNL method appears to be a good compromise between flexibility and simplicity. It can be readily calibrated with only a small amount of information about price elasticities and base year sales data. It allows for substantial flexibility in representing substitutions among vehicle types. On the other hand, it does not allow great flexibility in representing heterogeneous consumer preferences. This may be a fruitful area of future research and development, especially if it can be shown that more detailed representations of consumer tastes lead to more accurate predictions.

3. METHODOLOGY

This project constructs and calibrates a NMNL model along the line of Greene et al. (2005) and Bunch et al. (2011). Generalized cost coefficients are derived from the literature and NMNL properties. Given generalized cost coefficients, constant terms of the model are calibrated to baseline sales data such that the model prediction replicates baseline market share.

3.1 NESTING STRUCTURE

Choice alternatives in the CVCM are represented in detail, by make, model, engine and transmission, corresponding to the level of detail at which fuel economy measurements are made by the EPA. There are on the order of 1,000 choice alternatives. Individual vehicles are grouped into nests as in Figure 1 to allow differential substitution patterns within and between nests. The structure has 5 levels: Lev0 (Buy a new vehicle/Don't buy a new vehicle), Lev1 (Passenger Vehicles, Cargo Vehicles and Ultra Prestige vehicles), Lev2 (vehicle types: Two Seaters, Prestige Cars, Standard Cars, Prestige SUVs, MiniVans, Standard SUVs, Pickup Trucks, Vans, and Ultra Prestige Vehicles), Lev3 (vehicle classes (see Table 3) and Lev4 (vehicle configurations (one configuration is defined as a combination of make, model, engine size and transmission type)). Define Lev0 as the highest level and Lev4 as the lowest level. Right above Lev0 is root node (not drawn in Figure 1), which is the origin of the nesting structure/tree.
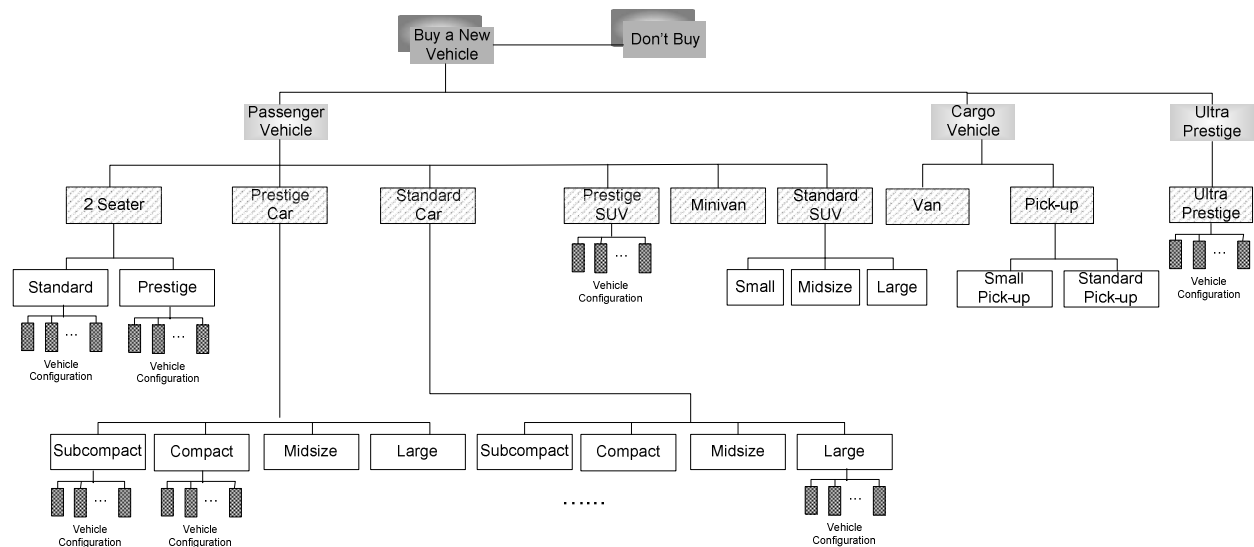
Figure 1 Nested Multinomial Logit Structure of Consumer Choice Model

Note: "Standard" is synonymous with "Non-Prestige"

The nesting structure in Figure 1 is defined according to general principles that group closer substitutes in a nest and ensure price sensitivity (price coefficient) and substitutability increase as one goes down to the bottom of the nesting structure[6]. The inclusion of the buy/no-buy option is necessary to predict impacts on total sales, not just the distribution of sales among makes,

---

[6] The requirement that price sensitivity increases as one goes down to the bottom is explained in Appendix A.

models and vehicle classes. Conditioning on buying a new vehicle, vehicle configurations are grouped according to functionality and size of vehicles and prestige/non-prestige. Thus lev1 distinguishes between passenger vehicles, cargo vehicles, and ultra-prestige vehicles (see its definition in Table 3), which are least substitutable. Lev2 further divides passenger vehicles into Two Seaters, Prestige Cars, Standard Cars, Prestige SUVs, Standard SUVs, and MiniVans, acknowledging increasing substitutability among these alternatives (e.g. Standard SUVs and MiniVans, which are both passenger vehicles, are closer substitutes than Standard SUVs and Small Pickup Trucks, because Small Pickup Trucks are cargo vehicles). Cargo vehicles are divided into Pickup Trucks and Vans. Lev3 continue dividing some nodes in lev2 by vehicle size or prestige/non-prestige.

The literature provides evidence that support our definition of nesting structure. A no-buy alternative is often included in previous studies (e.g., Berkovec, 1985; Berry, 1994; Berry et al., 1995; Goldberg, 1995; NERA, 2009). It is very common to segment vehicle market by vehicle size, functionality, and prestige/non-prestige (e.g. Lave and Train, 1979; Berkovec and Rust, 1985; Berkovec, 1985; Goldberg, 1995; Kleit, 2004; NERA, 2009). For example, Kleit (2004) classifies vehicles into small car, midsize car, large car, sports car, luxury car, small truck, large truck, small SUV, large SUV, minivan, and van, which is consistent with our class definition (Table 3). Moreover, our structure has advantages over other structures in the literature:

(1) It models vehicle market at a high level of detail, which enables the CVCM to potentially simulate the full range of sales mix shifts. The structure includes 5 levels, and choice alternatives are vehicle configurations (on the order of 1000), while the literature studies typically include two or three levels, and choice alternatives are vehicle size classes or makes/models (on the order of 200);

(2) The passenger and cargo vehicle distinction in Lev1 is fully compatible with EPA emissions standards' compliance categories for cars and trucks;

(3) Our structure has a more thorough treatment of prestige vehicles in consideration that they have different price sensitivities from non-prestige vehicles. In addition to grouping prestige two seaters, cars and SUVs into their own nests in Lev2 and Lev3, the structure also groups ultra-prestige vehicles into a nest in Lev1. The special treatment of ultra-prestige vehicles is to recognize that these vehicles have very distinct consumer demand and thus are hardly ever substitutes for other inexpensive vehicles. Technically speaking, positioning ultra-prestige vehicle nest in Lev1 allows us to assign a small price coefficient to these vehicles.

The structure in Figure 1 is implemented in the CVCM by default. Future versions of the CVCM could support user-defined structure. Alternative structures may have impacts on sales predictions. Sales in the level of vehicle configurations will be most sensitive to the structure change. The degree of sensitivity diminishes as the prediction is targeted at more aggregate levels.

Table 3 Vehicle Class Definition in the CVCM

CVCM Class	No. of Configurations[1]	Corresponding EPA Class
1. Prestige[2] Two-Seaters	27	Two Seaters
2. Prestige Subcompact Cars	49	Subcompact Cars, Minicompact Cars
3. Prestige Compact Cars and Small Station Wagons	71	Compact cars, Small Station Wagons
4. Prestige Midsize Cars and Station Wagons	66	Midsize Cars, Midsize Station Wagons
5. Prestige Large Cars	17	Large Cars
6. Two-Seater	26	Two Seaters
7. Subcompact Cars	58	Subcompact Cars, Minicompact Cars
8. Compact Cars and Small Station Wagons	82	Compact Cars, Small Station Wagons
9. Midsize Cars and Station Wagons	100	Midsize Cars, Midsize Station Wagons
10. Large Cars	29	Large Cars
11. Prestige SUVs	109	SUVs
12. Small[3] SUVs	17	SUVs
13. Midsize SUVs	72	SUVs
14. large SUVs	137	SUVs
15. MiniVans	19	MiniVans
16. Cargo/Large Passenger Vans	42	Cargo Vans, Passenger Vans
17. Small Pickup Trucks	49	Small Pickup Trucks
18. Standard Pickup Trucks	67	Standard Pickup Trucks
19. Ultra Prestige Vehicles[3]	93	See the definition (note 4) below

Notes:
(1) Number of configurations is the number of configurations which a CVCM class contains. It is not an attribute of the model itself, but specific to the vehicle data base to which the model is calibrated: a configuration is a record in the data base and a CVCM class consists of multiple records.
(2) Prestige and non-prestige classes are defined by vehicle price: the prestige are vehicles whose prices are higher than or equal to unweighted average price in the corresponding EPA class, and vice versa for non-prestige vehicles; these calculations are done after ultra-prestige vehicles (see below) are put in a separate nest. E.g., Prestige Two-Seater class is the set of relatively expensive vehicle configurations in EPA class of two seaters with prices higher than or equal to the unweighted average price of EPA two seaters.
(3) Non-prestige SUVs are divided into small, midsize and large SUVs by vehicle's footprint (small: footprint <43; midsize: 43<=footprint<46; large: footprint>=46)
(4) Ultra Prestige class is defined as the set of vehicles whose prices are higher than or equal to $75,000.

3.2 EQUATIONS

The CVCM includes a series of equations to define or calculate vehicle utilities, to calculate market share and sales of each vehicle configuration, and to estimate consumer surplus change brought by the installation of fuel economy technologies.

3.2.1 Prelude

We start from a review of Multinomial Logit (MNL) equations. The representative component of the utility expression for an alternative is defined in terms of four parts – the attributes x_k, attribute coefficients β_k, alternative specific constant α_j, and scale parameter μ. With the assumption that the variance of unobserved factors is distributed extreme value with variance $\dfrac{\pi^2}{6\mu^2}$ (Train, 2009), the utility of alternative j for individual n is

$$U_{nj} = V_{nj} + \varepsilon_{nj} = \alpha_j + \sum_r \beta_r x_{jr} + \varepsilon_{nj} = \alpha_j + \beta_p \sum_r \frac{\beta_r}{\beta_p} x_{jr} + \varepsilon_{nj} = \alpha_j + \beta_p G_j + \varepsilon_{nj}, \quad (15)$$

$$\mathrm{Var}(\varepsilon_{nj}) = \sigma^2 = \frac{\pi^2}{6\mu^2}, \quad (16)$$

where the sum G_j represents a "generalized cost" (Greene, 2001) for alternative j, β_p is the coefficient of vehicle price attribute and the scale parameter μ is proportional to the inverse of the standard deviation of the error term. The choice probability of alternative j is

$$P_{nj} = \frac{\exp \mu V_{nj}}{\sum_i \exp \mu V_{ni}} = \frac{\exp \mu(\alpha_j + \sum_r \beta_r x_{jr})}{\sum_i \exp \mu(\alpha_i + \sum_r \beta_r x_{ir})} = \frac{\exp \mu(\alpha_j + \beta_p G_j)}{\sum_i \exp \mu(\alpha_i + \beta_p G_i)}. \quad (17)$$

Note that the scale parameter μ and coefficients α_j and β_p are not separately identified and only the product of them can be estimated (Train, 2009). Thus in the CVCM, utility and choice probabilities have been expressed as

$$U_j = A_j + BG_j + \varepsilon_j, \quad (18)$$

$$P_j = \frac{\exp(A_j + BG_j)}{\sum_i \exp(A_i + BG_j)}. \quad (19)$$

with $A_j = \mu \alpha_j$ and $B = \beta_p \mu$. Coefficient B is called generalized cost coefficient since it reflects the derivative of utility with respect to price or generalized cost. Generalized cost coefficient is

proportional to scale parameter and thus inversely proportional to the standard deviation of error terms. Subscript n for individuals is omitted since the CVCM models the demand of a representative consumer.

3.2.2 Two-level CVCM Equations

The CVCMNMNL equations are first introduced in a simplified context with a two-level (vehicle configurations and vehicle classes) nested tree. Then full equations will be detailed in the next section.

The CVCM assumes that fuel economy and vehicle price are the only factors changing between model runs, and other attributes (e.g. performance and size) remain constant. This assumption is consistent with the current version of OMEGA which only predicts changes in fuel economy and vehicle prices. Other attributes can be included if the value of the attributes can be accurately quantified. The average value of unmeasured vehicle attributes is represented by an alternative-specific constant term. The constant for each alternative is calibrated to match baseline sales data. The utility[7] for vehicle j in class k is

$$U_{jk} = A_{jk} + B_k G_j + \varepsilon_{jk} = A_{jk} + B_k (C_{jk} - FS_{jk}) + \varepsilon_{jk}, \tag{20}$$

where

A_{jk}:   constant term for vehicle j in class k,
B_k: generalized cost coefficient parameter for vehicles in class k,
C_{jk} :   incremental cost for improving fuel economy of vehicle j, and
FS_{jk} :   the amount of fuel savings from improved fuel economy, valued by consumers when making purchase decisions.

The utility function for the class k is

$$U_k = A_k + \frac{B_{root}}{B_k} \ln \sum_{j \in k} \exp U_{jk} = A_k + \frac{B_{root}}{B_k} \ln \sum_{j \in k} \exp[A_{jk} + B_k (C_{jk} - FS_{jk})] \tag{21}$$

where A_k is constant term representing attributes shared by all alternatives in class k and B_{root} is the generalized cost coefficient for vehicle classes. Note that the log-sum term $\ln \sum_{j \in k} \exp U_{jk}$ is often referred to as the "inclusive value" in the literature (e.g. Train, 2009). Choice probability for alternative j is

$$P_j = P_{j|k} P_k, \tag{22}$$

---

[7]As seen in the appendix A, equation (20) only represents a component of the total utility that is unique to vehicle j. The utility component common to all vehicles in one class is captured by a class specific constant term.

with

$$P_{j|k} = \frac{\exp(A_j + B_k G_j)}{\sum_{j' \in k} \exp(A_{j'} + B_k G_{j'})} \tag{23}$$

and

$$P_k = \frac{\exp[A_k + \dfrac{B_{root}}{B_k} \ln \sum_{j \in k} \exp(A_j + B_k G_j)]}{\sum_{k'} \exp[A_{k'} + \dfrac{B_{root}}{B_{k'}} \ln \sum_{j \in k'} \exp(A_j + B_{k'} G_j)]}. \tag{24}$$

where $P_{j|k}$ is the conditional probability of choosing alternative j given that an alternative in class k is chosen, and P_k is the marginal probability of choosing an alternative in class k. Appendix A will show the equivalence of the NMNL specification here to more general formulations in the literature.

3.2.3 Full Scale CVCM Equations

We could list out NMNL equations for all the five levels. But a simpler alternative is to define utilities and calculate choice probabilities recursively based on the notations in Daly (2001). We reproduce the notation here for convenience:

- The tree function t(c) is used to define the nested logit structure: If c is a node in the tree, t(c) denotes the unique parent node at the higher level to which c is attached. For instance, Passenger Vehicle node is the parent of Standard Car node in Figure 1.
- The set ALL(c) denotes the set of nodes consisting of c and all its ancestors: ALL(c) = {c, t(c), t(t(c)), …, k| t(k) = root}
- Each node c can be considered an "alternative" in its own right. Nodes in the bottom level are "elemental alternatives", which are vehicle configurations. Nodes higher than the bottom level are viewed as "composite alternatives" that include all the elementary alternatives below it.
-

Utilities of nodes are then defined by

$$U_j = A_j + B_{t(j)} G_j \tag{25}$$

$$U_c = A_c + \frac{B_{t(c)}}{B_c} \ln \sum_{t(a)=c} \exp U_a = A_c + B_{t(c)} \overline{U}_c \tag{26}$$

with

$$\bar{U}_c = \frac{1}{B_c} \ln \sum_{t(a)=c} \exp U_a. \tag{27}$$

Equation (25) defines utilities for elementary alternatives including all vehicle configurations and No-Buy alternative. Equation (26) is recursive, defining the utility of node c as the summation of constant term A_c and generalized cost coefficient ($B_{t(c)}/B_c$) weighted log-sum term. The log-sum term is calculated over all the child nodes of c, where U_a is the utility of a child node a and again its utility can be expressed by equation (26).

In particular, the utility for root node (overall composite utility for the choice set) is

$$U_{root} = \frac{1}{B_{root}} \ln \sum_{t(a)=root} \exp U_a = \frac{1}{B_{root}} \ln(\exp U_{Buy} + \exp U_{NoBuy}) \tag{28}$$

with

$$U_{NoBuy} = A_{NoBuy}. \tag{29}$$

The utility function in equation (28) can be used to measure the consumer surplus change, consistent with Small and Rosen (1981):

$$\Delta CS = -\frac{1}{B_{root}} (\ln \sum_{t(a)=root} \exp U_a^1 - \ln \sum_{t(a)=root} \exp U_a^0) \tag{30}$$

where the superscripts 0 and 1 refer to before and after the change, and $-B_{root}$ is marginal utility of income.

The choice probability of each alternative is found by solving the following equation for P_c:

$$\ln P_c = \sum_{a \in ALL(c)} \{U_a - \ln \sum_{t(b)=t(a)} \exp U_b\}. \tag{31}$$

Specifically, the choice probabilities of bottom level elementary alternatives are calculated as the product of a series of probabilities:

$$P_j = P_{j|c(j)} P_{c|t(c)} ... P_{Buy|root}, \tag{32}$$

with

$$P_{c|t(c)} = \frac{\exp U_c}{\sum_{t(b)=t(c)} \exp U_b}, \tag{33}$$

where $P_{c|t(c)}$ is conditional probability of choosing node *c* given its parent node *t(c)* is chosen. In the CVCM, the market share of a vehicle segment is equivalent to the probability of choosing the corresponding node. Vehicle sales then equal the product of market size and market share:

$$N_c = M S_c = M P_c, \qquad\qquad (34)$$

where N_c is sales for the vehicle segment represented by node *c*, S_c is corresponding market share and *M* is market size, estimated by number of households.

The key input parameters for these equations include constant terms and generalized cost coefficients at each level of the nesting structure, change in vehicle price due to the installation of fuel economy technologies, and the value of fuel economy improvement perceived by consumers. The derivation of constant terms and generalized cost coefficients will be described in Section 3.4 on model calibration. Vehicle price change is assumed to be equal to increased vehicle cost, a direct output of the OMEGA. The assumption and calculation of consumer value of fuel economy will be presented in the next section.

3.3 VALUE OF FUEL ECONOMY

How consumers value fuel economy improvements has very significant implications for the costs and benefits of fuel economy and emissions policies. The accuracy of consumer choice models depends much on how close the assumption of value of fuel economy resembles the reality. However the literature has not achieved a consensus on this subject. On one hand, economically rational consumers would measure the value of fuel economy by the expected discounted present value of fuel saved over the full life of the vehicle. On the other hand, there is evidence that very few consumers actually make such quantitative assessments (Turrentine and Kurani, 2007). Greene et al. (2009) show that typical consumer loss aversion combined with the uncertainty of future fuel savings could lead to a significant undervaluing of future fuel savings relative to the expected present value. Greene (2010) concludes that econometric studies are nearly evenly divided about whether car buyers value fuel savings in accord with rational economic principles or significantly undervalue future fuel.

Reflecting this controversy, the National Research Council (2002) fuel economy study considered two alternative methods of estimating fuel savings valued by consumers, full lifetime discounted fuel savings and a 3-year simple payback. The OMEGA has calculated fuel savings as the payback from the first 5 years with 3% discount. In order to be consistent with the OMEGA, the CVCM implemented the same calculation method by default. However, users can always change the parameters (*r* and *L* in equation(35)) in the input file to reflect their own assumptions on fuel savings calculation.

Denote scenario 0 as the baseline scenario, with fuel economy at an initial value; denote scenario 1 as the policy scenario, where fuel economy changes over time in response to fuel economy and emissions policies. Define considered Fuel savings as fuel saved that the consumer takes into account in the vehicle purchase decision in policy scenario relative to baseline scenario:

$$FS_i(t) = \sum_{\tau=t+1}^{t+L} \frac{1}{(1+r)^{\tau-t}} P(\tau)M(\tau-t)\left[\frac{1}{\eta MPG_i^{\,0}(t)} - \frac{1}{\eta MPG_i^{\,1}(t)}\right] \tag{35}$$

where

$FS_i(t)$:   considered fuel savings of model year t vehicle i relative to its baseline configuration
$P(\tau)$:   price of fuel in year τ
$M(\tau-t)$:   annual miles traveled for a vehicle with age of $\tau-t$
r: consumer discount rate
η: OnRoad discount factor that discounts fuel economy (MPG) in order to reflect real-world driving conditions
L: assumed payback period, in years.

3.4 CALIBRATION

Generalized cost coefficients and alternative specific constant terms are key input parameters to the CVCM. Generalized cost coefficients can be directly assigned, as in NERA (2009), or can be derived from other measures, e.g. price elasticities, as in this CVCM. Constant terms represent baseline utilities before any changes to vehicles. It is necessary to calibrate constant terms such that the CVCM prediction replicates market shares in the baseline scenario.

3.4.1 Generalized Cost Coefficient Determination

3.4.1.1 Methods

Generalized cost coefficients in the CVCM have been determined based on multiple relationships and rules. Firstly generalized cost coefficients can be estimated from price elasticities according to the following relationship:

$$\eta_j = p_j B_c (1 - S_j), \tag{36}$$

$$B_c = \frac{\eta_j}{p_j(1 - S_j)} \approx \frac{\eta_c}{\overline{p}_c(1 - \overline{S}_c)} \tag{37}$$

where η_j is the own-price elasticity of demand for alternative j, p_j is the price of j, S_j is j's conditional market share given nest c is chosen, B_c is the generalized cost coefficient for alternatives in nest c, $\overline{p}_c$ is average price for alternatives in nest c, $\overline{S}_c$ is average conditional market share, and η_c is a representative value of η_js. Equation (36) is derived from the definition of elasticities and logit model equations (for further details, please refer to Train, 2009). Price elasticities can be chosen based on an evaluation of values found in the literature.

Secondly, theoretical requirement of NMNL on generalized cost coefficients provides useful information for determining generalized cost coefficients. The NMNL theory requires that the absolute value of generalized cost coefficients must increase as one goes down to the bottom

(vehicle configurations level) for the NMNL model to be consistent with utility maximization (see Appendix A):

$$|B_c| \geq |B_{t(c)}| \geq |B_{t(t(c))}| \geq ... \geq |B_{root}|, \tag{38}$$

where $B_{t(c)}$ is the generalized cost coefficient associated with the parent node of c. Thus generalized cost coefficients at bottom level provide upper bounds (in terms of absolute value) and generalized cost coefficient at the root node (choice to buy a new vehicle or not) provides a lower bound for all other generalized cost coefficients at intermediate nodes.

Thirdly, generalized cost coefficient of a nest has certain relationship with the price of that nest. We know that generalized cost coefficient is inversely proportional to the standard deviation of unobserved attributes in the nest (Appendix A). Prestige vehicle classes or nests may have large variance in unobserved attributes since consumers value these attributes very differently. Thus generalized cost coefficients of prestige vehicle classes or nests are lower in absolute value than those of non-prestige vehicle nests. This is consistent with the finding of Goldberg (1995) of lower generalized cost coefficients for higher-price market segments.

We further extend this relationship with evidence from empirical studies. Disaggregate vehicle type choice models (e.g. Train and Winston, 2007) typically include in the utility function the ratio of vehicle price (p_j) and household income (Y_n):

$$U_{jn} = \beta \frac{p_j}{Y_n} + ... + \varepsilon_{jn} = \frac{\beta}{Y_n} p_j + ... + \varepsilon_{jn}. \tag{39}$$

So generalized cost coefficient (β/Y_n here) is inversely proportional to income. Assuming that income elasticity of expenditure is 1 (i.e., expenditure on vehicle purchase is approximately proportional to income), we conclude that generalized cost coefficients are approximately inversely proportional to vehicle purchase expenditure and, roughly speaking, vehicle price.[8] That is,

$$\frac{B_c}{B_{c'}} \approx \frac{\bar{p}_{c'}}{\bar{p}_c}, \tag{40}$$

where c and c' represent two nests, and B and $\bar{p}$ are generalized cost coefficient and average price respectively. The CVCM models the choice of a representative consumer and cannot directly incorporate income difference at the household level. Equation (40) can act as a proxy to represent price sensitivity variation due to household income difference.

3.4.1.2 Calculation
The calculation of generalized cost coefficients according to Equation (37) requires the input of price elasticities. Table 4 4 has summarized elasticity values from relevant literature that study new vehicle demand and report elasticities explicitly. Although these literature elasticities are valuable, it is difficult to directly use them in the CVCM due to the following reasons: (1) literature studies and the CVCM have different nesting structures,[9] and (2) elasticities could be

---

[8] Our intention is not to derive a definitive relationship, but to obtain a rule of thumb from empirical observations, which would be useful to generalized cost coefficient calibration.

[9] Thus the categories presented in Table 4 do not correspond to the categories used in the nesting structure of the CVCM, but instead reflect the categories used in the cited studies.

quite different from one study to another depending on dataset and model assumptions. In view of these difficulties, one shall cautiously utilize these literature elasticities and also consider other constraints (equations (38) and (40)) to determine generalized cost coefficients. In the following sections, we will detail how generalized cost coefficients are calculated at each level of Table 5 by integrating all available information. Note that the choice levels in Table 4 do not exactly match those in Table 5. Roughly speaking, "Choice to Buy a New Vehicle or Not" in Table 4 corresponds to Level 0 of Table 5; "Choice of Market Segment" in Table 4 corresponds to Level 3 of Table 5; "Choice of Configurations" in Table 4 corresponds to Level 4 in Table 5. "Choice of Make/Model" in Table 4 has no direct corresponding level in Table 5 and is an intermediate level between level 3 and 4 of Table 5.

The overall price elasticity of automobile demand is set at -0.8 (Lev0 of Table 5), consistent with McCarthy (1996) and Levinsohn (1988). Following equation (37), the generalized cost coefficient for the buy/no buy decision is calculated and the value (-3.39E-05) serves as a **lower bound** (in absolute value) for all generalized cost coefficients in Table 5.

Not many studies report elasticities at vehicle configuration level (Level 4 of Table 5). So we first look at the make/model level, whose elasticities values are lower bounds[10] (in absolute value) of configuration level elasticities. Table 4 shows the average elasticity for choices among all individual makes and models is in the range of -2.3 to -4 (see "Choice of Make/Model" Section and Row "Average elasticity" in Table 4). Elasticities for choices among makes and models in each market segment vary, with the range of –3.3 to -4.7 for small, medium and large size segments, -1.2 to -3.7 for luxury vehicles and -1.2 to -4.2 for sport vehicles. Based on these estimates, we assume that price elasticities at make/model level are around -4 for non-luxury cars (-4 is about the central value of the literature estimates) and around -2 for luxury and sport cars (-2 is about the central estimate). Generalized cost coefficients (usually but not always ranked in the same way as elasticities) at vehicle configuration level (Level 4 of Table 5) shall be larger in absolute value than at make/model level. Therefore the representative value of elasticities is set at -5.0 at vehicle configuration level for non-prestige[11] cars (classes 6, 7, 8, 9, and 10 of Level 4 in Table 5) and -3.5 for prestige cars and two seaters (classes 1, 2, 3, 4, and 5 of Level 4 in Table 5. These values are within the range of literature estimates in Table 4 ("Choice of Configuration" section and studies of Berry et al., 1995 and EFTEC, 2008). Generalized cost coefficients for classes 1-10 are calculated according to equation (37) given elasticities are known. We don't have sufficient information to choose elasticities for other classes in Level 4 of Table 5 and will rely on equation (40) to calculate generalized cost coefficients. We select class 10 as the base class for non-prestige vehicles. Generalized cost coefficients of classes 12-18 are derived from class 10 generalized cost coefficient. Similarly for prestige vehicles, we select class 5 as the base class. Generalized cost coefficients of classes 11 and 19 are derived from class 5 generalized cost coefficient. Generalized cost coefficients in Level 4 serve as **upper bounds** (in absolute value) for other generalized cost coefficients in Table 5.

For elasticities at the vehicle class level (Level 3 of Table 5 and "Choice of Market Segment" of Table 4), Table 4 summarizes that own price elasticity is around -1.8 to -2.8 for small size

---

[10] Exactly speaking, generalized cost coefficients for choices among makes and models are lower bounds of generalized cost coefficients for choices among configurations. This relationship is approximately true for elasticities.

[11] Prestige cars in Table 3 have the same definition as luxury cars in Table 4.

segment, -1.3 to -3.5 for medium size segment, and -2.8 to -4.5 for large size segment. According to this observation, the representative value of price elasticities for choice among vehicle classes within Standard Car type is set at -3 (Row "Standard Car" in Level 3 of Table 5). Generalized cost coefficient is calculated according to equation (37). For luxury and sport cars, elasticities are around -1.7 to -3.5 in Table 4. Thus the representative value of price elasticities for choice among luxury/sport vehicle classes was initially set at -2.5 (Rows "Two Seater" and "Prestige Car" in Level 3 of Table 5), which is about the mean of literature estimates. However, calculated generalized cost coefficients based on these elasticities are larger in absolute values than their upper bounds and hence violate the theoretical requirement of NMNL (equation(38)). So price elasticities are adjusted to be -2.2 for Row "Prestige Car" and -1.3 for Row "Two Seater" so that[12] calculated generalized cost coefficients satisfy the constraint in (38). Again we are not trying to choose elasticity values for Prestige SUV, Standard SUV, and Pickup of Lev3 in Table 5. Instead, we selected Standard Car of Level 3 as the base. Generalized cost coefficients of Standard SUV and Pickup of Level 3 are derived from Standard Car generalized cost coefficient according to equation (40).

For Levels 2 and 1 of Table 5, we don't find relevant elasticity estimates in the literature. Thus the calibration of generalized cost coefficients is based on equations (38) and (40). First for choice of vehicle type within passenger category (Level 2-Passenger in Table 5), generalized cost coefficient is chosen to be -5.23e-5, which is the mean of its upper bound (Lev3-Two Seater: -7.08e-5) and lower bound (Lev0-RootNode: -3.38e-5). Then applying equation (40), generalized cost coefficient for choice of vehicle type within Cargo category (Lev2-Cargo) is calculated as -5.23e-5. Since there are no vehicle types within Ultra Prestige category, generalized cost coefficient of Lev2-Ultra Prestige is copied from Lev4-Ultra Prestige. For Lev1, generalized cost coefficient is simply set as the mean of its upper and lower bounds (-3.92e-5 and -3.38e-5 respectively).

So far all generalized cost coefficients are obtained from equation (37) or from equation (40) in the case that elasticities are not available, with equation (38) providing upper and lower bounds. On the other hand, unknown elasticities can be calculated once generalized cost coefficients are obtained. For convenience of implementing CVCM in C#, we simply provide C# program with all price elasticities in an input file (with some of the price elasticities back-calculated from the generalized cost coefficients as described above) and calculate all generalized cost coefficients using equation (37).

---

[12] A range of elasticities satisfy this constraint. We gradually increase the initial elasticity value of -3 by 0.1. The final value of -2.2 (-1.3 for the case of Two Seater) in the table is the first value in this process that satisfies the constraint.

Table 4 Own Price Elasticities of New Vehicle Demand in the Literature

		Own Price Elasticity of Demand	Values Used in Calibration
Choice of Configuration			
	Small	Berry et al. (1995): -6.4 for Mazda 323; Eftec (2008): -4.5	-5
	Midsize	Berry et al. (1995): -4.8 for Nissan Maxima; Eftec (2008): -5.4	-5
	Large	Berry et al. (1995): -4.8 for Honda Accord; Eftec (2008): -3.6	-5
	Luxury	Berry et al. (1995): -3.1 for Lexus LS400; Eftec (2008): -4.0	-3.5
	Sport	Eftec (2008): -1.6	-3.5 for two seaters
Choice of Make/Model	Average	Bordely (1993):-3.6; Goldberg (1995): -3.3; Goldberg (1998):-3.1;	
	Small	Bordley (1993): -3.4; Goldberg (1995): -3.5	-4
	Midsize	Bordley (1993): -3.3; Goldberg (1995): -4.6; Goldberg (1996,1998):-4	-4
	Large	Bordley (1993): -3.8; Goldberg (1995): -4.7; Goldberg (1996,1998):-4	-4
	Luxury	Bordley (1993): -3.7; Goldberg (1995): -2; Goldberg(1996):-1.2;	-2
	Sport	Bordley (1993): -4.2; Goldberg (1995): -1.4; Goldberg(1996,1998):-1.2	-2 for two seaters
	Truck	Goldberg (1995): -3.1□	
	Van	Goldberg (1995): -4.5□	
Choice of Market segment	Small	Bordley (1993):-1.9; Kleit (2002):-2.8; Cambridge (2008): -1.8	-3
	Midsize	Bordley (1993):-2.3; Kleit (2002):-3.5; Cambridge (2008): -1.3	-3
	Large	Bordley (1993):-3; Kleit (2002):-4.5; Cambridge (2008): -2.8	-3
	Luxury	Bordley (1993):-2.4; Kleit (2002):-1.7; Cambridge (2008): -3.5	-2.2
	Sport	Bordley (1993):-3.4; Kleit (2002):-2.3; Cambridge (2008): -1.8	-1.3 for two seaters
	Truck	Kleit (2002):-3 for small truck, -1.5 for large truck	
	SUV	Kleit (2002):-3 for small suv, -2 for large suv	
	Van	Kleit (2002):-2.4	
Choice to Buy a New Veh or Not		ranged from -0.8 to -1 Levinsohn(1988), Kleit (1990), McCarthy (1996,1998), Goldberg (1998)	-0.8
	Small	Berry et al. (1995): -6.4 for Mazda 323; Eftec (2008): -4.5	-5
	Midsize	Berry et al. (1995): -4.8 for Nissan Maxima; Eftec (2008): -5.4	-5
	Large	Berry et al. (1995): -4.8 for Honda Accord; Eftec (2008): -3.6	-5
	Luxury	Berry et al. (1995): -3.1 for Lexus LS400; Eftec (2008): -4.0	-3.5
	Sport	Eftec (2008): -1.6	-3.5 for two seaters

Table 5 Generalized Cost Coefficient Calibration

LEVEL 4	Choice of Make, Model, Engine Transmission Configuration within a Class						
Class	Name	Sales	Price	Share	No. Members	Ave. Share[1]	Elasticity	Slope
1	Prestige Two-Seater	79692	$50,888	0.47%	27	3.7%	-3.5	-7.14E-05
2	Prestige Subcompact	276351	$41,808	1.63%	49	2.0%	-3.5	-8.55E-05
3	Prestige Compact and Small Station	536024	$34,369	3.16%	71	1.4%	-3.5	-1.03E-04
4	Prestige Midsize Car and Station Wa	727577	$42,988	4.29%	66	1.5%	-3.5	-8.27E-05
5	Prestige Large	113968	$47,762	0.67%	17	5.9%	-3.5	-7.79E-05
6	Two-Seater	112099	$26,656	0.66%	26	3.8%	-3.5	-1.37E-04
7	Subcompact	1608947	$18,869	9.48%	58	1.7%	-5.0	-2.70E-04
8	Compact and Small Station Wagon	2392457	$17,901	14.10%	82	1.2%	-5.0	-2.83E-04
9	Midsize Car and Station Wagon	3180971	$21,132	18.75%	100	1.0%	-5.0	-2.39E-04
10	Large Car	752846	$24,217	4.44%	29	3.4%	-5.0	-2.14E-04
11	Prestige SUV	1011890	$46,765	5.96%	109	0.9%	-3.7	-7.95E-05
12	Small SUV	167691	$18,591	0.99%	17	5.9%	-4.9	-2.79E-04
13	Midsize SUV	1082846	$24,133	6.38%	72	1.4%	-5.1	-2.15E-04
14	Large SUV	2485225	$29,134	14.65%	137	0.7%	-5.1	-1.78E-04
15	Minivan	801143	$28,413	4.72%	19	5.3%	-4.9	-1.82E-04
16	Cargo / large passenger van	84530	$25,002	0.50%	42	2.4%	-5.1	-2.07E-04
17	Cargo Pickup Small	353636	$20,929	2.08%	49	2.0%	-5.1	-2.47E-04
18	Cargo Pickup Standard	984260	$28,444	5.80%	67	1.5%	-5.1	-1.82E-04
19	Ultra Prestige	214002	$94,930	1.26%	93	1.1%	-3.7	-3.92E-05
TOTAL[2]		16966155	$27,227	100.00%	1130			

LEVEL 3	Choice Among 19 Vehicle Classes within Vehicle Type						
Type	Name	Sales	Price[3]	Share	No. Members	Ave. Share	Elasticity	Slope
1	Two-Seater	191791	$36,725	1.13%	2	50.0%	-1.3	-7.08E-05
2	Prestige Car	1653920	$40,326	9.75%	4	25.0%	-2.2	-7.27E-05
3	Standard Car	7935221	$19,992	46.77%	4	25.0%	-3.0	-2.00E-04
4	Prestige SUV	1011890	$46,765	5.96%	1	100.0%	na	-7.95E-05
5	Standard SUV	3735762	$27,211	22.02%	3	33.3%	-2.7	-1.47E-04
6	Minivan	801143	$28,413	4.72%	1	100.0%	na	-1.82E-04
7	Cargo Van	84530	$25,002	0.50%	1	100.0%	na	-2.07E-04
8	Pickup	1337896	$26,457	7.89%	2	50.0%	-2.0	-1.51E-04
9	Ultra Prestige	214002	$94,930	1.26%	1	100.0%	na	-3.92E-05
TOTAL		16966155	$27,227	100.00%	18			

Level 2	Choice of Vehicle Type within Passenger or Cargo Categories						
Category	Name	Sales	Price	Share	No. Members	Ave. Share	Elasticity	Slope
1	Passenger	15329727	$26,362	90.35%	6	16.7%	-1.1	-5.23E-05
2	Cargo	1422426	$26,371	8.38%	2	50.0%	-0.7	-5.23E-05
3	Ultra Prestige	214002	$94,930	1.26%	1	100.0%	na	-3.92E-05

Level 1	Choice of Passenger,Cargo or Ultra Prestige Vehicle						
	Name	Sales	Price	Share	No. Members	Ave. Share	Elasticity	Slope
	Buy a new vehicle	16966155	$27,227	100.00%	3	33.3%	-0.7	-3.65E-05

Level 0	Choice to Buy a New Vehicle or Not						
		US HHs[4]	Price	Buy Share			Elasticity	Slope
	Root Node	129973385	$27,227	13.05%			-0.8	-3.38E-05

Note:
1)"Ave. Share" is the average of conditional shares of members in a nest. It is approximated by 1 over number of members
2)"Total" operation is not applicable to "price" column, which is sales weighted average price.
3) "Price" here reflects sales weighted average price.
4) "US HHs" is numer of households in the U.S. in base year

3.4.2 Constant Term Calibration

Given generalized cost coefficients, constant terms at each level of the nesting structure are calibrated to baseline sales data. Baseline market share and constants have the following relationship for any two vehicle configurations within the same vehicle class:

$$\frac{P_{i|k}^0}{P_{j|k}^0} = \frac{S_i^0}{S_j^0} = \frac{e^{A_{ik}}}{e^{A_{jk}}}, \forall i, j \in k \tag{41}$$

$$\Rightarrow A_{ik} - A_{jk} = \ln S_i^0 - \ln S_j^0, \forall i, j \in k$$

where superscript 0 represents baseline scenario, $P_{i|k}^0$ and $P_{j|k}^0$ are conditional probabilities of choosing vehicle i and j given class k has been chosen, and S_i^0 and S_j^0 are baseline market share of vehicles i and j. If we normalize one of the constants, e.g., A_{1k}, to be zero, then

$$A_{ik} = \ln S_{ik}^0 - \ln S_{1k}^0, \forall i \in k. \tag{42}$$

Vehicle class level constants can be derived from the following equation:

$$\frac{P_{k|h}^0}{P_{l|h}^0} = \frac{S_k^0}{S_l^0} = \frac{\exp[A_{kh} + \frac{B_h}{B_k} \ln(\sum_{i \in k} e^{A_{ik}})]}{\exp[A_{lh} + \frac{B_h}{B_l} \ln(\sum_{i \in l} e^{A_{il}})]}, \forall k, l \in nest\, h \tag{43}$$

where $P_{k|h}^0$ and $P_{l|h}^0$ are conditional probabilities of choosing vehicle class k and l given nest h has been chosen, S_k^0 and S_l^0 are base year market shares of vehicle classes k and l in the nest h, and A_{kh} and A_{lh} are class-specific constant terms. Normalizing the first class specific constant A_{1h} to be zero, we get

$$A_{kh} = \frac{B_h}{B_1} \ln(\sum_{i \in class1} e^{A_{i1}}) - \frac{B_h}{B_k} \ln(\sum_{i \in k} e^{A_{ik}}) + \ln S_k - \ln S_1, \forall k \in nest\, h \tag{44}$$

Again, we can use Daly (2001) notations to write a general equation. Denote c as a composite alternative and $t(c)$ as its parent. The following equation holds for any two composite alternatives in the same level of the nesting structure:

$$\frac{P_{c|t(c)}^0}{P_{b|t(b)}^0} = \frac{S_c^0}{S_b^0} = \frac{\exp[A_c + \frac{B_{t(c)}}{B_c} \ln(\sum_{t(a)=c} e^{U_a^0})]}{\exp[A_b + \frac{B_{t(b)}}{B_b} \ln(\sum_{t(a)=b} e^{U_a^0})]}, \forall c, b,\, \text{such that } t(c) = t(b) \tag{45}$$

where $P_{c|t(c)}^0$ and $P_{b|t(b)}^0$ are conditional probabilities of choosing alternatives c and b given their parent has been chosen, S_c^0 and S_b^0 are base year market shares of vehicle segments represented by c and b, and U_a^0 is baseline utility for an alternative a, as described by the recursive equation (26) with initial condition of

$$U_j^0 = A_j, \forall j \tag{46}$$

and

$$U_{NoBuy}^{0} = A_{NoBuy} = 0. \tag{47}$$

Constant terms can be solved from equation (45) given S_c^0, S_b^0, and generalized cost coefficients are known. The constant for the alternative of "not buying a new vehicle" is assumed to be 0.

4. IMPLEMENTATION AND USER GUIDE

The CVCM has been implemented in C# at the editor environment of Visual Studio 2010. The C# code reads input, calibrates the model parameters including constant terms and generalized cost coefficients, calculates utilities, choice probabilities, sales, and consumer surplus, and finally writes output to an Excel file. The C# code is distributed as a Windows installation file and users can install the program on the destination computers with Windows operating systems.[13]

4.1 USER INTERFACE

User interface of the CVCM is straightforward. The File menu has mainly two items: "Output Files to…" and "Open". The "Output Files to…" item specifies the folder of output files. The default output folder is CVCM installation folder\output. The "Open" item selects input file. Input files could be located in any folder. But by default, they are in CVCM installation folder\input. The CVCM installation program has copied example input files in the input folder. Each scenario has one input file, indicated by the file name. Select and open one input file to read in data. Then some of the data content will be displayed in the two tables of the user interface and users can check if the data are correctly read. The gray car button on the upper right corner will turn green. Users can click on the green button and run the program. After the run is finished, users can then select another input file to start another run.

The CVCM takes input on model parameters, vehicle characteristics in the baseline scenario (e.g. price, sales and fuel economy) and fuel economy improvement and associated incremental vehicle price in the policy scenario, where fuel economy changes over time in response to fuel economy and emissions policies. It then outputs predictions on sales and consumer surplus in the policy scenario.

All input and output are in Excel Files. As described in Section 4.1.2, the user can name the output file via the "GlobalParameter" sheet of the input file.

4.1.1 Input

A list of input data and data sources is as follows.

- *Vehicle database:* detailed database at vehicle configuration level. It includes vehicle identification information (e.g. make, model, and engine size), price, baseline sales, and baseline fuel economy.
- *Predictions on fuel economy and incremental price at vehicle configuration level:* key input data, commonly obtained from OMEGA output.
- *Generalized cost coefficients and alternative-specific constants:* model parameters. Generalized cost coefficients are derived from price elasticities and NMNL properties,

---

13 The current version of the CVCM does not work on non-windows operations systems (e.g. Mac and UNIX).

given elasticities are known from the literature. Constant terms are calibrated from baseline sales data and generalized cost coefficients.

- *Market size:* size of consumer market, which is typically approximated by the number of households. Household number projection for the United States is obtained from U.S. Census and American Community Survey.
- *Nesting structure:* default nesting structure is built in the model. In the future, users may be able to specify their own structure.
- *Fuel prices:* used to calculate fuel savings. Source: Annual Energy Outlook (AEO) 2010 from Energy Information Administration (EIA).
- *Annual and lifetime driving mileage for a typical car or truck:* used to calculate fuel cost and VMT weighted GHG emissions for manufacturers. Source: consistent with OMEGA assumptions.
- *Emission standards and vehicle footprint*: used to check manufacturer compliance with GHG emissions standards. This information is optional and requires linking the CVCM outputs to OMEGA. This linkage may be made available in future releases of the CVCM. At the current time, this field can be left blank.

Sample input files in the CVCM installation folder contain all the above information. Users should follow the format in these files to prepare their own input. An input file has 7 data sheets, listed as follows. The file also contains a sheet ("InputValidation") to validate the input. Click on "Validation Data" button in this sheet and error messages will prompt out if the input in the data sheets is not in right data type or within appropriate range. If the inputs fail the validation test, the implicit meaning is that the model nesting structure is not consistent with the parameters. An error message box will pop out and instruct users to check input files. Users are not able to run the model with invalid inputs.

4.1.1.1 Vehicle
Each row in this table contains attributes (see Table 6) of a vehicle configuration. CVCM classes are classified based on EPA classes, according to the relationship in Table 3. Users will need to provide data for the columns "predicted mpg" and "incremental price" based on OMEGA output or other sources.

Table 6 Format of Vehicle Sheet

veh id	manufacturer	nameplate	model	CVCM class	EPA class	fleet type	fuel type	footprint
baseline price	baseline mpg	baseline sales	predicted mpg	incremental price				

4.1.1.2 Manufacturer
It includes a list of manufacturer names, which must be consistent with column "manufacturer" in "vehicle" sheet.

4.1.1.3 Logit
This sheet lists price elasticities at each level of the nesting structure for the purpose of model calibration. Users can change the values of price elasticities, but not the nesting structure.

4.1.1.4 GlobalParameter
The structure of "GlobalParameter" sheet is as follows:

Table 7 Structure of "GlobalParameter" Sheet

Scenario Name	Payback Period	Discount rate	OnRoad/Tested MPG	Market Size

Scenario name defines the name of the output file. Payback period and discount rate are parameters for calculating the value of fuel economy improvement perceived by consumers. "OnRoad Discount" is used in fuel cost calculation to discounts EPA fuel economy (MPG) test value, which is displayed in fuel economy window stickers and used in the CVCM, to better reflect fuel economy under real-world driving conditions. Market Size data are used to calculate sales and calibrate logit model constants at the level of Buy/No-Buy.

4.1.1.5 Other Sheets

"VehicleUse" and "Fuel" sheets include parameters for calculating fuel cost. In "VehicleUse" sheet, Annual driving mileage of a car (truck) at certain age equals the product of VMT and survival rate. "Fuel" sheet simply records fuel prices with year 1 as the 1st year after redesign year.

"Target" sheet specifies footprint function parameters as in 2012-2016 EPA GHG emissions standards final rule (Table III.B.2–1 and Table III.B.2–2, page 25409 of EPA and NHTSA, 2010). The redesign year in the example input files is 2016. Thus parameters in this sheet reflect 2016 emissions standard.

4.1.2 Output

Each run will generate an output file, *with its name defined by the user* (at cell B2 of Global Parameter sheet of the Input file). An output file consists of two sheets: raw output and aggregate output.

"Raw Output" sheet first repeats the input data for the convenience of reading. The model output includes sales, market share, revenue (sales times the sum of vehicle price and incremental price), net price change (incremental price less fuel savings), and sales changes relative to the baseline scenario at the level of vehicle configuration.

"Aggregate Output" sheet outputs variables at more aggregate levels, including market-wide consumer surplus change, total sales, industry revenue, sales weighted average fuel economy and CO_2 emissions; manufacturer level sales, sales weighted average fuel economy and CO_2 emissions; sales at the level of passenger vehicle, cargo vehicle, or ultra-prestige[14], and sales at each vehicle class. Note that the fleet average fuel economy is calculated as a harmonic mean: with gpm_i as the gallons per mile for vehicle i,

$$FleetAvgFuelEcon = \frac{TotalSales}{\sum_i Sales_i * gpm_i}$$

Fleet average CO_2 values are calculated two ways: first, sales weighted:

---

[14] Passenger vehicle, cargo vehicle, and ultra-prestige distinction corresponds to level 1 of the nested choice structure.

$$SalesWtdCO2 = \frac{\sum_i Sales_i * CO2_i}{\sum_i Sales_i}$$

Secondly, it is also calculated as VMT-weighted, with the VMT based on the full lifetime undiscounted VMT of the vehicle; VMT differs by whether a vehicle is classified as a car or a truck for regulatory purposes:

$$VMTSalesWtdCO2 = \frac{\sum_i Sales_i * CO2_i * VMT}{\sum_i Sales_i * VMT_i}$$

4.2 INTERACTION WITH OMEGA

The CVCM could interact with the OMEGA at different degree. At this stage of model development, they run as two separate programs and pass information via excel files. In the future, the CVCM can be fully integrated into the OMEGA as one program.[15] The framework of interaction is as follows:

> Step 1: Run OMEGA model.
> Step 2: Collect data from OMEGA output and prepare input file for the CVCM.
> Step 3: Run CVCM
> > 3a: Calibrate CVCM using baseline sales data and price elasticities
> > 3b: Calculate sales, market share and consumer surplus
> > 3c: Output
> If the convergence criteria is met (i.e. emissions standards are complied), STOP here.
> Otherwise, Go back to step 1

---

[15] One possible way is to program the CVCM as a dynamic link library (dll) file and call this dll from the OMEGA.

REFERENCES

1. Allcott, H. and N. Wozny. 2009. "Gasoline Prices, Fuel Economy, and the Energy Paradox," unpublished manuscript, MIT Department of Economics, Cambridge, Massachusetts, November 16.
2. Anderson, S.P., A. De Palma and J.F. Thisse. 1988. "A Representative Consumer Theory of the Logit Model," *International Economic Review*, vol. 29, no. 3, pp. 461-466.
3. Austin, D. and T. Dinan. 2005. "Clearing the Air: The Costs and Consequences of Higher CAFE Standards and Increased Gasoline Taxes," *Journal of Environmental Economics and Management*, vol. 50, pp. 562-582.
4. Baltas, G. and P. Doyle. 2001. "Random Utility Models in Marketing Research: A Survey," *Journal of Business Research*, vol. 51, pp. 115-125.
5. Bento, A.M., L.H. Goulder, E. Henry, M.R. Jacobsen and R.H. von Haefen. 2009. "Distributional and Efficiency Impacts of Gasoline Taxes," *American Economic Review*, vol. 99, no. 3, pp. 667-699.
6. Bento, A.M., L.H. Goulder, E. Henry, M.R. Jacobsen and R.H. von Haefen. 2005. "Distributional and Efficiency Impacts of Gasoline Taxes: An Econometrically Based Multi-market Study," *AEA Papers and Proceedings*, vol. 95, no. 2, pp. 282-287.
7. Berkovec, James. 1985. "Forecasting Automobile Demand Using Disaggregate Choice Models," *Transportation Research B* 19B(4), pp. 315-329.
8. Berkovec, James and John Rust. 1985. "A Nested Logit Model of Automobile Holdings for One Vehicle Households," *Transportation Research B* 19B(4), pp. 275-285.
9. Berry, S.T. 1994. "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 25(2), PP. 242-262.
10. Berry, S., J. Levinsohn and A. Pakes. 1995. "Automobile Prices in Market Equilibrium," *Econometrica*, vol. 63, no. 4, pp. 841-890.
11. Bordley, R.F. 1993. "Estimating Automotive Elasticities from Segment Elasticities and First Choice/Second Choice Data," *The Review of Economics and Statistics*, vol. 75, no. 3, pp. 455-462.
12. Boyd, J. and R. Mellman. 1980. "The Effect of Fuel Economy Standards on the U.S. Automotive Market: An Hedonic Demand Analysis," *Transportation Research*, Vol. 14A, No. 5-6, pp. 367-378
13. Brownstone, D., D.S. Bunch, T.F. Golob and W. Ren. 1996. "A transactions choice model for forecasting demand for alternative-fuel vehicles", *Research in Transportation Economics,* Vol.4, pp. 87-129.
14. Brownstone, D., D.S. Bunch and K. Train. 2000. "Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-fuel Vehicles," *Transportation Research B*, pp. 315-338.
15. Bunch, D.S., D.L. Greene, T.E. Lipman and S. Shaheen. 2011. "Potential Design, Implementation, and Benefits of a Feebate Program for New Passenger Vehicles in California," State of California Air Resources Board and the California Environmental Protection Agency, Sacramento, California, available at http://76.12.4.249/artman2/uploads/1/Feebate_Program_for_New_Passenger_Vehicles_in_California.pdf.
16. Cambridge Econometrics. 2008. "Demand for Cars and their Attributes", final report for the UK Department of Transport, Cambridge, UK, 23 January.

17. Cardell, N.S. and F.C. Dunbar. 1980. "Measuring the Societal Costs of Downsizing," *Transportation Research A*, vol. 14A, pp. 423-434.
18. Carrasco, J. A. and J. D. Ortuzar. 2002."Review and Assessment of the Nested Logit Model," *Transport Reviews*, Vol. 22, No. 2, 197-218.
19. Daly, A. 2001. "Alternative tree logit models: comments on a paper of Koppelman and Wen." Transportation Research Part B-Methodological 35(8): 717-724.
20. Daly, A. J. and Zachary, S. 1978. *Improved multiple choice models.* In D. A. Hensher and M. Q. Dalvi (eds), Determinants of Travel Choice (Westmead: Saxon House), pp. 335-357
21. Donndenlinger, J.A. and H.E. Cook. 1997. *Methods for Analyzing the Values of Automobiles*, SAE Technical Paper 970762, Society of Automotive Engineers, Warrendale, Pennsylvania.
22. The Economics for the Environment Consultancy (EFTEC). 2008. *Demand for Cars and their Attributes*, Final Report, Economics for the Environment Consultancy, Ltd., London, UK, January.
23. (EPA/NHTSA) Environmental Protection Agency and National Highway Traffic Safety Administration. 2010. *Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards*, Federal Register, Vol. 75, no. 88, pp. 25324-25728.
24. Goldberg, P.K. 1998. "The Effects of the Corporate Average Fuel Efficiency Standards in the U.S." *The Journal of Industrial Economics*, vol. XVLI, no. 1, pp. 1-33.
25. Goldberg, P.K. 1996. *The Effects of the Corporate Average Fuel Efficiency Standards*, Working Paper 5673, NBER, Cambridge, Massachusetts, July.
26. Goldberg, P.K. 1995. "Product Differentiation and Oligopoly in International Markets: the Case of the U.S. Automobile Industry," *Econometrica*, vol. 63, no. 4, pp. 891-951.
27. Gramlich, J. 2008. "Gas Prices and Endogenous Product Selection in the U.S. Automobile Industry," manuscript, Department of Economics, Yale University, New Haven, Connecticut, November 20, 2008.
28. Greene, D.L. 2011. *Uncertainty, Loss Aversion and Markets for Energy Efficiency*, *Energy Economics*, vol. 33, pp. 608-616, 2011.
29. Greene, D.L. 2010. *How Consumers Value Fuel Economy: A Literature Review*, EPA-420-R-10-008, U.S. Environmental Protection Agency, Washington, D.C., March.
30. Greene, D.L. 2009. "Feebates, Footprints and Highway Safety," *Transportation Research D*, vol. 14, pp. 375-384.
31. Greene, D.L. 2001. *TAFV Alternative Fuels and Vehicles Choice Model Documentation*, ORNL/TM-2001/134, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
32. Greene, D.L. 1998. "Survey Evidence on the Importance of Fuel Availability to the Choice of Alternative Fuel Vehicles," *Energy Studies Review*, vol. 8, no. 3, pp. 215-231.
33. Greene, D.L. 1994. *Alternative Fuels and Vehicles Choice Model*, ORNL/TM-12738, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
34. Greene, D.L., J. German and M.A. Delucchi. 2009. *Fuel Economy: The Case for Market Failure.* in D. Sperling and J.S. Cannon, eds., Reducing Climate Impacts in the Transportation Sector, Springer Science and Business Media.
35. Greene, D.L. and J.T. Liu. 1988. "Automotive Fuel Economy Improvements and Consumer Surplus," *Transportation Research A*, vol. 22A, no. 3 pp. 203-218.

36. Greene, D.L., P.D. Patterson, M. Singh and J. Li. 2005. "Feebates, Rebates and Gas-guzzler Taxes: A Study of Incentives for Increased Fuel Economy," *Energy Policy*, vol. 33, no. 6, pp. 757-776.

37. Harrison, D., A. Nichols and B. Reddy. 2008. "Evaluation of NHTSA's Benefit-Cost Analysis of 2011-2015 CAFE Standards," National Economic Research Associates, Boston, Massachusetts, June 30. Available on the internet at http://www.nera.com/67_5346.htm .

38. Jacobsen, M.R. 2010. "Evaluating U.S. Fuel Economy Standards in a Model with Producer and Household Heterogeneity," manuscript, Department of Economics, University of California, San Diego, January.

39. Jaffe, A. and R. Stavins. 1994."The Energy-efficiency Gap - What Does it Mean?" *Energy Policy*, Vol. 22, no. 10, pp. 804-810.

40. Kleit, A.N. 2004. "Impacts of Long-Range Increases in the Fuel Economy Standards," *Economic Inquiry*, vol. 42, no. 2. Pp. 279-294.

41. Kleit, A.N. 2002a. "Impacts of Long-Range Increases in the Corporate Average Fuel Economy Standard," Working Paper 02-10, AEI-Brookings Joint Center for Regulatory Studies, Washington, D.C.

42. Kleit, A.N. 2002b. "Impacts of Long-Range Increases in the Corporate Average Fuel Economy Standard," The Pennsylvania State University, February 7, 2002. Available on the internet at http://www.heartland.org/custom/semod_policybot/pdf/11537.pdf .

43. Kleit, A.N. 1990. "The Effect of Annual Changes in Automobile Fuel Economy Standards," *Journal of Regulatory Economics*, 2:2 (June 1990) 151-172.

44. Klier, T. and J. Linn. 2008. "The Price of Gasoline and the Demand for Fuel Efficiency: Evidence from Monthly New Vehicle Sales Data," manuscript, Federal Reserve Bank of Chicago, Chicago, Illinois, September.

45. Knittel, C.R. and K. Metaxoglou. 2008. "Estimation of Random Coefficient Demand Models: Challenges, Difficulties and Warnings," Working paper 14080, National Bureau of Economic Research, Cambridge, Massachusetts, June.

46. Lave, C. and K. Train. 1979. "A Disaggregate Model of Auto-Type Choice," *Transportation Research*, Vol. 13A, no. 1, pp. 1-9.

47. Levinsohn, J. 1988. "Empirics of Taxes on Differentiated Products: The Case of Tariffs in the U.S. Automobile Industry," *NBER Chapters, in: Trade Policy Issues and Empirical Analysis*, pages 9-44. National Bureau of Economic Research, Inc.

48. Maddala, G.S. 1992. *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge, UK.

49. McCarthy, P.S. 1996. "Market Price and Income Elasticities of New Vehicle Demands," *The Review of Economics and Statistics*, vol. LXXVII, no. 3, pp. 543-547.

50. McCarthy P.S. and R.S. Tay. 1998. "New Vehicle Consumption and Fuel Efficiency: A Nested Logit Approach," *Transportation Research-E*, Vol. 34, No. 1, pp. 39-51.

51. McFadden, D. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior," pp. 105-142 in P. Zarembka, ed., *Frontiers in Econometrics*, Academic Press, New York.

52. Moon, H.R., M. Shum and M. Weidner. 2010. "Estimation of Random Coefficients Logit Demand Models with Interactive Fixed Effects," Department of Economics, University of Southern California, Los Angeles, May 27, 2010.

53. (NRC) National Research Council. 2002. "Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards," Report of the Committee, National Academy Press, Washington, D.C.

54. NERA Economic Consulting. 2009. "Evaluation of NHTSA's Benefit-Cost Analysis of 2011-2015 CAFE Standards."

55. Sawhill, J.W. 2008. "Are Capital and Operating Costs Weighed Equally in Durable Goods Purchases? Evidence from the U.S. Automobile Market," discussion paper, Department of Economics, University of California at Berkeley, Berkeley, California, April.

56. Small, K. A. and H. S. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." Econometrica 46(1): 105-130.

57. Train, K. 2009. *Discrete Choice Methods with Simulation*, Second Edition, Cambridge University Press, Massachusetts.

58. Train, K. 1993. *Qualitative Choice Analysis*, MIT Press, Cambridge, Massachusetts.

59. Train, K.E. and C. Winston. 2007. "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers," *International Economic Review*, vol. 48, no. 4, pp. 1469-1496.

60. Turrentine, T. and K. Kurani. 2007. "Car buyers and fuel economy?" *Energy Policy*, vol. 35, pp. 1213-1223.

61. Williams, H. C. W. L. 1977. "On the formation of travel demand models and economic evaluation measures of user benefit." *Environment and Planning*, 9A, 285-344.

APPENDIX A: DERIVATION OF NESTED LOGIT MODEL EQUATIONS AND RELEVANT PROPERTIES

The primary purpose of this appendix is to provide a general form of nested logit model equations and demonstrate CVCM equations as one specific instance of the general form. The secondary purpose is to derive conditions on structural parameters in nested logit models.

Without loss of generality, we consider a two-level nesting structure for the convenience of discussion, which is the most common case in the literature. Formulations for two-level models can be extended to multi-level cases.

The following formulation framework is consistent with William (1977), Daly and Zachary (1978) and Train (2009). Let j denote an elemental alternative in the nested tree and c the upper level composite alternative (or nest) to which j belongs. The utility of alternative j is

$$U_j = V_c + V_{j|c} + \varepsilon_j = V_c + V_{j|c} + \varepsilon_c + \varepsilon_{j|c}, \tag{A-1}$$

where the observed component of utility is decomposed into two parts: a part labeled V_c that is constant for all alternatives within the nest c, and a part labeled $V_{j|c}$ that varies over alternatives within the nest. The error term ε_j is also divided into two independent components of ε_c and $\varepsilon_{j|c}$. The following assumptions are made. Errors $\varepsilon_{j|c}$ are identically and independently distributed (iid) Gumbel with scale parameter μ_c. Errors ε_c are distributed such that total errors ε_j are distributed Gumbel with scale parameter μ_{root}.

We first see how the above definition and assumptions imply a relationship of scale parameters μ_c and μ_{root}. The variance of total errors is the sum of variance of two error components:

$$\mathrm{Var}(\varepsilon_j) = \mathrm{Var}(\varepsilon_c) + \mathrm{Var}(\varepsilon_{j|c}), \tag{A-2}$$

which, because the variance of the Gumbel distribution is $\dfrac{\pi^2}{6\mu}$ with μ as the scale parameter , can be expressed as

$$\frac{\pi^2}{6\mu_{root}} = \mathrm{Var}(\varepsilon_c) + \frac{\pi^2}{6\mu_c}. \tag{A-3}$$

Since $\mathrm{Var}(\varepsilon_c)$ is non-negative, the above implies the following *structural condition* (Williams, 1977):

$$\mu_{root} \le \mu_c. \tag{A-4}$$

Then we compare variations of choice probability expressions. Choice probability for alternative j (Carrasco and Ortuzar, 2002) is

$$P_j = P_{j|c} P_c, \qquad (A\text{-}5)$$

with

$$P_{j|c} = \frac{\exp(\mu_c V_{j|c})}{\sum_{j' \in c} \exp(\mu_c V_{j'|c})} \qquad (A\text{-}6)$$

and

$$P_c = \frac{\exp[\mu_{root} V_c + \frac{\mu_{root}}{\mu_c} \ln \sum_{j \in c} \exp(\mu_c V_{j|c})]}{\sum_{c'} \exp[\mu_{root} V_{c'} + \frac{\mu_{root}}{\mu_{c'}} \ln \sum_{j \in c'} \exp(\mu_{c'} V_{j|c'})]} = \frac{\exp[\mu_{root} V_c + \frac{\mu_{root}}{\mu_c} I_c]}{\sum_{c'} \exp[\mu_{root} V_{c'} + \frac{\mu_{root}}{\mu_{c'}} I_{c'}]}, \qquad (A\text{-}7)$$

where $P_{j|c}$ is the conditional probability of choosing alternative j given that an alternative in nest c is chosen, P_c is the marginal probability of choosing an alternative in nest c, and I is so called inclusive value of nest c.

Choice probability form in (A-5) to (A-7) is consistent with the one in Train (2009) (equations (4.4) and (4.5) of chapter 4) if the scale parameter μ_{root} is normalized to 1 and μ_c replaced by $\frac{1}{\lambda_k}$, the so-called log-sum coefficient. Next we wish to show that CVCM formulation is a specific instance of this general form. Assume utility function $V_{j|c}$ takes a specific form similar to the one in CVCM:

$$V_{j|c} = \alpha_j + \beta p_j \qquad (A\text{-}8)$$

where α_j is alternative specific constant and p is vehicle price or generalized cost. Then choice probability can be expressed as

$$P_{j|c} = \frac{\exp[\mu_c(\alpha_j + \beta p_j)]}{\sum_{j' \in c} \exp[\mu_c(\alpha_{j'} + \beta p_{j'})]} \qquad (A\text{-}9)$$

and

$$P_c = \frac{\exp[\mu_{root}Y_c + \dfrac{\mu_{root}}{\mu_c}\ln\sum_{j\in c}\exp(\mu_c(\alpha_j + \beta p_j))]}{\sum_{c'}\exp[\mu_{root}Y_{c'} + \dfrac{\mu_{root}}{\mu_{c'}}\ln\sum_{j\in c'}\exp(\mu_{c'}(\alpha_{j'} + \beta p_{j'}))]}. \tag{A-10}$$

On the other hand, choice probability in the CVCM is

$$P_{j|c} = \frac{\exp(A_j + B_c G_j)}{\sum_{j'\in c}\exp(A_{j'} + B_c G_{j'})} \tag{A-11}$$

and

$$P_c = \frac{\exp[A_c + \dfrac{B_{root}}{B_c}\ln\sum_{j\in c}\exp(A_j + B_c G_j)]}{\sum_{c'}\exp[A_{c'} + \dfrac{B_{root}}{B_{c'}}\ln\sum_{j\in c'}\exp(A_j + B_{c'} G_j)]}. \tag{A-12}$$

If one defines

$$A_j = \mu_c \alpha_j \tag{A-13}$$

$$A_c = \mu_{root} Y_c \tag{A-14}$$

$$B_c = \beta \mu_c \tag{A-15}$$

$$B_{root} = \beta \mu_{root} \tag{A-16}$$

and

$$p_j = G_j, \tag{A-18}$$

then CVCM equations are equivalent to the general form in (A-9) to (A-10). Parameters B_c and B_{root} are called generalized cost coefficients in the CVCM, since they reflect the derivative of utility with respect to price (or generalized cost). Equation (A-15) and (A-16) indicate that the absolute value of generalized cost coefficients are proportional to scale parameters and thus proportional to the inverse of standard deviation of random errors.

APPENDIX B: MODEL SENSITIVITY ANALYSIS

Among the model's parameters, the most important ones are price elasticities and those defining how consumers value fuel savings from fuel economy improvement. In this section, we examine the sensitivity of model results to variation of price elasticities and consumers' evaluation of fuel savings.

The input data (baseline sales and predicted fuel economy) are from an OMEGA run which simulates the scenario of light duty vehicles meeting 2016 EPA greenhouse gas emissions standards. The OMEGA results were provided by EPA in November, 2011. Sensitivity analysis results are specific to the data, which is why we use a realistic OMEGA data set. Since OMEGA tends to (but does not precisely) equalize the marginal cost of installing fuel economy technologies across all vehicles, we expect the impact of fuel economy improvement on sales mix to be small.

We will start by describing the distribution of price elasticities and then present sensitivity analysis results.

B.1 THE DISTRIBUTION OF OWN PRICE ELASTICITIES

Generalized cost coefficients are calibrated from assumed price elasticities (see Table 5). The elasticities represent the average elasticities of alternatives in a nest. Each vehicle has its own price elasticity depending on its price and market share as well as the generalized cost coefficient for its nest. Once the CVCM is calibrated, the actual elasticity of an alternative can be calculated based on equation (36). Alternatively, we can also get actual price elasticities through simulation. For example, the own price elasticity of a vehicle can be obtained by calculating the relative sales change of the vehicle in response to 1% change in its price. We calculated own price elasticities of 1130 vehicle configurations using equation (36).[16] The distribution of these individual elasticities is shown in Figure 2.

---

[16] Equation (36) is derived in the case of simple logit models. Thus elasticities calculated using this equation for nested logit models (e.g. CVCM here) are only an approximation to true values. However the error is very small.
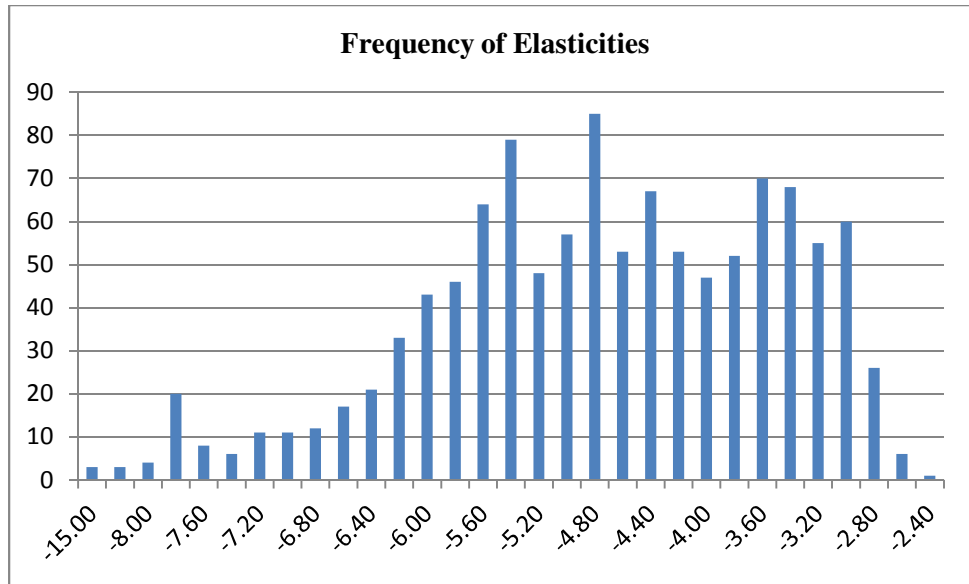
Figure 2 Distribution of Own Price Elasticities

Most individual vehicle price elasticities are in the range of -6 to -3. Only 10 vehicle configurations have price elasticities less than -8.0, as displayed in Table 8. Those vehicles are either ultra-prestige vehicles with very high prices or relatively expensive cars in their classes.

Table 8 List of Vehicles with Very High Elasticities (in absolute value)

manufacturer	nameplate	model	vehicle class	generalized cost coefficient	elasticity	price
Daimler	MERCEDES-BENZ	SLR	Ultra Prestige	-0.000038	-18.9	497750
BMW	ROLLS-ROYCE	PHANTOM	Ultra Prestige	-0.000038	-15.6	409000
BMW	ROLLS-ROYCE	PHANTOM EWB	Ultra Prestige	-0.000038	-15.4	405000
BMW	ROLLS-ROYCE	PHANTOM DROPHEAD COUPE	Ultra Prestige	-0.000038	-13.0	342000
VOLKSWAGEN	BENTLEY	AZURE	Ultra Prestige	-0.000038	-12.6	332585
VOLKSWAGEN	BENTLEY	ARNAGE RL	Ultra Prestige	-0.000038	-10.1	266585
VOLKSWAGEN	BENTLEY	ARNAGE	Ultra Prestige	-0.000038	-8.5	224585
Ford	FORD	MUSTANG	Subcompact	-0.000264	-8.3	31525
TOYOTA	LEXUS	IS 250	Subcompact	-0.000264	-8.3	31220
TOYOTA	LEXUS	IS 250	Subcompact	-0.000264	-8.2	31220
Mitsubishi	MITSUBISHI	ECLIPSE SPYDER	Subcompact	-0.000264	-8.0	30224
Mitsubishi	MITSUBISHI	ECLIPSE SPYDER	Subcompact	-0.000264	-8.0	30224
Ford	VOLVO	V50 FWD	Compact	-0.000289	-8.0	27560
Mazda	MAZDA	MAZDA RX-8	Subcompact	-0.000264	-8.0	30108

Table 9 provides additional descriptive statistics for elasticities of vehicle configurations within each class. In particular the medians of these elasticities are comparable to elasticity inputs shown in Table 5.

Table 9 Descriptive Statistics of Elasticities

Vehicle Classes	No. of vehicles	Min	1st Quartile	Median	3rd Quartile	Max	Std. deviation
Prestige Two-Seater	27	-5.5	-4.2	**-3.6**	-3.4	-3.2	0.6
Prestige Subcompact	54	-6.3	-4.6	**-3.6**	-3.2	-2.8	0.9
Prestige Compact and Small Station Wagon	71	-7.1	-4.3	**-3.7**	-3.4	-3.0	0.8
Prestige Midsize Car and Station Wagon	58	-6.1	-4.5	**-3.9**	-3.2	-2.8	0.9
Prestige Large	17	-5.4	-3.6	**-3.5**	-2.9	-2.7	0.8
Two-Seater	26	-5.9	-5.5	**-4.4**	-3.8	-2.0	1.1
Subcompact	76	-8.3	-7.0	**-6.0**	-5.1	-3.0	1.4
Compact and Small Station Wagon	82	-8.0	-6.9	**-5.4**	-4.6	-3.8	1.3
Midsize Car and Station Wagon	85	-7.8	-6.1	**-5.2**	-4.6	-3.3	1.2
Large Car	29	-6.6	-5.9	**-5.5**	-4.9	-4.3	0.6
Prestige SUV	108	-6.1	-4.1	**-3.6**	-3.2	-2.8	0.7
Small SUV	17	-7.3	-6.1	**-5.3**	-5.0	-4.5	0.8
Midsize SUV	78	-7.0	-5.7	**-5.1**	-4.7	-3.4	0.7
Large SUV	132	-6.5	-6.0	**-5.4**	-4.7	-2.9	0.8
Minivan	19	-6.5	-5.3	**-4.6**	-4.4	-3.7	0.8
Cargo / large passenger van	42	-6.2	-5.8	**-5.7**	-5.4	-4.7	0.4
Cargo Pickup Small	49	-6.7	-5.6	**-4.9**	-4.5	-3.8	0.8
Cargo Pickup Standard	67	-7.8	-5.6	**-5.3**	-4.6	-3.8	0.9
Ultra Prestige	93	-18.9	-4.8	**-3.6**	-3.2	-2.9	2.9

B.2 THE DISTRIBUTION OF CROSS PRICE ELASTICITIES

We have also obtained cross price elasticities at vehicle class level through simulation. The demand elasticity of class k with respect to the price of class l is calculated as the relative change in class k sales (total sales for all vehicles in class k) given 1% price change for all vehicles in class l. The cross elasticities are shown in Table 10, where own price elasticities are in bold text and large cross elasticities are in red. The values reported in Table 10 are comparable to those in Table 1 in the chapter of literature review.

Table 10 Price Elasticities at Vehicle Class Level

Class Name		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Prestige Two-Seater	1	-3.47	0.01	0.01	0.01	0.01	0.43	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00
Prestige Subcompact	2	0.02	-2.90	0.16	0.16	0.16	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.00	0.00	0.00	0.00
Prestige Compact and Small Station Wagon	3	0.02	0.24	-2.25	0.24	0.24	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01
Prestige Midsize Car and Station Wagon	4	0.04	0.40	0.40	-2.73	0.40	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.01	0.01	0.01	0.01
Prestige Large	5	0.01	0.08	0.08	0.08	-3.33	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00
Two-Seater	6	0.36	0.00	0.00	0.00	0.00	-1.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subcompact	7	0.05	0.05	0.05	0.05	0.05	0.05	-3.10	0.67	0.67	0.67	0.05	0.05	0.05	0.05	0.05	0.01	0.01	0.01	0.01
Compact and Small Station Wagon	8	0.06	0.06	0.06	0.06	0.06	0.06	0.72	-2.72	0.72	0.72	0.06	0.06	0.06	0.06	0.06	0.02	0.02	0.02	0.02
Midsize Car and Station Wagon	9	0.09	0.09	0.09	0.09	0.09	0.09	1.19	1.19	-3.01	1.19	0.09	0.09	0.09	0.09	0.09	0.03	0.03	0.03	0.03
Large Car	10	0.05	0.05	0.05	0.05	0.05	0.05	0.59	0.59	0.59	-4.24	0.05	0.05	0.05	0.05	0.05	0.01	0.01	0.01	0.01
Prestige SUV	11	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	-2.38	0.06	0.06	0.06	0.06	0.02	0.02	0.02	0.02
Small SUV	12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-2.67	0.09	0.09	0.00	0.00	0.00	0.00	0.00
Midsize SUV	13	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	1.15	-2.45	1.15	0.06	0.02	0.02	0.02	0.02
Large SUV	14	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	1.48	1.48	-3.02	0.08	0.02	0.02	0.02	0.02
Minivan	15	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	-1.52	0.01	0.01	0.01	0.01
Cargo / large passenger	16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1.25	0.05	0.05	0.00
Cargo Pickup Small	17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	-2.71	0.39	0.00
Cargo Pickup Standard	18	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.31	2.58	-1.60	0.02
Ultra Prestige	19	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-3.84

B.3 SENSITIVITY ANALYSIS

Eight cases are defined for the sensitivity analysis (see Table 10). Case 1 Baseline is the case to which other cases are compared. Cases 2-4d assume vehicles are required to improve fuel economy and vehicle prices are increased consequently. The fuel economy improvements and price increases of individual vehicle configurations are the same for all policy cases 2-4d, as provided by the OMEGA output dataset. Case 2 Reference uses default CVCM assumptions. It assumes consumers value the first 5 years of fuel savings using an annual discount rate of 3%. The default elasticities in Table 5 were used to calibrate the model. Varying the length of payback period and the discount rate generates policy cases 3a and 3b, and varying the elasticities generates policy cases 4a-4d. We examine the impact of these different assumptions on consumer surplus change, fleet average MPG and total sales.

Table 11 Sensitivity Analysis Results

Case ID	Name	Payback Period	Disc. Rate	Elasticities	Avg. Veh. Net Value ($)	Consumer Surplus Change ($/veh.)	Fleet MPG	Total Sales (millions)
1	Baseline	NA	NA	NA	0	0	27.35	16.65
2	Reference	5	0.03	default	1284	1227	33.67	17.29
3a	Lower value of fuel savings	**2**	**0**	default	23	17	33.66	16.66
3b	Higher value of fuel savings	**15**	**0.3**	default	4054	3591	33.65	18.64
4a	50% lower elasticities	5	0.03	default ***0.5**	1250	1222	33.7	16.96
4b	25% lower elasticities	5	0.03	default ***0.75**	1267	1225	33.69	17.12
4c	25% higher elasticities	5	0.03	default ***1.25**	1301	1230	33.65	17.45
4d	50% higher elasticities	5	0.03	default ***1.50**	1318	1232	33.64	17.62

The estimated consumer surplus change is highly sensitive to how consumers are assumed to value fuel savings from fuel economy improvements relative to the baseline case. The greater the amount of fuel savings that consumers consider when buying their vehicles, the larger the column "average vehicle net value" is. The vehicle net value is defined as the value of fuel savings taken into account minus vehicle price increase. Consumer surplus is largest in case 3b, where consumers are assumed to take into account fuel savings over the full expected lifetime of a vehicle. Price elasticities have much smaller impacts on consumer surplus. The higher the price elasticities, the larger the consumer surplus change is.

The impacts on total sales follow the same pattern as consumer surplus change. Total sales are most sensitive to the value of fuel savings perceived by consumers, with largest sales in case 3b. Total sales also increase when demand is more price elastic because the OMEGA data imply large gains in net values. The assumed price elasticity of new vehicle demand is -0.8.

Fleet average MPG is robust to all the variables varied in the sensitivity tests. In general, the differences are less than one tenth of a MPG among sensitivity test cases 2-4d. Fleet average MPG provided by OMEGA, if weighted by baseline sales, is 33.73, which is higher than all sensitivity test case fleet MPGs. This suggests the existence of a very small sales mix rebound effect: In the OMEGA data supplied, the greatest benefits of improved fuel economy (fuel savings minus vehicle price increase) tend to accrue to lower fuel economy vehicles. Thus, we see a sales mix shift towards lower fuel economy vehicles, as shown by Table 11. The table displays market shares in the baseline and reference cases for each baseline MPG decile. The share of lower fuel economy vehicles (decile 1 and 2) increases, while the share of higher fuel economy vehicles (decile 9 and 10) decreases.

Table 12 Market Shares by MPG Decile

MPG Decile	1	2	3	4	5	6	7	8	9	10
Baseline	5.1%	20.4%	31.2%	24.3%	13.7%	2.7%	0.6%	0.0%	0.3%	1.6%

| Reference | 5.8% | 21.2% | 30.4% | 24.7% | 13.4% | 2.4% | 0.6% | 0.0% | 0.2% | 1.3% |

The rebound effect can be quantified using the following equation

$$\varphi^s = \frac{MPG' - MPG^s}{MPG' - MPG^0},$$
(48)

where
S: index for a policy case
MPG^0: fleet MPG in the baseline case
MPG^s: fleet MPG in the policy case (average OMEGA MPG weighted by sales in the policy case)
MPG': fleet average OMEGA MPG weighted by the baseline sales.

If MPG^s is smaller than MPG', then there is a rebound effect. Table 12 indicates that the rebound effect is small, in the range of 1%, and higher price elasticities tend to magnify the rebound effect.

Table 13 Rebound Effect

Case ID	Name	Fleet MPG	Rebound Effect
2	Reference	33.67	0.9%
3a	Lower value of fuel savings	33.66	1.1%
3b	Higher value of fuel savings	33.65	1.3%
4a	50% lower elasticities	33.70	0.5%
4b	25% lower elasticities	33.69	0.6%
4c	25% higher elasticities	33.65	1.3%
4d	50% higher elasticities	33.64	1.4%

In summary, the sensitivity analysis results suggest that fleet MPG is robust to assumptions about price elasticities and the value of fuel economy perceived by consumers. Consumer surplus and total sales are very sensitive to perceived value of fuel economy and not very sensitive to variation in price elasticities at lower levels in the nesting structure.