Proceedings of the ASME 2013 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2013
August 4-7, 2013, Portland, Oregon USA

DETC2013-13372

SENSITIVITY OF VEHICLE MARKET SHARE PREDICTIONS TO ALTERNATIVE DISCRETE CHOICE MODEL SPECIFICATIONS

C. Grace Heckmann
Graduate Student
Carnegie Mellon University
Mechanical Engineering
Pittsburgh, Pennsylvania
15213 United States
(563) 580-7557
christine.grace.heckmann@gmail.com

Jeremy J. Michalek
Associate Professor
Carnegie Mellon University
Mechanical Engineering
Engineering and Public Policy
Pittsburgh, PA 15213
United States
(412) 268-3765
jmichalek@cmu.edu

W. Ross Morrow
Assistant Professor
Iowa State University
Mechanical Engineering
Ames, IA 50011
United States
(515) 294-4690
wrmorrow@iastate.edu

Yimin Liu
Technical Expert
Ford Motor Company
Dearborn, MI 48121
United States
(313) 322-3307
Yliu59@ford.com

ABSTRACT

When design decisions are informed by consumer choice models, uncertainty in the choice model and its share predictions creates uncertainty for the designer. We take a first step in investigating the variation in and accuracy of market share predictions by characterizing fit and forecast accuracy of multinomial logit, mixed logit, and nested logit models over a variety of utility function specifications for the US light duty new vehicle market. Using revealed preference data for years 2004-2006, we estimate a multinomial logit model for each combination of a chosen set of utility function covariates found in the literature. We then use each of the models to predict vehicle shares for the 2007 market and examine several metrics to measure fit and predictive accuracy. We find that the best models selected using any of the proposed metrics outperform random guessing yet retain substantial error in fit and prediction for individual vehicle models. For example, with no information (random guessing) 30% of share predictions are within 0.2% absolute share error in a market with an average share of ~0.4%, whereas for the best models 70% are within 0.2% (for the 2007 vehicle market this translates to an error of ~33,000 units sold). Share predictions are sensitive to the presence of utility covariates but less sensitive to the form. Models that perform well on one metric tend to perform well on the other metrics as well. In particular, models selected for best fit have comparable forecast error to those with the best forecasts, and residual error in model fit is a major source of forecast error.

1. INTRODUCTION AND MOTIVATIONS

Motivated by a call to base design decisions explicitly on predictions of their downstream consequences for the firm [1], researchers have proposed a variety of methods to predict the influence of design decisions on firm profit. The majority of these efforts apply discrete choice methods [2] to predict consumer choice as a function of product attributes and price, using choice predictions to guide or even optimize design decisions [3–11]. Such methods rely on the accuracy of choice predictions: uncertainty in choice predictions creates uncertainty about which designs are best [5,12]. Given the many sources of uncertainty in such models, Frischknecht et al. [8] question the suitability of using choice models in a design context.

Among the most popular product domains for application of choice modeling is automotive design. A national focus on alternative vehicle adoption has resulted in a variety of research reports that attempt to predict the vehicle market landscape as far as 40 years into the future. The variety of sources - equity research houses, consulting firms, academic institutions, and public sector government agencies - illustrates the interest in vehicle demand prediction. However despite these considerable research efforts, predictions for the market share of annual new vehicle purchases for several powertrains over the next 40 years vary wildly (see Appendix A). If such market models are employed in the context of new vehicle design, researchers must be aware of the degree of and implications of forecast uncertainty.

Prediction variation can arise from factors like dissimilar data sources and assumptions, and even within a single model prediction uncertainty may result from a multitude of factors including limited market data histories in the case of alternative vehicle powertrains, statistical estimation procedures, and model specification. In this paper we begin an investigation of predictive power and uncertainty in model predictions by focusing on utility function form and structural specifications (e.g. multinomial logit, mixed logit and nested logit).

Measuring the forecast accuracy of a particular model in a particular year may have idiosyncratic components. Forecast accuracy will depend on model specification as well as the particular factors affecting sales in the year of interest. A comprehensive study would examine a range of models representative of those used in the literature to model and predict choice, including models with alternative utility function specifications (including population heterogeneity and Bayesian hierarchical specifications), estimation and calibration procedures (e.g.: maximum likelihood, instrumental variables, or generalized method of moments), and data sources (aggregate sales data vs. individual-level data and stated vs. revealed choice data). We characterize share forecast accuracy of multinomial logit choice models in the automotive market with a variety of utility function specifications informed by the literature, each fit to aggregate market sales data from 2004-2006 and tested for accuracy of predicting share in the 2007 market. We apply several metrics to answer the following research questions:

1.) How widely do predictions vary based on utility function specification?
2.) How can we quantify share prediction error, and do different evaluation metrics lead to different selections of the "best predictive model"?

The second question is included as a necessary discussion in the course of answering the first. If we are to judge which utility or structural specification is the "best" at making predictions, we must have a clear definition of "best". We also compare the best multinomial logit models to mixed and nested logit models with similar utility function specifications, as well as to models estimated on alternative years and vehicle class sub-segments.

Section 2 contains an overview of the development and notable applications of choice models and their appearance in design literature. Section 3 describes our data set and detailed experimental setup so that a reader could replicate our work. It also discusses the metrics which we use to evaluate our models. Section 4 compares the results from estimating various utility and structural model specifications using the metrics described in the methodology section. Section 5 concludes with a summary of our findings, especially those relevant to designers, and Section 6 outlines some of the topics which require further investigation for a complete understanding of discrete choice model uncertainty.

2. LITERATURE REVIEW

Broadly, there are two schools of research in the vehicle demand literature. The first is concerned foremost with predicting future vehicle demand shares, usually at an aggregate level like vehicle class or powertrain type, and often without transparency about the assumptions and models used to make the forecast. We henceforth refer to this type of literature as "forecasting". The second school is interested in model construction and in vehicle and consumer attribute coefficient estimation especially as it pertains to willingness to pay and demand elasticity in past markets. We henceforth refer to this type of literature as "explanatory". Appendix B compares publications of each type. Engineering design in the decision based design and market systems areas requires models that can predict response to a new design. The forecasting literature is typically not used due to lack of transparency and documentation of data and modeling assumptions. Rather, models from the explanatory literature are applied in a predictive context.

Forecasting studies are conducted by private or government research entities or issued in report format from an academic research institute (see Appendix B). Reports are not generally peer reviewed and rarely contain a full mathematical description of the model, making it impossible to reproduce the model without additional information. Some reports include sensitivity cases formed with variations on model assumptions; for example, the EIA Annual Energy Outlook [13] contains base, low and high alternative vehicle future market share as a result of base, low and high future oil prices. This type of sensitivity only captures uncertainty about model input parameters and assumes that model specification and estimated coefficients are known. In practice, model specifications for choice contexts as complex as automotive purchases are always uncertain, and the relevant question is whether or not the model is sufficient for its intended function.

The bulk of the new vehicle purchase demand literature is explanatory, conducted by academic researchers and published in peer-reviewed academic journals (see Appendix B). This literature extensively discusses model estimation and to a lesser degree model selection, including potential sources of error from model misspecification. Usually researchers compare the goodness of fit across several specifications in order to determine which model best represents a known, current reality. However, most of this literature does not attempt to make predictions about future vehicle market-share penetration or evaluate models with predictive capabilities in mind (Frischknecht et al. [8] is a rare exception). In general, models that fit the existing data best may not necessarily be the best at predicting counterfactuals [14].

The earliest applications of economic models to predict overall automotive demand focused on macroeconomic variables and, as Train [15] highlights, only included price. These studies are referred to as aggregate studies because the level of granularity of predictions was at the whole market or vehicle class level as opposed to individual make-models. Disaggregate studies evolved to predict the number of vehicles an individual household would choose to own [15]. For example, Lave and Train [16] advanced this work by proposing a disaggregate model of vehicle class purchase choice based on consumer characteristics and additional vehicle characteristics, such as fuel economy, weight, size, number of seats, and horsepower. A wide variety of models followed over the next three decades.

We compare our model specifications to several well-known models from the automotive demand literature:

1.) Boyd and Mellman [17] propose incorporating a hedonic (random coefficient) demand function as an extension of simple logit models. In order to mitigate the independence of irrelevant alternatives (IIA) property [2], they argue that tastes for vehicle attributes could be allowed to vary over the population without explicit inclusion of consumer characteristics. Several studies in the design literature have adopted this model [11,18–21].

2.) Berry et al. [22] propose a canonical model informally referred to as the "BLP model" in the econometrics literature. It signaled a shift in the treatment of price endogeneity with the alternative-specific constant in revealed preference data. Related literature includes Nevo [23] which dissects the mathematical and computational details for researchers who wish to estimate BLP-type models. Dubé et al. [24], Knittel and Metaxoglou [25], and Skrainka and Judd [26] explore numerical estimation issues with structural models using the BLP model as an example.
3.) Brownstone and Train [27] propose several new vehicle purchase choice models using the results of a multi-part study conducted in Southern California and described in Bunch et al. [28]. Related literature includes Brownstone et al. [29] which follows up by combining revealed preference data with the stated preference data. McFadden and Train [30] prove that "Under mild regularity conditions, any discrete choice model derived from random utility maximization has choice probabilities that can be approximated as closely as one pleases by a [mixed multinomial logit] (MMNL) model," and uses a model from Brownstone and Train [27] as an illustration.
4.) Whitefoot and Skerlos [11] investigate the effect of fuel economy standards on vehicle size and employ a logit model with coefficients drawn from the literature.

It is important to note that we use the preceding literature models to inform comparison models of our creation; we do not recreate any of them exactly due to limited availability of data or specifics about estimation methodologies.

Other new-vehicle purchase models include [31–38]. Appendix C summarizes the covariates used in these models. These are the covariates used to inform the utility function combinations we test.

Other applications of discrete choice models for vehicle demand include predicting variations on vehicle demand in a variety of other contexts such as vehicle holdings (decision to buy or sell) and vehicle pricing [36,39–42], vehicle purchase financing options [43], and used-car and sequential choices [44]. Choo and Mokhtarian [45] propose a model based entirely on consumer usage and demographic variables. Zhang et al. [46] proposes an agent based modeling approach. Greene [47] conducts a literature review on consumer willingness to pay for fuel economy.

The engineering community has used discrete choice models broadly [3–7,20,48–51], especially in the context of product design optimization. Frischknecht et al. [8] specifically addresses their use in this context and find that the optimal design and resultant product profit are sensitive to the choice of model specification.

3. METHODOLOGY

Our overall goals are to examine the robustness of multinomial logit model predictions over various utility function specifications and to compare the predictions across the structural specifications of logit, mixed logit and nested logit. For conciseness we refer to the multinomial logit model as simply the "logit" model, whereas mixed and nested logit specifications are explicitly identified as such. We identify a universe of covariates informed by the literature and form combinations of them such that we have defined all possible linear utility function specifications from these covariates. (Note that we do not include an alternative specific constant and therefore do not use additional covariates as instruments to account for endogeneity). We then estimate the logit coefficients on US consumer vehicle purchase data from 2004-2006 and predict market shares for each of the vehicles in the US purchase data from 2007.

Using the metrics described in Section 3.3, we rank the predictive power of the models and compare their effectiveness across utility function specification for each of the metrics. We examine:

1.) What model would we have selected from 2004-2006 as the "best" based on goodness-of-fit if we didn't yet have the 2007 data? If we had selected that model, how well would we have done in predicting 2007?

2.) Which model estimated from 2004-2006 data predicts 2007 shares best based on metrics of forecast accuracy, regardless of how well it fits the estimation data?

3.) How well do models estimated on 2004-2006 data perform in predicting 2007 data, relative to models fit on 2007 data?

For a sub-set of utility specifications we determine to be "best" for logit models, we also estimate mixed and nested logit models. These models are then compared to the logit models on each of the metrics in order to evaluate the effect of structural specification.

3.1 The Data Set

Our data set draws vehicle attribute information from Ward's Automotive Index [52] and aggregate sales data from Polk [53] for vehicle sales during the 2004-2007 period. Other studies have used a variety of data sources (including these) as well as individual researcher conducted stated preference surveys. Our data set is revealed preference data exhaustively covering all new vehicle purchases in the United States. We use 2004-2006 data for estimation and 2007 for prediction because three years of data should be sufficient to predict a successive year.

Our models consider only new vehicle buyers, thus there is no outside good (option to not purchase any vehicle). Inclusion of an outside good allows a choice model to endogenously determine market size. Excluding it models only share among the vehicles purchased. We ignore the outside good here because it is difficult to define the correct market [22] and if the outside good is included, then further uncertainty in prediction arises from the need to account for macroeconomic conditions. For example, sales volume typically drops during a recession. As part of a model to predict future US vehicle stock, Greenspan and Cohen [54] include variables such as "household formations, the unemployment rate, and the prices of new vehicles, repairs and gasoline" to estimate new vehicle sales. Nevertheless, our model is not insulated entirely from macroeconomic issues. For example, if the overall US economy declines in a future year, consumers may be more price-sensitive and opt to purchase less expensive cars, which affects share predictions through the model's price coefficient. We are aware that there are many factors that drive share which are not included in our models, but we are interested in how well a modeler can predict when relying primarily on available vehicle attribute data.

Within the intent of addressing research question one, "How widely can predictions vary based on how covariates enter the utility function or if they are omitted entirely?", we begin by surveying the automotive demand model literature to determine the universe of independent variables historically used in the discrete choice models. While most of the automotive demand literature compares their respective results in the context of a few alternate utility function specifications, to our knowledge there is no research conducting a combinatorial search over possible utility covariates.

In general there are an infinite number of covariates that could be included in a utility function. We only consider new light duty vehicle purchases and restrict our attention to covariates identified in the automotive demand literature over the last fifty years. The universal set of possible utility covariates generated by compiling a list of covariates from literature is a realistic, finite set. We assume that the utility function is linear in parameters, which is a standard assumption in almost all logit models because it ensures that the log-likelihood function is concave [2].

Appendix C summarizes the new car purchase automotive demand models. From this list of covariate possibilities, we choose which subset should be tested in a comprehensive but manageable search. Many of the models include demographic or consumer usage covariates, but because Ward's Automotive Index data [52] does not include individual-level choices we ignore demographics. For some of the demographic information like gender or income an aggregate distribution over the US population is available, but because we do not know which consumers selected which vehicles, sampled consumer attributes are unlikely to accurately determine specific individuals' sensitivity to vehicle attributes. There are vehicle attributes we omit because they are not available in our data sources:

- *Indirect vehicle attributes* like consumer reports ratings for handling and safety – These would be unknown at the time of prediction.
- *Vehicle and battery maintenance costs*- These covariates are used primarily when predicting alternative vehicle share and they will not vary substantially across conventional and hybrid powertrains.
- *Acceleration time (seconds)*- We indirectly test inclusion of acceleration through functions of horsepower and weight. Note that horsepower/weight correlates well with 0-60 second acceleration time for cars well but poorly for trucks.
- *Range*- This covariate is used primarily when predicting alternative vehicle share and will not vary substantially across conventional and hybrid powertrains. A related fuel economy covariate is included.
- *Top speed*- We use an alternative measure of performance through horsepower and weight.
- *Number of seats*- We use vehicle class, which is closely related to seating.
- *2-year retained value*- Like the consumer rating data this would not be known at the time of prediction.
- *All alternative-vehicle specific attributes* (e.g. dummies for hybrid or electric power trains)- These are not relevant to our data set which includes conventional vehicles and only a limited number of hybrid power trains

As discussed in the introduction, predicting alternative vehicle share is the focus of much current research. Due to the restrictions imposed by our data set we do not predict their shares in this study, though potential future work would include investigating how the findings of this study transfer to prediction of vehicles with limited market histories. We would expect that the forecasts would typically be less accurate than for conventional vehicles because consumers have less developed opinions of alternative vehicle technologies [55–58].

The highlighted covariates in Appendix C are those which remain after omitting demographic, usage, indirect, and unavailable attributes. Some studies group price and fuel economy variables into discrete levels of each rather than treating them as continuous variables. We consider all covariates (except for class and brand dummies) to be continuous variables because, unlike controlled conjoint experiments, the market data do not fit into a small number of discrete levels. Price is always included as a covariate and can take either of the forms listed in Table 1; vehicle class dummies are also always included. The other highlighted covariates in Appendix C can take one of the forms listed in Table 1 or can be excluded entirely. A constant is necessarily excluded because it does not vary over the data when an outside good is not included. Given these covariate options, there are 9,000 possible utility specifications for the logit model outlined in Table 1. Note that the option "not included" means that the covariate is excluded from the utility function entirely.

There is correlation in many of these covariates, Such correlations can induce bias in the estimated coefficients if not corrected for [59]. However, while this presents difficulties in drawing inferences from the coefficients (e.g. willingness-to-pay) it does not affect the ability to make predictions from the model so long as the correlations in the training data would also be present in the prediction set. For vehicle markets, this is very likely to hold for near-term predictions at least.

Operating cost includes the macroeconomic variable of retail gas price. Though we aim to exclude non-vehicle attributes, this covariate was so prevalent in the literature that we thought it important to include for comparison.

TABLE 1- COVARIATE FORMS TESTED IN UTILITY FUNCTION SPECIFICATIONS

		Functional form options			
Covariate	**Option 1**	**Option 2**	**Option 3**	**Option 4**	**Option 5**
Price		price ($)	price + op cost	ln(price)	
Operating cost[1]	not included	fuel cost/mile	miles/fuel cost	miles/gallon	gallons/mile
Acceleration[2]	not included	horsepower/weight (hp/wt)	wt/hp	$\exp(c_1*(hp/wt)\text{^}c_2)$	hp
Size	not included	length	width	length-width	length*width
Style	not included	(length*width)/height			
Luxury	not included	dummy if air-conditioning is standard			
Transmission	not included	dummy if auto. transmission is standard			
Manufacturer[3,4]	not included	dummy for country of origin	dummy for brand		
Vehicle class[5]		dummies for vehicle class			

[1] Fuel cost is average annual gas price [60] in 2004 dollars, adjustment based on the Consumer Price Index [61]

[2] c1= -0.00275 and c2= -0.776 as in the EIA Annual Energy Outlook 2011 [13]

[3] Country of origin includes: United States, Europe, and Asia; excludes United States dummy for identification

[4] Brand includes: Acura, Audi, BMW, Buick, Cadillac, Chevrolet, Chrysler, Dodge, Ford, GMC, Honda, Hummer, Hyundai, Infiniti, Isuzu, Jaguar, Jeep, Kia, Land Rover, Lexus, Lincoln, Mazda, Mercedes, Mercury, Mitsubishi, Nissan, Oldsmobile, Pontiac, Porsche, Saab, Saturn, Scion, Subaru, Suzuki, Toyota, Volkswagen, Volvo; excludes Acura dummy for identification

[5] Class includes: Compact, midsize sedan, full size sedan, luxury sedan, SUV, luxury SUV, pickup, minivan, van, and sports; van is excluded for identification

.

3.2 Model Estimation

As discussed previously our utility function is linear in parameters:

$$\boldsymbol{u_{ij}} = \mathbf{x}_j'\boldsymbol{\beta} + \boldsymbol{\varepsilon_{ij}} \tag{1}$$

where u_{ij} is the utility of vehicle j for consumer i, $\mathbf{x}_j$ is the attribute vector of vehicle j, $\boldsymbol{\beta}$ is the vector of model parameters to be estimated, and ε_{ij} is an error term. Following standard assumptions, if ε_{ij} is independently identically distributed (iid) and follows a type I extreme value distribution, then the probability P_j that a consumer will choose vehicle j can be expressed as:

$$P_j = \frac{\exp(\mathbf{x}_j'\boldsymbol{\beta})}{\sum_{k=1}^{N}\exp(\mathbf{x}_k'\boldsymbol{\beta})} \tag{2}$$

where N is the number of vehicles in the estimation data set. This is the (multinomial) logit formula. The likelihood of the estimated parameters L is defined as the probability of generating the observed data given the estimated parameter values:

$$L(\hat{\boldsymbol{\beta}}|\mathbf{x}) = \prod_{k=1}^{N} {P_k}^{s_k} \tag{3}$$

where s_k is the sales of vehicle k. The maximum likelihood estimator of the parameters $\hat{\boldsymbol{\beta}}$ is the value of the vector which maximizes L. The monotonic transformation $\ln(L)$ is typically used as the objective function for convenience. For more detail on logit models and their estimation see Train [2].

The mixed logit, or random coefficients logit, model is similar to the logit model except the individual β's are allowed to vary over the population to represent heterogeneous consumer preferences. In our case we assume that they are independently normally distributed:

$$\boldsymbol{\beta} \sim N(\boldsymbol{\mu}, \Sigma) \tag{4}$$

where Σ is a diagonal matrix, and maximum likelihood estimates the elements of $\boldsymbol{\mu}$ and Σ using numerical integration [2]. This specification relaxes the IIA restriction.

Our nested logit specification divides the vehicles into groups or nests by vehicle class and fits a logit model to each of the classes. We assume that the utility functional form is the same for each nest, but in this specification each nest can have a different coefficient value for a given attribute. For example, the β for price will be different for midsize cars than it is for pickups. However, within a nest β is fixed. A nested logit exhibits the IIA property for products within a nest, but relaxes the IIA restriction for products in different nests.

3.3 Evaluation Metrics

After fitting each of the model specifications, we evaluate prediction error using the Kullback-Leibler divergence (KL) [62], the equivalent average likelihood (EAL) a cumulative distribution of error tolerance (CDFET), and the average share error (ASE), and we compare the goodness-of-fit across the Akaike Information Criterion (AIC) [63] and the Bayesian Information Criterion (BIC) [64]. Each of these metrics is described below. It should be noted that the KL divergence, EAL, CDFET, and ASE could also be used as goodness-of-fit metrics, but we use them only in the context of evaluating predictions, which is our primary focus. While the AIC and BIC could be used as goodness-of-prediction metrics, they were not designed for this purpose nor have the authors ever seen them used in this context. We compare models selected as best by these metrics to one another and to literature informed benchmark models.

Goodness-of-prediction metrics

Kullback-Leibler divergence (KL):

$$\boldsymbol{KL}(\mathbf{P}||\mathbf{Q}) = \sum_{\boldsymbol{i=1:N}} \boldsymbol{ln}\left(\frac{\boldsymbol{P_i}}{\boldsymbol{Q_i}}\right)\boldsymbol{P_i} \tag{5}$$

where N is the number of outcomes, P_i is the observed probability of realizing outcome i and Q_i is the predicted probability of outcome i. The Kullback-Leibler distance measures the difference between a predicted distribution and the true distribution. It represents the information lost when **Q** is used to approximate **P**, and it is a non-symmetric measure (the distance from **P** to **Q** is not the same as the distance from **Q** to **P**) [65]. It can take on any positive real value, and lower values of the metric- meaning there is less divergence between distributions - are preferred. If the hypothesized model is identical to the true model, then the Kullback-Leibler distance is zero [66]. It is the measure most closely related to the maximum likelihood because the value of the parameters which maximize the likelihood will tend towards the minimizer of the KL distance [66].

Equivalent average likelihood (EAL):

$$\boldsymbol{EAL = L^{1/N}} \tag{6}$$

where L is the likelihood and N is the number of outcomes. This metric "normalizes" likelihood to a single choice in the data by assessing the common likelihood one would have to achieve for each data point in order to generate the same net likelihood as the model. The EAL can take on real values between 0 and 1, and larger values are preferred.

Note that the KL and EAL metrics will have different values for the same model, but they will rank a set of models identically (see Appendix D for proof). We provide both values in comparison tables so that the reader can get a sense of the intuitive meanings of both metrics, but the "KL-best-model" is always identical to the "EAL-best-model".

Error tolerance CDF (CDFET): The absolute error between the actual share values and the predicted share values is calculated for each model, and the cumulative distribution function (CDF) of error is plotted. This can be used to compare models on the metric "If I'm willing to accept a share prediction error of 0.02%, for example, which model has the most predictions with error less than this amount", or inversely "what tolerance band must I be comfortable with in order to capture 25% of the predictions the model made". This metric evaluates a model in terms of various risk tolerance levels. We use absolute error as opposed to relative error because relative error overemphasizes prediction errors for vehicles with small market shares.

Average share error

$$\boldsymbol{ASE = \frac{1}{N}\sum_{i=1}^{N}\left|s_{i,actual} - s_{i,predicted}\right|} \tag{7}$$

where N is the number of outcomes, s_{actual} is the observed share, and $s_{predicted}$ is the share predicted by the model. We report this as a summary statistic in the model comparison table but do not use it as a basis for "best model" selection because it does not holistically capture distribution divergence. It will not distinguish between a model with large error for one vehicle alternative vs. the same degree of error spread out among many vehicle alternatives.

Goodness-of-fit metrics

Akaike information criterion (AIC):

$$\boldsymbol{AIC = 2ln(L) - 2k} \tag{8}$$

where L is the likelihood and k is the number of model parameters. AIC can take on the value of any negative real number, has no standalone meaning, and is only useful as compared to the AIC of other candidate models fit to the same data set. Greater values of AIC are preferred.

AIC is a variation on likelihood but penalizes overfitting. When more parameters are included in a function, the likelihood necessarily increases so that selecting a model based on likelihood will simply yield the model with the most parameters. The AIC rewards inclusion of variables that improve likelihood significantly, but it penalizes inclusion of extra variables that offer less information. Though this is not a guarantee that there is no overfitting, it helps to avoid selecting models which fit the data well but predict poorly [66].

The AIC is inconsistent for large sample sizes because the log-likelihood will increase linearly with the number of observations while the penalty term is proportional to the number of parameters. Finite sample corrected AIC formulas do exist, but there is no proof that they apply to general likelihood models as opposed to linear regression and autoregressive models [66].

Bayesian information criterion (BIC):

$$\boldsymbol{BIC = 2ln(L) - ln(N) * k} \tag{9}$$

where L is the likelihood, N is the number of outcomes, and k is the number of model parameters. The BIC is similar to the AIC but with a stronger penalty for increasing numbers of covariates. The BIC corrects the inconsistency in the AIC, but the AIC is more efficient. Greater values of BIC are preferred [66]. Derivations and consistency proofs for the KL, AIC and BIC metrics can be found in [66].

4. RESULTS

Of the 9,000 tested utility function specifications, for 8,927 (99%) Matlab's fminunc algorithm converged to likelihood-maximizing coefficients. 72 returned an exit flag indicating possible convergence and 1 failed to converge. Only the 8,927 models which successfully converged were considered as candidate models. The candidate models were ranked from best to worst on each of the metric dimensions discussed in the evaluation metrics section. There were no models with identical values of any metric (no ties). In the following results "best models" refer to the models ranked as number one for a given metric.

4.1 Metric Comparison

The traditional goodness-of-fit metrics- LL/KL/EAL and AIC, BIC - select the same "best" model when evaluated on the estimation data set. Furthermore they select the same model as "best" when applied to the prediction data set, though the best estimation model and best prediction model are different. The goodness-of-prediction CDFET metric selects distinct models for the best model at the error tolerance bar levels of 25%, 50% and 75%. These three models are also distinct from the best estimative and predictive models under the AIC, BIC, and KL/EAL criteria.

Figure 1 is the CDF of the error tolerance for the best models. The x-axis is the absolute difference between the predicted share and the actual share and the y-axis is the proportion of make-model share predictions which have an error less than the corresponding value on the x-axis. For example, in Figure 1 the point (0.25%, 75%) indicates that 75% of the share predictions made by the model deviate from the observed share by less than 0.25%. The "no info" line is calculated by assigning an equal share to all vehicles and represents the case where no model is used to inform predictions. The class line is the model which includes only class dummies. The static line is a model which assumes that 2007 make-model shares were held constant from their 2006 values and all vehicles introduced in 2007 received an equal proportion of the remaining market. From Figure 1 we can see that for low error tolerance levels, all of the best models perform similarly (the CDF lines are nearly on top of one another) and no model clearly dominates another (lies to the left of the other curves over a significant range in the figures). The best attribute-driven models outperform no-info predictions for more than 80% of vehicles and outperform the class only model for all vehicles. This result validates that attribute-driven models are predictively superior to making random guesses or simple metrics like vehicle class. However, all of the attribute driven models are inferior to the static model, suggesting that for this data set the best predictor of the future is assuming it is identical to the past.

Figure 2 is the CDF of the error tolerance for the worst models, instead of the best. The worst models all perform similarly to one another and essentially match the class only model. While a model could be posed that predicts worse than the no-info model, we do not observe it in our utility specifications. Comparing the best and worst model groupings to the class only model across the figures reveals that the best models do predict at least marginally better than the worst models.

The best models and worst models differ most noticeably in their omission of covariates. The best models include some form of almost every covariate, whereas the worst models omit covariates entirely. For example, the worst model as selected by the AIC and KL metrics applied to the estimated data only contains the covariates price and class. Conversely, if we compare only models that contain some form of price, operating cost, acceleration, size covariates and class and brand dummies (style, luxury and automatic transmission dummies could be excluded), then we see no practical difference in the predictive power of the best and worst models. The models' predictive power is robust to attribute functional form but is sensitive to the exclusion of attributes. See Appendix E for selected model coefficients estimates.

Because it is difficult to understand implications of small share variations at the make-model level, we show the vehicle class shares predicted by each of the best predictive models and each of the worst predictive models as selected by the metrics, as well as the no-info and class only models' predicted shares, and compare them to the actual vehicle class shares in Figure 3. These are obtained by summing the individual vehicle shares over each class. We see that even though the AIC/BIC/KL and CDFET metrics select different best models, the variation in the share predictions is small at the vehicle class level. Furthermore even if the worst models are used, at this level of aggregation it does not result in a large discrepancy. The biggest difference in class share prediction for the best and worst models is between the "Best 75th percentile" and "Worst 75th percentile" models for the pickup class, a difference of about 3.9%. The greatest deviation from the actual class share is in the "Best AIC/BIC/KL" model for the SUV class with a difference of about 4.8%. Comparing the best and worst models to the class only model it appears that much of the predictive power in these sets is due to the vehicle class dummies as they have similar share breakdowns as the class only model. Notably, all of the models predict better at the class level than the no information model.

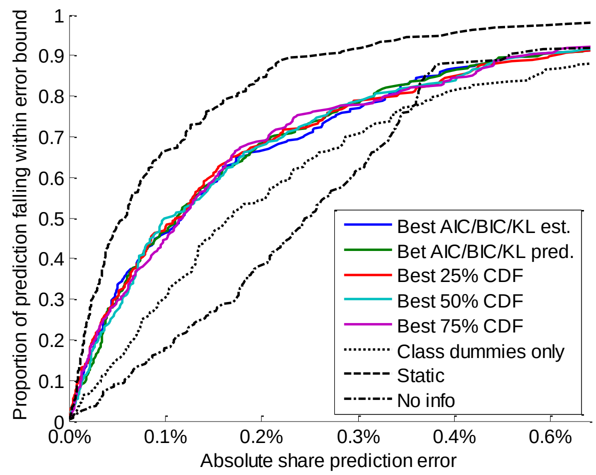

FIGURE 1-THE METRICS SELECT MODELS THAT PREDICT BETTER THAN THE CLASS-ONLY AND NO-INFORMATION MODELS FOR LOW ERROR TOLERANCES

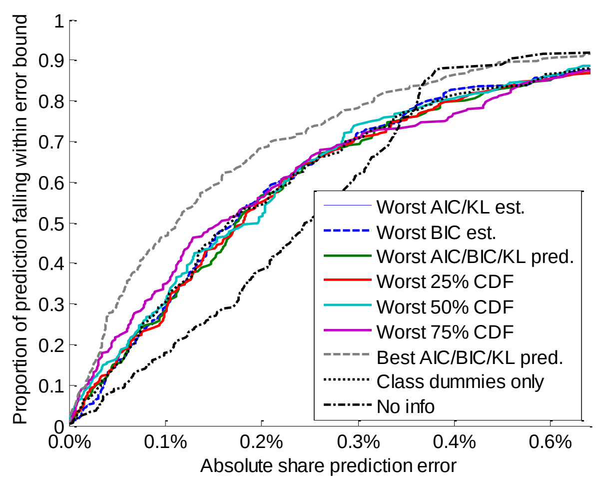

FIGURE 2- THE WORST MODELS AS SELECTED BY THE METRICS PREDICT BETTER THAN THE NO-INFORMATION MODEL FOR LOW ERROR TOLERANCES BUT ONLY AS WELL AS MODELS WITH CLASS DUMMIES ONLY

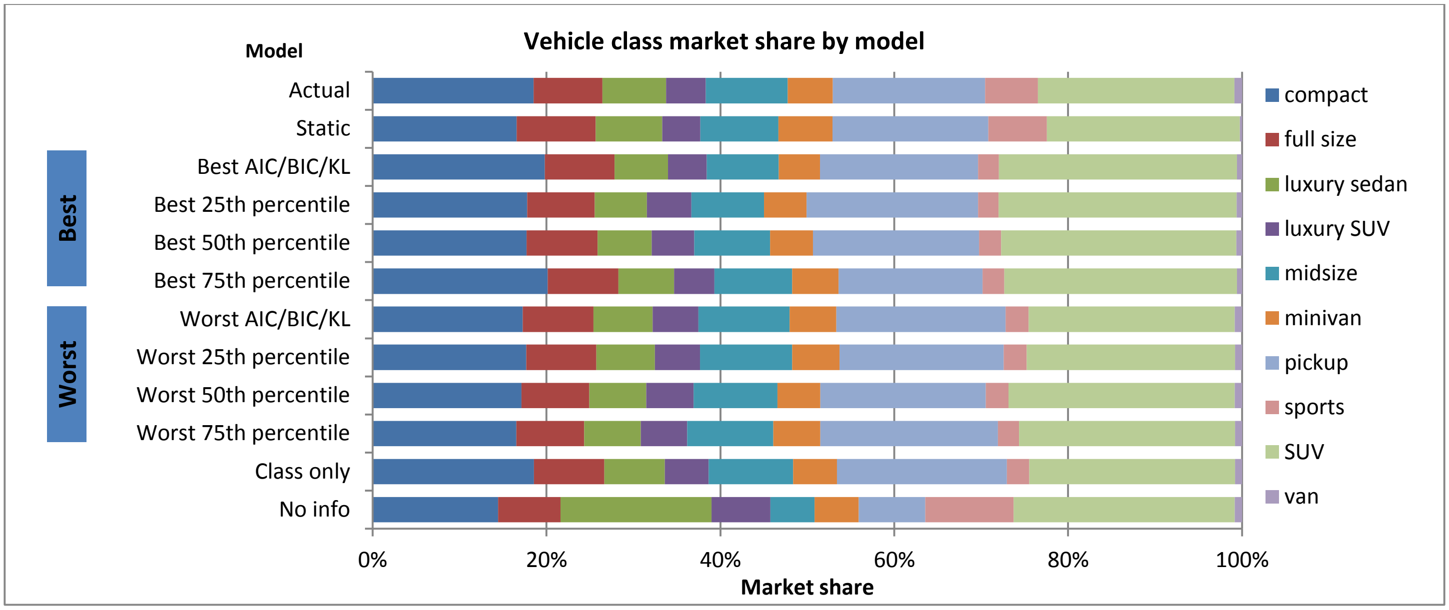

FIGURE 3- VEHICLE CLASS MARKET SHARE FOR BEST AND WORST MODELS COMPARED TO CLASS ONLY AND NO-INFORMATION MODELS AND ACTUAL CLASS MARKET SHARE

4.2 Mixed logit, nested logit, and segment specifications

Using each of the utility functions from the best estimative and best predictive logit models, we fit mixed and nested logit models as described in the methodology section. Figure 4 compares results for these structural specifications. Due to computational limitations, we did not run all 9,000 utility form combinations for the mixed and nested logit structural specifications. Rather we used the results from the logit model output to inform our tested combinations.

We see in Figure 4 that the mixed and nested logit specifications do not predict market share meaningfully better than the logit model. While models with more parameters (mixed logit) and models that mitigate the IIA substitution pattern (mixed logit and nested logit) have the potential to perform better, the cases tested do not appear to offer any substantial predictive benefits in this case.

Appendix F contains plots of actual versus predicted shares for the best predictive logit model functional form for (1) a model fit to 2006 data used to predict 2007 and (2) a model fit directly to 2007 data. We carry out these alternative experiments to verify that there was not a fundamental shift in consumer attitudes that occurred in 2007 and that our prediction quality is not sensitive to the number of years of data used to estimate the model. We see no meaningful difference in the 2007 prediction quality between the models fit to 2004-2006 data, 2006 data only, and 2007 data only. We interpret this as confirmation that our results are not artifacts of the data set we tested. Additionally, the poor prediction quality even when the logit model is fit to the 2007 data itself, which is the most informative data set possible for 2007 predictions, confirms that the major source of error in prediction is error in fit. Without data on missing covariates that influence choice, such as vehicle aesthetics, it is difficult to explain choice behavior at the make-model level with only the available covariates. Future work will examine alternative-specific constants and model calibration to address these discrepancies. See Appendix G for mixed and nested logit coefficient estimates.

We repeated the process of fitting the 9,000 utility function model alternatives to 2004-2006 midsize vehicles only. When the data set is pared to include only midsize vehicles the 2007 midsize vehicle prediction quality improves significantly (see Appendix F). Similarly to the full data set models, the midsize models that included all covariates showed no significant difference in prediction quality. The improvement in prediction quality at the segment level suggests that while these models may not be well suited to describing or predicting market share for a broad market of products that have a wide range of characteristics, they may nevertheless be useful for designing vehicles with a narrower portfolio of competitors.

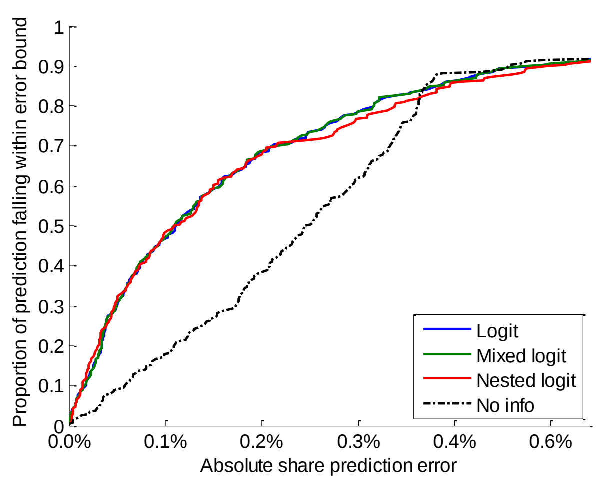

FIGURE 4- THE MIXED LOGIT AND NESTED LOGIT STRUCTURAL SPECIFICATIONS PREDICT NO BETTER THAN THE LOGIT SPECIFICATION WHEN ALL THREE ARE FIT USING THE UTILITY FUNCTIONAL FORM FROM THE BEST AIC/BIC/KL PREDICTIVE LOGIT MODEL

5. CONCLUSIONS

In the literature review section we posed the questions "How widely do predictions vary based on utility function specification?" as well as "How can we quantify share prediction error, and do different evaluation metrics lead to different selections of the 'best predictive model?'" In examining the CDF plots, the vehicle class market share predictions, and the actual versus predicted share plots in Appendix F there is not much distinction in error between models selected based on alternative metrics. Models that perform well on each metric also tend to perform well on the other metrics, and models that perform poorly on one metric also tend to perform poorly on the other metrics. In particular, the logit models selected by AIC and BIC metrics (based on the fit data) have comparable forecast error to those with the best predictive metrics. Share predictions are sensitive to the presence of utility covariates but less sensitive to the form of those covariates. Furthermore, all of the models are superior to the no-information model (which assumes equal share for all vehicle alternatives), so that while all models have error, any model that we tested is better than random guessing. The worst models omitted covariates entirely while there was not much difference in prediction quality among models which included all of the covariates in a variety of functional forms.

In the methodology section we posed the questions "What model would we have selected from 2004-2006 as the 'best' based on goodness-of-fit if we didn't yet have the 2007 data? If we had selected that model, how well would we have done in predicting 2007?" and "Which model predicts 2007 shares best based on metrics of forecast accuracy?" In Appendix E the goodness-of-fit selection metrics and goodness-of-prediction selection metrics used in this paper do not vary significantly across the models selected by each. This suggests that a goodness-of-fit selection criterion (which is known at the time of model estimation) may be sufficient for selecting a model for prediction. More work is needed to assess if this result can be generalized to different periods and forecast horizons.

We find that with no information (predicting all models have equal share of 0.42%), 70% of the share predictions have error greater than 0.2% (an error of about 33,000 vehicles sold for 2007). In contrast, in the best models 30% of share predictions have error greater than 0.2%. The level of 0.2% was chosen as an example, and can be evaluated at other cutoff points to reflect a designer's error tolerance. The more flexible mixed and nested logit models did not offer meaningful improvements for the new vehicle market using the types of covariates that have been used in the literature, although we only examined mixed logit models with diagonal covariance matrices. We caution that these types of models may not be well suited to predict vehicle shares for new design or policy evaluation [7,9–11] and that these same limitations may apply to other product design domains which use discrete choice models [4–6]. Reducing the data set to include only midsize vehicles improved the predictive capabilities of the model. Consequently we would recommend designers carefully consider the scope of research questions and test predictive accuracy before using choice models to inform design decisions. Access to other vehicle attribute information and/or consumer covariates could improve share predictions.

6. LIMITATIONS

Our investigation is a first step in a larger goal of characterizing prediction uncertainty in discrete choice models. We are concerned primarily with utility function specification using vehicle attributes in the context of multinomial logit, (diagonal) mixed logit, and nested logit models. All of our models have error resulting from misspecification and missing information. For example, we use manufacturer suggested retail price (MSRP) data to represent vehicle price, even though transaction prices tend to be lower due to negotiation. We also lack information on attributes that are important in some vehicle classes (like towing capacity for trucks), and we lack information and quantification of some key purchase drivers, such as aesthetics. We lack individual-level choice data with consumer covariates, such as demographics or usage variables [9], which can help explain choice behavior and improve predications when predictions of future population covariates are available. Nevertheless, such limitations are common in choice models used to assess the vehicle market or guide design choices, and our aim is to better understand the implications of these choices and the accuracy of resulting models. More research is needed to assess a wider scope of modeling alternatives, such as methods for handling price endogeneity (instrumental variables), model calibration (with alternative-specific constants), alternative estimation methods (e.g.: generalized method of moments, Bayesian methods), alternative heterogeneity specifications (e.g.: latent class models, full covariance mixed logit models, mixture models, and generalized logit models that account for scale and coefficient heterogeneity [67]).

The framework of our study uses random utility discrete choice models that treat consumers as rational utility maximizers whose utility function can be partly observed by the modeler. While this is a popular approach to modeling consumer choice, there are key criticisms of its basic assumptions. For instance, Axsen and Kurani discuss how preferences change over time [68] and change with cultural symbolism [55] and social interactions [69]. Furthermore, MacDonald et al. suggest that consumers' preferences for attributes do not exist a-priori but rather that products are evaluated on a case by case basis [56]. Morrow et al. [10] and others have suggested that vehicle choice behavior may be better represented by a consider-then-choose model where consumers do not maximize their utility over the universe of available choices but rather screen all choices to narrow them down to a reasonable subset over which utility maximization applies. And more broadly, the Lucas critique warns against use of aggregated historical data to predict outcomes in counterfactual future scenarios [70]. More disaggregated modeling of individual-level behavior may mitigate these effects.

REFERENCES

[1] Hazelrigg G. A., 1998, "A framework for decision-based engineering design," Journal of mechanical design, **120**(4), pp. 653–658.

[2] Train K. E., 2009, Discrete Choice Models with Simulation, Cambridge University Press, New York, NY.

[3] Wassenaar H. J., and Chen W., 2003, "An Approach to Decision-Based Design With Discrete Choice Analysis for Demand Modeling," Journal of Mechanical Design, **125**(3), p. 490.

[4] Wassenaar H. J., Chen W., Cheng J., and Sudjianto A., 2005, "Enhancing Discrete Choice Demand Modeling for Decision-Based Design," Journal of Mechanical Design, **127**(4), p. 514.

[5] Besharati B., Luo L., Azarm S., and Kannan P. K., 2006, "Multi-Objective Single Product Robust Optimization: An Integrated Design and Marketing Approach," Journal of Mechanical Design, **128**(4), p. 884.

[6] Williams N., Azarm S., and Kannan P. K., 2008, "Engineering Product Design Optimization for Retail Channel Acceptance," Journal of Mechanical Design, **130**(6), p. 061402.

[7] Shiau C.-S. N., and Michalek J. J., 2009, "Optimal Product Design Under Price Competition," Journal of Mechanical Design, **131**(7), p. 071003.

[8] Frischknecht B. D., Whitefoot K., and Papalambros P. Y., 2010, "On the Suitability of Econometric Demand Models in Design for Market Systems," Journal of Mechanical Design, **132**(12), p. 121007.

[9] He L., Chen W., Hoyle C., and Yannou B., 2012, "Choice Modeling for Usage Context-Based Design," Journal of Mechanical Design, **134**(3), p. 031007.

[10] Morrow W. R., and Macdonald E. F., 2012, "Models in Decision-Based Design Optimization," Proceedings of the ASME 2012 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, American Society of Mechanical Engineers, ed., Chicago, IL.

[11] Whitefoot K. S., and Skerlos S. J., 2012, "Design incentives to increase vehicle size created from the U.S. footprint-based fuel economy standards," Energy Policy, **41**, pp. 402–411.

[12] Resende C. B., Grace Heckmann C., and Michalek J. J., 2012, "Robust Design for Profit Maximization With Aversion to Downside Risk From Parametric Uncertainty in Consumer Choice Models," Journal of Mechanical Design, **134**(10), p. 100901.

[13] U.S. Energy Information Administration, 2011, "Annual Energy Outlook 2011 with Projections to 2035", Washington, D.C.

[14] Gigerenzer G., and Brighton H., 2009, "Homo Heuristicus: Why Biased Minds Make Better Inferences," Topics in Cognitive Science, **1**(1), pp. 107–143.

[15] Train K., 1979, "Consumers' responses to fuel-efficient vehicles: a critical review of econometric studies," Transportation, **8**, pp. 237–258.

[16] Lave C., and Train K., 1979, "A Disaggregate Model of Auto-type Choice," Transportation Research Part A: General, **13**(1), pp. 1–9.

[17] Boyd J. H., and Mellman R. E., 1980, "The Effect of Fuel Economy Standards on the US Automotive Market: A Hedonic Demand Analysis," Transportation Research Part A: General, **14**, pp. 367–378.

[18] Chen C., 2001, "Design for the environment: A quality-based model for green product development," Management Science, **47**(2), pp. 250–263.

[19] Michalek J. J., Papalambros P. Y., and Skerlos S. J., 2004, "A Study of Fuel Efficiency and Emission Policy Impact on Optimal Vehicle Design Decisions," Journal of Mechanical Design, **126**(6), p. 1062.

[20] Shiau C.-S. N., and Michalek J. J., 2009, "Should Designers Worry About Market Systems?," Journal of Mechanical Design, **131**(1), p. 011011.

[21] Shiau C.-S. N., Michalek J. J., and Hendrickson C. T., 2009, "A structural analysis of vehicle design responses to Corporate Average Fuel Economy policy," Transportation Research Part A: Policy and Practice, **43**(9-10), pp. 814–828.

[22] Berry S., Levinsohn J., and Pakes A., 1995, "Automobile Prices in Market Equilibrium," Econometrica, **63**(4), pp. 841–890.

[23] Nevo A., 2000, "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," **9**(4), pp. 513–548.

[24] Dubé J., Fox J., and Su C., 2009, "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation," (Working paper), pp. 1–51.

[25] Knittel C. R., and Metaxoglou K., 2012, "Estimation of Random Coefficient Demand Models: Two Empiricists' Perspective," (Working paper).

[26] Skrainka B., and Judd K., 2011, "High Performance Quadrature Rrules: How Numerical Integration Affects a Popular Model of Product Differentiation," (Working paper).

[27] Brownstone D., and Train K., 1999, "Forecasting new product penetration with flexible substitution patterns," Journal of Econometrics, **89**, pp. 109–129.

[28] Bunch D. S., Bradley M., and Golob T. F., 1993, "Demand for Clean Fuel Vehicles in California A Discrete Choice Stated Preference Pilot Project," Transportation Research Part A: Policy and Practice, **27A**(3), pp. 237–253.

[29] Brownstone D., Bunch D. S., and Train K., 2000, "Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-fuel Vehicles," Transportation Research Part B: Methodological, **34**(5), pp. 315–338.

[30] McFadden D., and Train K., 2000, "Mixed MNL Models for Discrete Response," Journal of Applied Economics, **15**, pp. 447–470.

[31] Dagsvik J. K., Wetterwald D. G., and Aaberge R., 1996, "Potential Demand for Alternative Fuel Vehicles", Oslo, Norway.

[32] Dagsvik J. K., and Liu G., 2009, "A Framework fo Analyzing Rank-ordered Data with Application to Automobile Demand," Transportation Research Part A: General, **43**, pp. 1–12.

[33] Greene D. L., Patterson P. D., Singh M., and Li J., 2005, "Feebates, Rebates and Gas-guzzler taxes: A Study of Incentives for Increased Fuel Economy," Energy Policy, **33**(6), pp. 757–775.

[34] McCarthy P. S., 1996, "Market price and income elasticities of new vehicle demands," The Review of Economics and Statistics, **78**(3), pp. 543–547.

[35] Musti S., and Kockelman K. M., 2011, "Evolution of the Household Vehicle Fleet: Anticipating Fleet Composition, PHEV Adoption and GHG Emissions in Austin, Texas," Transportation Research Part A: Policy and Practice, **45**(8), pp. 707–720.

[36] Selby B., and Kockelman K., 2012, "Microsimulating Automobile Markets: Evolution of Vehicle Holdings and Vehicle Pricing Dynamics," Transportation Research Forum, **51**(2), pp. 83–96.

[37] Train K., and Winston C., 2007, "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers," International Economic Review, **48**(4), pp. 1469–1497.

[38] Vance C., and Mehlin M., 2009, "Tax Policy and CO2 Emissions-An Econometric Analysis of the German Automobile Market", Bochum, Germany.

[39] Manski C., and Sherman L., 1980, "An Empirical Analysis of Household Choice Among Motor Vehicles," Transportation Research Part A: General, **14**(1976), pp. 349–366.

[40] Mannering F., and Winston C., 1985, "A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization," The RAND Journal of Economics, **16**(2), pp. 215–236.

[41] Berkovec J., and Rust J., 1985, "A Nested Logit Model of Automobile Holdings for One Vehicle Households," Transportation Research Part B: Methodological, **19**(4), pp. 275–285.

[42] Berkovec J., 1985, "Forcasting Automobile Demand Using Disaggregate Choice Models," Transportation Research Part B: Methodological, **19B**(4), pp. 315–329.

[43] Mannering F., Winston C., and Starkey W., 2002, "An Exploratory Analysis of Automobile Leasing by US Households," Journal of Urban Economics, **52**(1), pp. 154–176.

[44] Goldberg P. K., 1998, "The Effects of the Corporate Average Fuel Efficiency Standards in the US," The Journal of Industrial Economics, **46**(1), pp. 1–33.

[45] Choo S., and Mokhtarian P. L., 2004, "What Type of Vehicle do People Drive? The Role of Attitude and Lifestyle in Influencing Vehicle Type Choice," Transportation Research Part A: Policy and Practice, **38**(3), pp. 201–222.

[46] Zhang T., Gensler S., and Garcia R., 2011, "A Study of the Diffusion of Alternative Fuel Vehicles: An Agent-Based Modeling Approach*," Journal of Product Innovation Management, **28**(2), pp. 152–168.

[47] Greene D., 2010, "How Consumers Value Fuel Economy: A Literature Review", Washington, D.C.

[48] Kumar D., Chen W., and Simpson T. W., 2009, "A Market-driven Approach to Product Family Design," International Journal of Production Research, **47**(1), pp. 71–104.

[49] Michalek J. J., Feinberg F. M., and Papalambros P. Y., 2005, "Linking Marketing and Engineering Product Design Decisions via Analytical Target Cascading," Journal of Product Innovation Management, **22**(1), pp. 42–62.

[50] Luo L., Kannan P. K., and Ratchford B. T., 2007, "New Product Development Under Channel Acceptance," Marketing Science, **26**(2), pp. 149–163.

[51] Michalek J. J., Ceryan O., Papalambros P. Y., and Koren Y., 2006, "Balancing Marketing and Manufacturing Objectives in Product Line Design," Journal of Mechanical Design, **128**(6), p. 1196.

[52] Ward's Automotive Group, 2008, "Detailed US Automotive Sales Data."

[53] Polk, 2008, "'Automotive Sales Data'."

[54] Greenspan A., and Cohen D., 1999, "Motor Vehicle Stocks, Scrappage, and sales," Review of economics and statistics, **81**(3), pp. 369–383.

[55] Heffner R. R., Kurani K. S., and Turrentine T. S., 2007, "Symbolism in California's early market for hybrid electric vehicles," Transportation Research Part D: Transport and Environment, **12**(6), pp. 396–413.

[56] MacDonald E. F., Gonzalez R., and Papalambros P. Y., 2009, "Preference Inconsistency in Multidisciplinary Design Decision Making," Journal of Mechanical Design, **131**(3), p. 031009.

[57] MacDonald E. F., Gonzalez R., and Papalambros P., 2009, "The construction of preferences for crux and sentinel product attributes," Journal of Engineering Design, **20**(6), pp. 609–626.

[58] Axsen J., and Kurani K. S., 2012, "Interpersonal influence within car buyers' social networks: Applying five perspectives to plug-in hybrid vehicle drivers," Environment and Planning A, **44**, pp. 1047–1065.

[59] Wooldridge J. M., 2006, Introductory Econometrics: A Modern Approach, Thomas Higher Education, Mason, OH.

[60] 2013, "U.S. All Grades All Formulations Retail Gasoline Prices," Energy Information Administration Data [Online]. Available: http://tonto.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMM_EPM0_PTE_NUS_DPG&f=A. [Accessed: 01-Aug-2013].

[61] "Annual Average Indexex," Bureau of Labor Statitstics Data [Online]. Available: http://www.bls.gov/cpi/cpi_dr.htm#2007. [Accessed: 01-Aug-2013].

[62] Kullback S., and Leibler R. A., 1951, "On information and sufficiency," The Annals of Mathematical Statistics, **22**(1), pp. 79–86.

[63] Akaike H., 1973, "Information theory and an extension of the maximum likelihood principle," Second International Symposium on Information Theory, B.N. Petrov, and F. Csaki, eds., Akademiai Kiado, pp. 267–281.

[64] Schwarz G., 1978, "Estimating the dimension of a model," Annals of Statistics, **6**(2), pp. 461–464.

[65] Burnham K. P., and Anderson D. R., 2002, Model Selection and Multi-model Inference: A Practical Information-theoretic Approach, Springer, New York, NY.

[66] Claeskens G., and Hjort N. L., 2008, Model Selection and Model Averaging, Cambridge University Press, New York, NY.

[67] Fiebig D. G., Keane M. P., Louviere J., and Wasi N., 2009, "The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity," Marketing Science, **29**(3), pp. 393–421.

[68] Axsen J., Mountain D. C., and Jaccard M., 2009, "Combining Stated and Revealed Choice Research to Simulate the Neighbor Effect: The Case of Hybrid-electric Vehicles," Resource and Energy Economics, **31**(3), pp. 221–238.

[69] Axsen J., and Kurani K. S., 2011, "Interpersonal Influence in the Early Plug-in Hybrid Market: Observing Social Interactions with an Exploratory Multi-method Approach," Transportation Research Part D: Transport and Environment, **16**(2), pp. 150–159.

[70] Lucas R., 1976, "Econometric policy evaluation: A critique," Carnegie-Rochester conference series on public policy, **1**, pp. 19–46.

[71] Becker T. A., Sidhu I., and Tenderich B., 2009, "Electric Vehicles in the United States A New Model with Forecasts to 2030", Berkeley, CA.

[72] Balducci P. J., 2008, "Plug-in Hybrid Electric Vehicle Market Penetration Scenarios", Richland, WA.

[73] Duvall M., and Knipping E., 2007, "Environmental Assessment of Plug-In Hybrid Electric Vehicles", Palo Alto, CA.

[74] Greene D. L., Duleep K. G., and McManus W., 2004, “Future Potential of Hybrid and Diesel Powertrains in the U.S. Light-Duty Vehicle Market”, Oak Ridge, TN.

[75] Graham R., 2001, “Comparing the Benefits and Impacts of Hybrid Electric Vehicle Options”, Palo Alto, CA.

[76] National Research Council, 2008, Transitions to Alternative Transportation Technologies: a Focus on Hydrogen, The National Academies Press, Washington, D.C.

[77] Lin Z., and Greene D., 2010, “A Plug-in Hybrid Consumer Choice Model with Detailed Market Segmentation,” Transportation Research Board 89th Annual Meeting, Transportation Research Board, Washington, D.C.

[78] Moawad A., Sharer P., and Rousseau A., 2011, “Light-Duty Vehicle Fuel Consumption Displacement Potential up to 2045”, Chicago, IL.

[79] Duvall M., 2011, “Transportation Electrification: A Technology Overview”, Palo Alto, CA.

[80] Musti S., and Kockelman K., 2011, “Evolution of the Household Vehicle Fleet: Anticipating Fleet Composition, PHEV Adoption and GHG Emissions in Austin, Texas,” Transportation Research Part A: Policy and Practice, **45**(8), pp. 707–720.

[81] Santini D. J., and Vyas A. D., 2005, “Suggestions for a New Vehicle Choice Model Simulating Advanced Vehicles Introduction Decisions (AVID): Structure and Coefficients”, Chicago, IL.

APPENDIX A

Predictions of new vehicle market shares over the next 40 years vary wildly.

Figs. 5-8 plot predictions in vehicle market shares from five selected studies (EIA Annual Energy Outlook 2011 [13], CET Electric Vehicles in the United States- A model with Forecasts to 2030 [71], PNNL Plug-in Hybrid Electric Vehicle Market Penetration Scenarios [72], EPRI Environmental Assessment of PHEVs [73], and ORNL Future Potential of Hybrid and Diesel Powertrains in the US Light-Duty Vehicle Market [74]). The point estimates represent the base case and the error bars represent the range of predictions over sensitivity cases (if any). Because of the myriad of assumptions and variations in methodology, these predictions are not shown on the same plot as to avoid inviting comparisons of comprehensive future vehicle market landscapes; each prediction should be viewed as a singular data point that a report proposes may be realized at the moment in time shown.

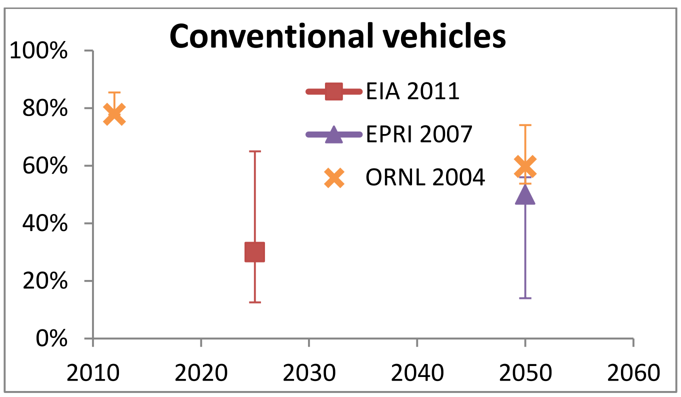

FIGURE 5 – CONVENTIONAL VEHICLE (CV) SHARE OF ANNUAL NEW VEHICLE PURCHASES

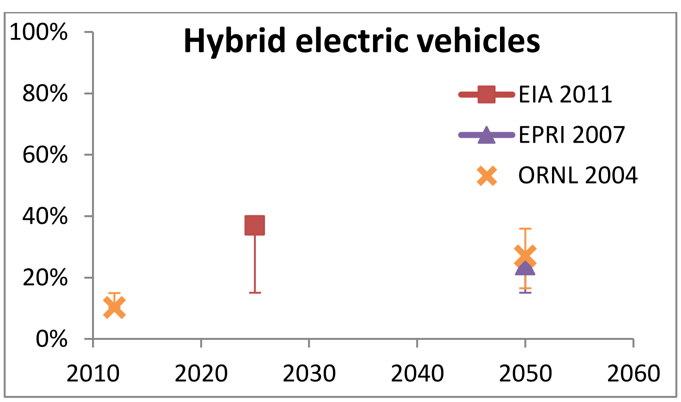

FIGURE 6 – HYBRID ELECTRIC VEHICLE (HEV) SHARE OF ANNUAL NEW VEHICLE PURCHASES

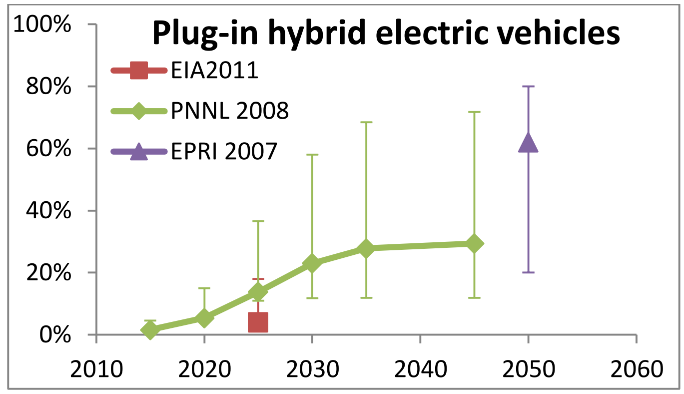

FIGURE 7 – PLUG-IN HYBRID ELECTRIC (PHEV) SHARE OF ANNUAL NEW VEHICLE PURCHASES

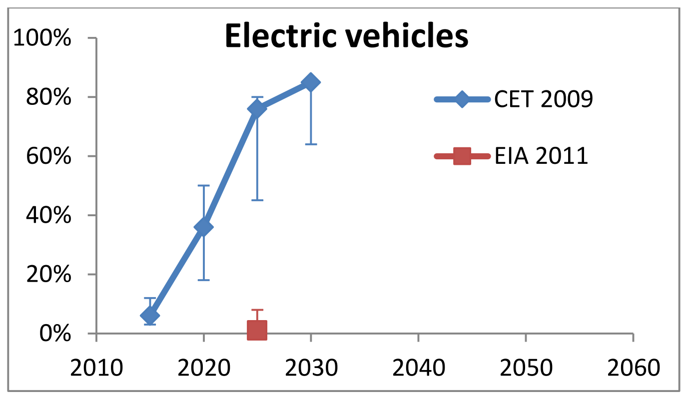

FIGURE 8 – ELECTRIC VEHICLE (EV) SHARE OF ANNUAL NEW VEHICLE PURCHASES

Note: Abbreviations key- (EIA) Annual Energy Outlook 2011 [13]; (CET) Electric Vehicles in the United States- A model with Forecasts to 2030 [71]; (PNNL) Plug-in Hybrid Electric Vehicle Market Penetration Scenarios [72]; (EPRI) Environmental Assessment of PHEVs [73]; (ORNL) Future Potential of Hybrid and Diesel Powertrains in the US Light-Duty Vehicle Market [74]

APPENDIX B

Comparison of predictive and explanatory vehicle demand literature

		Author type				Goal		Study		
Author (Individual or Institute)	**Year**	**Academic**	**Government**	**Other research (see note 1)**	**Consulting firm**	**Explanatory**	**Predictive**	**Reproducible from documentation**	**Peer-reviewed journal publication**	**Uses discrete choice model**
Lave and Train [16]	1979	x				x		x	x	x
Boyd and Mellman [17]	1980				x	x		x	x	x
Berry et al. [22]	1995	x				x		x	x	x
Dagsvik et al. [31]	1996		x			x		x		x
McCarthy [34]	1996	x				x		x	x	x
Brownstone and Train [27]	1999	x				x		x	x	x
Electric Power Research Institute [75]	2001			x		x				See note 2
Choo and Mokhtarian [45]	2004	x				x		x	x	x
Oak Ridge National Laboratory [74]	2004		x				x	x		x
Greene et al. [33]	2005		x			x		x	x	x
Electric Power Research Institute [73]	2007			x			x			See note 2
Train and Winston [37]	2007	x				x		x	x	x
National Research Council of the National Academies [76]	2008			x			x			
Pacific Northwest National Laboratory [72]	2008		x				x			See note 3
Center for Entrepreneurship and Technology (UC Berkeley) [71]	2009	x					x			
Dagsvik and Liu [32]	2009		x			x		x	x	x
Lin and Greene [77]	2009		x				x	x	See note 4	x
Vance and Mehlin [38]	2009	x				x		x		x
Frischknecht et al. [8]	2010	x				x		x	x	x
Argonne National Laboratory [78]	2011		x			x		x		x
Electric Power Research Institute [79]	2011			x			x			
Energy Information Administration [13]	2011		x				x	x		x
Musti and Kockelman [80]	2011	x				x		x	x	x
Zhang et al. [46]	2011	x				x		x	x	x
Whitefoot and Skerlos [11]	2012	x				x		x	x	x

Note: 1.) Independent research agencies may receive government funding; 2.) Report indicates that model was "choice based market model" but is not explicit about model type; 3.) Model is an extension of the ORNL 2004 [74] report, so it is at least partially based on choice modeling but extension methodology is not explicitly described; 4.) Published in conference proceedings

APPENDIX C

Literature discrete choice model survey

Demand covariates	**Lave 1979** [16]	**Boyd 1980** [17]	**Berry 1995** [22]	**Dagsvik 1996** [31]	**Mc-Carthy 1996** [34]	**Brown -stone 1999** [27]	**ORNL 2004** [74]	**ANL 2005** [81]	**Greene 2005** [33]	**Train 2007** [37]	**Dagsvik 2009** [32]	**Vance 2009** [38]	**Frischk-necht 2010** [8]	**EIA AEO 2011** [13]	**Musti 2011** [80]	**Zhang 2011** [46]	**White-foot 2012** [11]
Price																	
price		x		x			x	x	x	x				x	x		x
price + fuel cost/50000mi		x															
ln(income-price)			x														
price/ln(income)						x											
levels of price																x	
income-price/month											x						
price/income	x				x					x		x	x				
(price/income)^2	x																
2 year retained value										x							
Operating cost																	
fuel cost/mi (cost/km)	x				x	x	x	x				x		x			
mi/fuel cost (km/cost)			x														
mpg (L/km)		x		x													
1/mpg										x	x		x				x
levels of mpg																x	
levels of miles/charge																x	
NPV of fuel savings									x								
Maintenance cost																	
repair rating		x															
battery replacement $								x						x			
vehicle and battery maintenance $								x						x			
Acceleration																	
hp (kw)					x					x	x	x					
hp/wt			x				x			x			x				x
wt/hp		x															
f(hp/wt)														x			
known seconds						x											

Demand covariates	Lave 1979	Boyd 1980	Berry 1995	Dagsvik 1996	Mc-Carthy 1996	Brown-stone 1999	ORNL 2004	ANL 2005	Greene 2005	Train 2007	Dagsvik 2009	Vance 2009	Frischk-necht 2010	EIA AEO 2011	Musti 2011	Zhang 2011	White-foot 2012
Other performance																	
handling rating		x															
range						x											
1/range							x	x						x			
top speed				x		x		x									
Size																	
length					x												
width										x							
length-width										x							
length*width			x									x	x				x
(len*wid)^2-2*len*wid													x				
luggage space relative CV						x		x						x			
# of seats	x										x						
Constant																	
constant			x				x										
Intangibles																	
style: (length+width)/height		x															
luxury: noise rating		x															
luxury: dummy A/C standard			x														
safety: dummy crash-test rating					x												
quality: consumer satisfaction rating					x												
quality: reliability rating										x							
transmission: dummy auto is standard										x							
Manufacturer																	
indicator of country of origin					x					x		x	x				
indicator of firm					x					x							
Power train																	
indicator for power source(s)				x		x					x		x	x	x	x	
pollution relative to CV						x											

Demand covariates	Lave 1979	Boyd 1980	Berry 1995	Dagsvik 1996	Mc-Carthy 1996	Brown -stone 1999	ORNL 2004	ANL 2005	Greene 2005	Train 2007	Dagsvik 2009	Vance 2009	Frischk-necht 2010	EIA AEO 2011	Musti 2011	Zhang 2011	White-foot 2012
Vehicle class																	
class indicator like compact, sedan, etc.	x				x	x	x		x	x		x	x		x	x	
sub class indicator like small, standard, luxury	x						x										
External environment																	
fuel availability: indexed to CV							x	x						x			
fuel availability: proportion of stations that can refuel						x											
fuel availability: dummy for can refuel at home								x						x			
make-model availability relative to CV														x			
fraction vehicles equipped to be home or reserve power								x									
policy incentive (HOV lane exemption, rebate, etc.) or penalty (tax)								x	x			x					
Usage																	
commute						x											
household size	x					x							x		x		
number of household vehicles	x														x		
population density					x								x		x		
geographic location					x												
vehicle miles traveled	x																
Demographics																	
age	x			x	x					x					x		
gender				x											x		
education	x					x											
income (not interacted with price)	x														x		
Transaction																	
search process					x					x							
financing										x							
Data source	**SP**	**RP**	**RP**	**SP**	**SP**	**SP**	**RP**	**RP/SP**	**RP**	**SP**	**SP**	**RP**	**RP**	**N/A**	**SP**	**SP**	**N/A**

Note: for "N/A" data sources expert elicitations or literature surveys were used; SP stands for "stated preference" and RP stands for "revealed preference", CV stands for conventional vehicle, HOV stands for high occupancy vehicle (carpool)

APPENDIX D

The Kullback-Leibler and Equivalent Average Likelihood Metrics will rank models identically

If we write KL from Equation 5 and EAL from Equation 6 as in Equation 10 we can see that it consists of a constant k which is a function of the data itself less the natural log of the EAL. Since this is a monotonic transformation, both will result in the same rankings.

$$\boldsymbol{KL(P||Q) = \sum_{i=1:N} ln\left(\frac{P_i}{Q_i}\right) P_i} \tag{5}$$

$$EAL = L^{1/N} \tag{6}$$

$$\begin{aligned} KL(P||Q) &= \sum_{i=1:N} ln\left(\frac{P_i}{Q_i}\right) P_i \\ &= \sum_{i=1:N} \left(P_i ln(P_i) - P_i ln(Q_i)\right) \\ &= \sum_{i=1:N} P_i ln(P_i) - \sum_{i=1}^{N} \frac{n_i}{N} ln(Q_i) \\ &= \sum_{i=1:N} P_i ln(P_i) - \frac{1}{N} ln\left(\prod_{i=1}^{N} (Q_i)^{n_i}\right) \\ &= k - ln(EAL) \end{aligned} \tag{10}$$

APPENDIX E

Estimated coefficients and evaluation metrics for selected models with discussion

For each of the covariates listed in the coefficient table, a numerical value in the row indicates that the covariate was included in the utility function and the value is the coefficient estimate; the blank covariate rows for each model indicate that the covariate was not included in the specification. The "brand dummies included" row contains an "x" if the 36 brand dummies were estimated, but they are not listed for brevity. The magnitude of the covariates was generally on the order of one. All of the coefficients were statistically significant at the two-tailed α=0.01 level.

Price- The price coefficients for all of the model specifications was negative as expected.

Operating cost- Two of the models returned a negative coefficient sign for "mi./fuel cost". We would initially expect this coefficient to be positive, meaning consumers prefer greater values of the covariate or to be able to drive more miles for less money. However, the negative signs occur only when the price form also includes the operating cost. The operating cost coefficient acts as a modifier in this case, and the coefficients must be viewed together, not separately, for interpretation.

The positive coefficient for "gal./mi." in the case of the "Best AIC/BIC/KL predictive model" is also unexpected as it indicates that consumers prefer lower fuel economy. This is potentially related to consumer preference for larger cars, but we partially control for that with the inclusion of class dummies and a size covariate, both of which are present in this model. It also may be related to the preference for higher performance. We partly control for this with the inclusion of acceleration metrics, but the simple hp/wt metrics may not capture all performance issues important to the consumer that are negatively correlated with fuel economy (such as towing capacity, 0-60mph acceleration time, 0-30mph time, 30-60mph time, top speed, all-wheel drive, etc.).

Acceleration- All of the signs are as expected for all of the acceleration forms (f(hp/wt) monotonically increases as the ratio of hp/wt increases).

Size- A size covariate is included in all but one of the best models, and it only appears in the form (length*width). The positive sign indicates that consumers prefer larger cars when vehicle class is controlled for.

Style- Larger values of the covariate represent cars which are relatively lower to the ground as compared to their footprint, e.g. a sports car would be expected to have a larger value of this covariate than a sedan. The best models give mixed estimations on the sign of the coefficient.

Luxury (A/C dummy) and transmission (standard auto dummy)- When included these coefficients are small, though statistically significant, so the impact of either on the utility is minimal.

Manufacturer- All of the best models include the 36 brand dummies. Even AIC and BIC rank models with these additional covariates included despite the metrics' penalties for overfitting.

Class- Class dummies were necessarily included in all of the model specifications, meaning there was no combination tested that did not include them. The van dummy was omitted for identification, so all of the class coefficient estimates represent the utility consumers derive from choosing a vehicle in the respective class over a van. For all of the literature and best models, the coefficients on SUV, luxury SUV and pickup are always positive, whereas for other classes they are of mixed sign.

Table E.1- Coefficient estimates for selected logit models

		<u>Literature informed models</u>				<u>Best estimation set fit</u>	<u>Best prediction set fit</u>			
	Covariate units	**BM-A like**	**BM-B/C like**	**BLP-like**	**Whitefoot-like**	**AIC/BIC/KL/EAL**	**AIC/BIC/KL/EAL**	**CDFET 25th percentile**	**CDFET 50th percentile**	**CDFET 75th percentile**
Cost to consumer										
PRICE										
price	10k $	-0.38			-0.47			-0.42	-0.47	
price+$/50000mi	10k $ +10k $/50k mi		-0.28			-0.44	-0.40			-0.34
ln(price)	ln(10k $)			-1.65						
Operating										
$/mi	$/10 mi									
mi/$	10 mi/$			7.46		-5.67				-7.37
mpg	10 mi/gal	0.03								
gal/mi	gal/10 mi				-48.09		27.15	-39.61		
Performance										
Acceleration										
hp/wt	hp/10 lbs			1.50	1.46	1.04			0.92	
wt/hp	10 lbs/hp	-0.48	-0.40							
f(hp/wt)	exp(hp/wt)						2.76			
hp	hp							0.00		0.00
Size										
Physical										
length	ft									
width	ft									
length-width	ft									
length*width	100 sq-ft			5.69	6.42	8.45	5.20	7.80	4.78	
Intangibles										
Style										
(length*width)/height	100-ft	23.21	22.68			-19.87		-17.85		15.55
Luxury										
a/c std	1			0.11		0.00				
Transmission										
automatic std.	1					0.02		-0.11		
Manufacturer										
geographical (US, Asia, Europe)										
Europe	1									
Asia	1									
brand dummies included	1					x	x	x	x	x
Class dummies										
compact	1	-0.28	-0.46	1.30	1.35	2.78	2.04	2.59	2.26	0.72
fullsize	1	-0.95	-1.18	0.92	0.60	2.04	1.15	2.01	1.27	-0.29
luxury sedan	1	-1.15	-1.51	0.86	0.67	2.68	1.90	2.64	1.98	0.17
luxury SUV	1	0.54	0.37	1.30	1.43	1.65	1.34	1.57	1.37	0.66
midsize	1	-0.22	-0.44	1.63	1.44	2.88	2.04	2.86	2.17	0.55
minivan	1	0.00	-0.23	0.90	0.68	1.23	0.92	1.28	1.10	0.24
pickup	1	0.67	0.56	1.44	1.39	1.85	1.52	1.83	1.54	0.87
sports	1	-1.80	-2.05	0.40	0.33	1.84	0.89	1.70	0.85	-1.17
SUV	1	0.33	0.10	1.51	1.45	2.09	1.76	2.08	1.83	0.69
Total number of covariates		**13**	**12**	**14**	**13**	**52**	**49**	**51**	**48**	**49**

Table E.2- Metrics for selected logit models

			Literature informed models				Best estimation set fit	Best prediction set fit			
Metric	**Min. value over all models**	**Max. value over all models**	**BM-A like**	**BM-B/C like**	**BLP-like**	**Whitefoot-like**	**AIC/BIC/KL/EAL**	**AIC/BIC/KL/EAL**	**CDFET 25th percentile**	**CDFET 50th percentile**	**CDFET 75th percentile**
ESTIMATION SET											
AIC (10e7)	-5.2524	-4.9917	-5.1950	-5.2024	-5.1688	-5.1509	-4.9917	-5.0064	-4.9970	-5.0053	-5.0300
BIC (10e7)	-5.2524	-4.9917	-5.1950	-5.2024	-5.1688	-5.1509	-4.9917	-5.0064	-4.9970	-5.0053	-5.0300
KL	0.1933	0.4572	0.3991	0.4066	0.3726	0.3545	0.1933	0.2082	0.1987	0.2071	0.2320
EAL	0.0049	0.0064	0.0052	0.0052	0.0053	0.0054	0.0064	0.0063	0.0064	0.0063	0.0061
ASE	0.0020	0.0031	0.0029	0.0029	0.0028	0.0028	0.0020	0.0021	0.0021	0.0021	0.0022
PREDICTION SET											
KL	0.2499	0.5016	0.4294	0.4287	0.4112	0.3921	0.2644	0.2499	0.2742	0.2641	0.2892
EAL	0.0046	0.0059	0.0049	0.0049	0.0050	0.0051	0.0058	0.0059	0.0058	0.0058	0.0057
ASE	0.0024	0.0033	0.0030	0.0030	0.0030	0.0029	0.0025	0.0024	0.0025	0.0024	0.0025
CDF cutoff											
Within 25%	0.0004	0.0010	0.0007	0.0007	0.0007	0.0007	0.0004	0.0004	0.0004	0.0005	0.0004
Within 50%	0.0011	0.0022	0.0019	0.0020	0.0019	0.0016	0.0013	0.0012	0.0012	0.0011	0.0013
Within 75%	0.0026	0.0043	0.0037	0.0038	0.0038	0.0036	0.0031	0.0030	0.0029	0.0029	0.0026
Within 100%	0.0251	0.0365	0.0345	0.0342	0.0296	0.0284	0.0264	0.0269	0.0273	0.0271	0.0280

1.) Boyd and Mellman A (BM-A) includes price, gal/mi, repair rating, (len+wid)/hei, hp/wt, noise rating and handling rating; Boyd and Mellman B (BM-B) includes price + fuel cost/50000 miles, repair rating, (len+wid)/hei, hp/wt, and noise rating, Boyd and Mellman C (BM-C) is the same as BM-B but also includes a handling rating; BLP includes a constant, ln(income-price), hp/wt, len*wid, a dummy for air conditioning as a standard feature, miles/fuel cost, and an alternative specific constant; Whitefoot includes price, gallons/mile, hp/wt, len*wid and an alternative specific constant
2.) The AIC, BIC, KL, and EAL metrics select the same model "best model" so they are included as one column in the table
3.) All coefficients are statistically significant at the α=.01 level
4.) US geographical dummy is excluded for identification
5.) Van dummy is excluded for identification
6.) There are 36 brand dummies so for conciseness the coefficient estimates are not included in this table, but an "x" in the "brand dummies included" row indicates that they were estimated as part of the model; the Acura brand dummy is excluded for identification
7.) The boxed metrics indicate the metric value when it was the selection criterion for the model

APPENDIX F

Actual versus predicted shares for the best AIC/BIC/KL predictive logit model for the entire data set and for the midsize vehicle segment only

In all of the figures the solid line represents the space where the predicted share would be identical to the actual share. In Figure 9 - Figure 11 the offset dotted lines represent a prediction error of +/- 0.2% and in Figure 12 the offset dotted lines represent a prediction error of +/-2%. Note that Figure 9 - Figure 11 use the utility function form from the best AIC/BIC/KL predictive model estimated on all 2004-2006 data. Figure 12 uses the utility function form from the best AIC/BIC/KL predictive model estimated on midsize vehicle only 2004-2006 data.

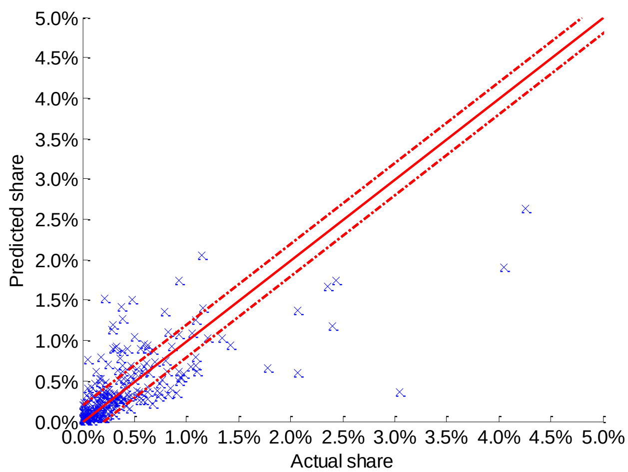

FIGURE 9 – PREDICTION ACCURACY (3 YR): ACTUAL VERSUS PREDICTED 2007 SHARE FOR THE BEST AIC/BIC/KL PREDICTIVE LOGIT MODEL ESTIMATED ON 2004-2006 DATA

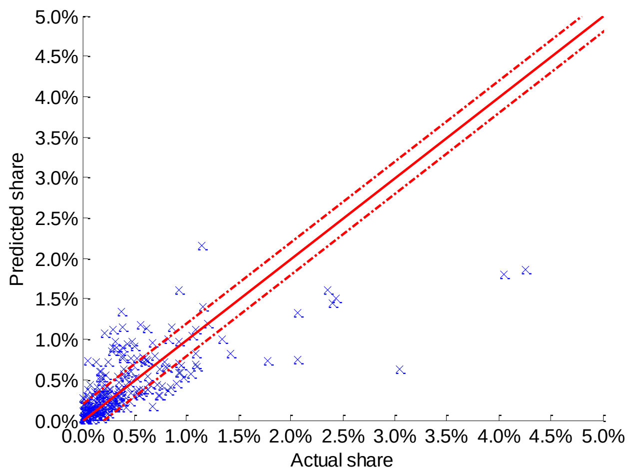

FIGURE 10 – PREDICTION ACCURACY (1 YR): ACTUAL VERSUS PREDICTED 2007 SHARE FOR THE BEST AIC/BIC/KL PREDICTIVE LOGIT MODEL ESTIMATED ON 2006 DATA

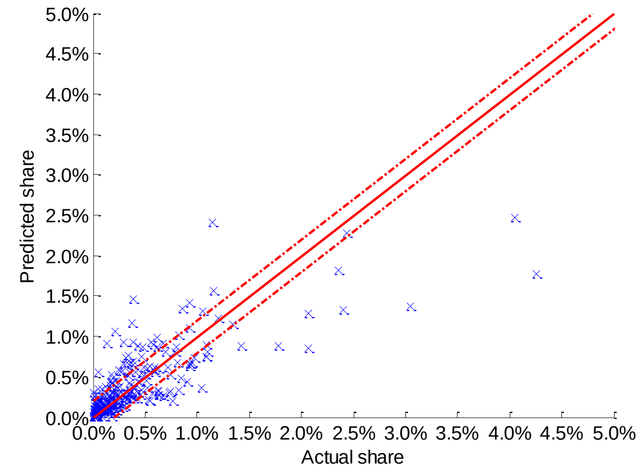

FIGURE 11 – MODEL FIT: ACTUAL VERSUS PREDICTED 2007 SHARE FOR THE BEST AIC/BIC/KL PREDICTIVE LOGIT MODEL ESTIMATED ON 2007 DATA

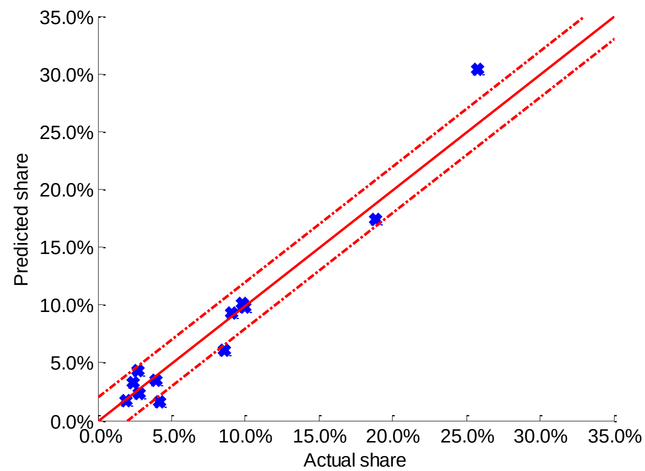

FIGURE 12 – PREDICTION ACCURACY (3 YR, MIDISZE): ACTUAL VERSUS PREDICTED MIDSIZE VEHICLE 2007 SHARE FOR THE BEST AIC/BIC/KL MIDSIZE PREDICTIVE LOGIT MODEL ESTIMATED ON 2004-2006 DATA

APPENDIX G

Selected results from mixed and nested logit model estimation

The evaluation metrics are compared for logit, mixed logit, and nested logit models fit to 2004-2006 data and used to predict 2007 data. The columns designated "Est." represent the model fit using the utility functional form from the **logit** model with the best AIC/BIC/KL metrics calculated from the estimation data. Similarly, the columns designated "Pred." represent the model fit using the utility functional form from the **logit** model with the best AIC/BIC/KL metrics calculated from the prediction data. The "estimation set" metrics are the metrics evaluated for each model on the 2004-2006 estimation data and the "prediction set" metrics are the metrics evaluated for each model on the 2007 prediction data.

Table G.1- Logit, mixed logit and nested logit metric comparison

	Logit		Mixed logit		Nested logit	
Metric	**Est.**	**Pred.**	**Est.**	**Pred.**	**Est.**	**Pred.**
ESTIMATION SET						
AIC (10e7)	-4.9917	-5.0064	-4.9911	-5.0063	-4.9916	-5.0070
BIC (10e7)	-4.9917	-5.0064	-4.9911	-5.0064	-4.9916	-5.0070
KL	0.1933	0.2082	0.1927	0.2081	0.1932	0.2087
EAL	0.0064	0.0063	0.0064	0.0063	0.0064	0.0063
ASE	0.0020	0.0021	0.0020	0.0021	0.0020	0.0021
PREDICTION SET						
KL	0.2644	0.2499	0.2641	0.2500	0.3123	0.2764
EAL	0.0058	0.0059	0.0058	0.0059	0.0056	0.0058
ASE	0.0025	0.0024	0.0024	0.0023	0.0026	0.0025

For the estimated coefficients shown, all models estimated use the utility function form from the best AIC/BIC/KL predictive model estimated on all 2004-2006 data, with a slight structural modification in the nested logit model. We have chosen the vehicle classes as the nests and they are incorporated by means of the λ parameter in Eq. 11 as opposed to representing them as class dummies in the utility function:

$$P_j = \frac{\exp\left(v_j / \lambda_k\right)\left(\sum_{i \in N_k} \exp\left(v_i / \lambda_k\right)\right)^{\lambda_k - 1}}{\sum_{l=1}^{K}\left(\sum_{m \in N_l} \exp\left(v_m / \lambda_l\right)\right)^{l}} \quad (11)$$

where j indexes the products, v is the observed utility of each product, the N represent nests, and the λ are the nest specific parameters to be estimated. In this formulation, no class needs to be excluded for identification. Any attribute coefficient for a given nest can be found by dividing the nominal mean nested logit coefficient in the table by the nest's "class dummy". This modified coefficient is comparable to the logit and mixed logit mean coefficients.

Table G.2- Estimated coefficients for logit, mixed logit and nested logit models

	Logit	Mixed logit		Nested logit
	Mean	**Mean**	**St. dev.**	**Nominal mean**
Physical attributes				
price+$/50000mi	-0.40	-0.40	0.00*	-0.67
gpm	27.15	26.57	1.20	60.50
f(hp/wt)[1]	2.76	2.87	0.28*	3.37**
length*width	5.20	5.19	1.04	7.76
Class dummies				
compact	2.04	2.04		1.76
fullsize	1.15	1.16		1.45
luxury sedan	1.90	1.92		1.80
luxury SUV	1.34	1.35		1.57
midsize	2.04	2.06		1.83
minivan	0.92	0.93		1.30
pickup	1.52	1.53		1.38
sports	0.89	0.89		1.47
SUV	1.76	1.77		1.54
van				0.50
Brand dummies				
Audi	-1.02	-1.02		-1.80
BMW	0.62	0.62		0.88
Buick	0.06	0.07		-0.08
Cadillac	-0.17	-0.17		-0.39
Chevrolet	0.78	0.79		1.01
Chrysler	0.44	0.45		0.53
Dodge	0.65	0.66		0.93
Ford	1.14	1.14		1.70
GMC	0.11	0.11		0.27
Honda	1.08	1.08		1.58
Hummer	-0.53	-0.52		-0.87
Hyundai	0.04	0.05		-0.13
Infiniti	-0.41	-0.41		-0.81
Isuzu	-2.08	-2.08		-3.25
Jaguar	-1.50	-1.50		-2.55
Jeep	1.03	1.03		1.34
Kia	-0.38	-0.38		-0.79
Land Rover	-0.48	-0.48		-0.87
Lexus	0.11	0.12		0.20
Lincoln	-0.69	-0.69		-1.07
Mazda	-0.29	-0.28		-0.59
Mercedes	0.80	0.79		1.14
Mercury	-0.54	-0.53		-1.01
Mitsubishi	-0.84	-0.83		-1.56
Nissan	0.28	0.29		0.30
Oldsmobile	-1.08	-1.05		-1.91
Pontiac	0.01	0.02		-0.16
Porsche	0.64	0.63		1.03
Saab	-1.17	-1.18		-2.12
Saturn	-0.07	-0.06		-0.23
Scion	-0.34	-0.34		-0.81
Subaru	-0.17	-0.15		-0.41
Suzuki	-1.38	-1.37		-2.58
Toyota	0.80	0.81		1.16
Volkswagen	-0.15	-0.15		-0.52
Volvo	-0.80	-0.80		-1.35

Note: All parameters are significant at the α=0.01 level except starred parameters

*Not significant

**Significant at the α=0.10 level

[1] f(hp/wt)=exp(c1*(hp/wt)^c2) where c1= -0.00275 and c2= -0.776 as in the EIA Annual Energy Outlook 2011 [13]