Carnegie Mellon University

CARNEGIE INSTITUTE OF TECHNOLOGY

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF Doctor of Philosophy

TITLE Vehicle Demand Forecasting with Discrete Choice Models:

2 Logit 2 Quit

PRESENTED BY Christine Grace Haaf

ACCEPTED BY THE DEPARTMENT OF

Mechanical Engineering

ADVISOR, MAJOR PROFESSOR DATE

DEPARTMENT HEAD DATE

APPROVED BY THE COLLEGE COUNCIL

DEAN DATE

Vehicle Demand Forecasting with Discrete Choice Models:

2 Logit 2 Quit

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Mechanical Engineering

Christine Grace Haaf

B.S., Mechanical Engineering, Northwestern University
M.S., Mechanical Engineering, Carnegie Mellon University

Carnegie Mellon University
Pittsburgh, PA

December, 2014

Acknowledgements

I would like to thank my academic advisor, Jeremy Michalek, for his continuous support throughout graduate school. Jeremy— you did the best job possible of knowing when to let me get away with hand-wavy explanations of concepts so that I could develop them later and not be momentarily overwhelmed and when to push for specific evidence of outrageous statistical claims.

Thank you to the rest of my committee who have provided comments on oh-so-many drafts of this work over the last 2 years. Ross and Elea— thank you for harkening back to your graduate student days and reading some of my raw code. I acknowledge my other coauthors on published journal papers which have come from work included here or completed prior to my thesis work: Yimin Liu and Camilo Resende.

Thank you to my fellow graduate students from the Design Decisions Laboratory for their helpful research feedback and academic advice over the years. Thank you to John Rula for his suggestion of this thesis' subtitle. Thank you to my housemates in graduate school, Will Schroeder and Alex Styler, who became my first friends in Pittsburgh because they accepted my barter of banana bread for internet. Thank you to Matt Crespi, Thomas Goldring, and Anjali Sehrawat who pulled all nighters with me on homework early on in our programs and made graduate school feel like college all over again. Thank you to fellow members of the Graduate Student Assembly, especially executive committee members David Bergman, Will Boney, Ashlie Henery, Jason Imbrogno, Ruth Poproski, and Andy Schultz. It was an honor serving with you.

Thank you to my friends and family for your support during this process. Lastly, immeasurable thanks to my husband, Tim. I may have been babbling incoherently about math for which you had no interest, but you listened patiently and I appreciate it. I could not have finished without your encouragement in all of the forms it took.

This research was supported in part by:

- National Science Foundation (grant #1064241)
- Ford Motor Company
- Toyota Motor Corporation

The views expressed are those of the authors and not necessarily those of the sponsors.

Committee

Professor Jeremy J. Michalek (Chair)
Mechanical Engineering and Engineering and Public Policy

Inês Azevedo
Engineering and Public Policy

Jonathan Cagan
Mechanical Engineering

Elea Feit
Marketing, Drexel University

W. Ross Morrow
Mechanical Engineering, Iowa State University (2010-2014)
Market and Product Research Analyst, Ford Motor Company (2014)

Table of Contents

List of Tables

List of Figures

Abstract

Discrete choice models (DCMs) are used to forecast demand in a variety of engineering, marketing, and policy contexts, and understanding the uncertainty associated with model forecasts is crucial to inform decision-making. This thesis evaluates the suitability of DCMs for forecasting automotive demand. The entire scope of this investigation is too broad to be covered here, but I explore several elements with a focus on three themes: defining how to measure forecast accuracy, comparing model specifications and forecasting methods in terms of prediction accuracy, and comparing the implications of model specifications and forecasting methods on vehicle design. Specifically I address several questions regarding the accuracy and uncertainty of market share predictions resulting from choice of utility function and structural specification, estimation method, and data structure assumptions.

I[1] compare more than 9,000 models based on those used in peer-reviewed literature and academic and government studies. Firstly, I find that including more model covariates generally improves predictive accuracy, but that the form those covariates take in the utility function is less important. Secondly, better model fit correlates well with better predictive accuracy; however, the models I construct— representative of those in extant literature— exhibit substantial prediction error stemming largely from limited model fit due to unobserved attributes. Lastly, accuracy of predictions in existing markets is neither a necessary nor sufficient condition for use in design.

[1] 1st person singular is used for consistency throughout the abstract, introduction and conclusion, but all studies are based on published or working papers with co-authors who are named in the respective sections.

Much of the econometrics literature on vehicle market modeling has presumed that biased coefficients make for bad models. For purely predictive purposes, the drawbacks of potentially mitigating bias using generalized method of moments estimation coupled with instrumental variables outweigh the expected benefits in the experiments conducted in this dissertation. The risk of specifying invalid instruments is high, and my results suggest that the instruments frequently used in the automotive demand literature are likely invalid. Furthermore, biased coefficients are not necessarily bad for maximizing the *predictive* power of the model. Bias can even aid predictions by implicitly capturing persistent unobserved effects in some circumstances.

Including alternative specific constants (ASCs) in DCM utility functions improves model fit but not necessarily forecast accuracy. For frequentist estimated models all tested methods of forecasting ASCs improved share predictions of the whole midsize sedan market over excluding ASC in predictions, but only one method results in improved long term new vehicle, or entrant, forecasts. As seen in a synthetic data study, assuming an incorrect relationship between observed attributes and the ASC for forecasting risks making worse forecasts than would be made by a model that excludes ASCs entirely. Treating the ASCs as model parameters with full distributions of uncertainty via Bayesian estimation is more robust to selection of ASC forecasting method and less reliant on persistent market structures, however it comes at increased computational cost. Additionally, the best long term forecasts are made by the frequentist model that treats ASCs as calibration constants fit to the model post estimation of other parameters.

1. Introduction and Motivation

Vehicle demand forecasts play a critical role in the development and adoption of vehicle technologies. They inform policy such as in the development of the national Corporate Average Fuel Economy (CAFE) standards[2] [1]; they support engineering design decisions such as how consumers view the tradeoff between performance and fuel economy [2,3]; and they serve as necessary components of broader energy models like those that estimate anticipated electricity or oil consumption [4–7].

The literature employs many methods to forecast vehicle demand. For example, Becker *et al.* [8] forecast new vehicle sales based on the Bass model [9], which predicts new technology diffusion based on maximum market size and the purchases of early adopters. The National Research Council's 2008 and 2010 reports "Transitions to Alternative Transportation Technologies" [10,11] extend the forecasts from the Argonne VISION model [12], which is based on the Annual Energy Outlook [13], by changing input assumptions. Similarly, Balducci [14] explores the implications of future economic scenarios on the projections of Greene *et al.* [15] and the Department of Energy's *Plug-in Hybrid Electric Vehicle R&D Plan* [16]. Zhang *et al.* [17] demonstrate an agent-based approach (in conjunction with a logit model) to study the diffusion of alternative fuel vehicles.

Among the methods available for forecasting vehicle demand, discrete choice models (DCMs) are the most prevalent in the literature (most of the preceding methods named involve a DCM at least indirectly, see Appendix A for more detail). A key

[2] CAFE standards are indirectly informed by a nested logit model. The fuel price and transportation demand of the CAFE analysis is based on the Annual Energy Outlook's [112] National Energy Modeling System (NEMS) that uses a nested logit model in the Consumer Vehicle Choice Submodule (CVCS) [113].

advantage of DCMs is the ability to map consumers' preferences for attributes to product performance in the market. Furthermore, they can be used to evaluate potential market responses to changes in these attributes, which is especially important for product design decision making. Consumer usage and demographic information can be included for additional flexibility, and the models can be estimated on data at various levels of aggregation from individual level surveys to product annual sales data.

The typical DCM form specifies utility u_{ijt} that consumer i observes for product j in market t to be linear in parameters:

$$u_{ijt} = \mathbf{x}'_{jt}\boldsymbol{\beta}_{it} + \xi_{jt} + \varepsilon_{ijt} \tag{1}$$

where $\boldsymbol{\beta}_{it}$ is the vector of attribute preference coefficients for consumer i in market t, $\mathbf{x}_{jt}$ is the vector of attribute values for product j observed by the researcher, ξ_{jt} is the alternative-specific constant (ASC) representing the average utility of the attributes of product j unobserved by the researcher but known to consumers, and ε_{ijt} is the consumer-specific random error term representing variation in consumer utility for attributes (and other factors) unobserved by the researcher. Specific models vary in the way they represent the distribution of $\boldsymbol{\beta}_{it}$ over the population of consumers, the assumed correlation between $\mathbf{x}_{jt}$ and ξ_{jt}, the functional form of the covariates, and the assumed distributional form of the error term ε_{ijt}.

I pose a general case developed by Fiebig *et al.* [18] and relate popular models in the automotive demand literature to special cases of this general model [18]. The subscript t in Eq. 1 is dropped for notational simplicity since this thesis does not address the evolution of preferences over time. The general case here models $\boldsymbol{\beta}_i$ as being distributed over the population with both scale heterogeneity and taste heterogeneity:

$$\boldsymbol{\beta}_i = \sigma_i^{\varepsilon}\boldsymbol{\mu} + \gamma\boldsymbol{\eta}_i + (1-\gamma)\sigma_i^{\varepsilon}\boldsymbol{\eta}_i \tag{2}$$

where σ_i^{ε} is the individual-specific standard deviation of the error term ε_{ijt} representing scale heterogeneity (in Eq. 1 the variance of ε_{ijt} is normalized to 1), $\boldsymbol{\mu}$ represents the mean taste parameter vector, $\boldsymbol{\eta}_i$ is a random vector representing the taste variation over the population of consumers, and γ is a parameter between 0 and 1 that determines the extent to which the model represents pure taste heterogeneity versus pure scale heterogeneity. Various restrictions of this model create special cases that are familiar in the literature (Table 1).

DCM predictions for the market share of annual new vehicle purchases by powertrain vary wildly [4,5,8,14,15]. Forecast discrepancies can arise from many sources including noisy data, finite data, omitted variables, changes in preferences or market conditions between estimation and prediction, and misspecification of the choice process [19]. Model misspecification is virtually guaranteed in most revealed-preference contexts, and while some studies conduct sensitivity analyses with regards to input parameters [4,5,8,14,15], less well understood are the impacts of the inherent statistical properties of DCMs as they apply to engineering and policy decisions. There are many sources of forecast uncertainty, but most work only conducts sensitivity analyses for a few, discrete parameter inputs— like low, base, and high gas price scenarios— and represents the resulting share forecast uncertainty as respective point estimates. In contrast, this work propagates multiple sources of uncertainty— finite sample estimated coefficient variation, prediction of unobservable ASCs, utility function and structural specification— through to share forecasts and characterizes share predictions as distributions of potential outcomes rather than point estimates.

Little validation has been done with regards to the appropriateness of using DCMs for automotive demand forecasting (Frischknecht *et al.* [20] is an exception), and this has created a barrier for adoption by some communities, e.g. US emissions policy modelers [21]. The studies presented in this thesis are intended to advance validation efforts and address a broad need for investigation into the conditions under which DCMs are likely to make good or bad forecasts and market structures or time frames for which the models are more likely to be accurate versus misleading.

The three studies examine a number of these and other DCM issues as they specifically apply to vehicle demand forecasting. The data used in the case studies is of a form frequently encountered by automotive demand researchers: revealed preference sales data at the aggregate market level with many observations (vehicles) per market relative to the number of markets (years) available, and the estimation and forecasting techniques are based on extant vehicle demand literature methods.

The three studies cover various elements of three guiding research themes: (1) defining how to measure forecast accuracy (study 1); (2) comparing model specifications and forecasting methods in terms of prediction accuracy, e.g. utility function (study 1) and taste heterogeneity (study 3) specifications; and (3) comparing design implications of different model specifications and forecasting methods. For the third theme I focus on the implications of mitigating coefficient bias arising from endogeneity (study 2) and evaluating forecast accuracy for all vehicles versus new entrants only (study 1).

In the first study (Chapter 2) I test 9,000 utility function specifications informed by the vehicle demand literature as well as logit, mixed logit, and nested logit structural specifications, and I evaluate models in terms of fit and forecast accuracy according to several measures. In the second study (Chapter 3) I estimate a mixed logit model that includes an ASC in the utility function by maximum likelihood estimation (MLE) and generalized method of moments with IVs (GMM-IV), several techniques to predict future market ASCs are proposed, and the forecasts from models estimated by MLE-C with ASCs fit post estimation and GMM-IV with ASCs fit simultaneously are compared. In the third study (Chapter 4) I use Bayesian estimation to incorporate specification of the ASC as an estimation parameter with a full distribution of uncertainty and to enable estimation of a mixed logit model with correlated model coefficients as opposed to independent coefficients.

Table 1 — Discrete choice model forms

Model	$\boldsymbol{\beta}_i$	Error term ε_{ijt} distributional assumption	Taste parameter heterogeneity specification	Estimated params. (excl. ξ)	Estimation technique by ξ assumption ξ=0	$\xi \neq 0$ $\mathrm{Cov}(\xi_{jt}, \mathbf{x}_{jt}) = \mathbf{0}$	$\xi \neq 0$ $\mathrm{Cov}(\xi_{jt}, \mathbf{x}_{jt}) \in \mathbb{R}^M$
Generalized multinomial logit[1] (G-MNL)	$\boldsymbol{\beta}_i = \sigma_i^\varepsilon \boldsymbol{\mu}$ $\ldots + \gamma \boldsymbol{\eta}_i + \ldots$ $(1-\gamma)\sigma_i^\varepsilon \boldsymbol{\eta}_i$	$EV1(0, \sigma_i^\varepsilon)$ $\sigma_i^\varepsilon \sim LN(1, \tau^2)$	$\mathrm{Cov}(\beta_m, \beta_n) \in \mathbb{R}, \forall m, n$	γ, τ^2 $\boldsymbol{\mu}, \boldsymbol{\Sigma}_\beta$			
Hierarchical mixed logit (MIXL)[2]	$\boldsymbol{\beta}_i = \boldsymbol{\mu} + \boldsymbol{\eta}_i$ $\ldots + \boldsymbol{\Delta}\mathbf{z}_i$	EV1(0,1) (iid)	$\mathrm{Cov}(\beta_m, \beta_n) \in \mathbb{R}, \forall m, n$	$\boldsymbol{\mu}, \boldsymbol{\Sigma}_\beta, \boldsymbol{\Delta}$	Bayes [61]	Bayes [101]	
MIXL[3]	$\boldsymbol{\beta}_i = \boldsymbol{\mu} + \boldsymbol{\eta}_i$	EV1(0,1) (iid)	$\mathrm{Cov}(\beta_m, \beta_n) \in \mathbb{R}, \forall m, n$	$\boldsymbol{\mu}, \boldsymbol{\Sigma}_\beta$	Bayes [99,105]	Bayes [103,104,109] ♠	
			$\mathrm{Cov}(\beta_m, \beta_n) = 0, \forall m, n$	$\boldsymbol{\mu}, \boldsymbol{\sigma}^\beta$	MLE [20,59] Bayes [17,55] ♦	♣ ♠	GMM [33,35–37,49,52–54,58], MLE [47], Other [49] ♣
Mixed probit	$\boldsymbol{\beta}_i = \boldsymbol{\mu} + \boldsymbol{\eta}_i$	N(0,1) (iid)	$\mathrm{Cov}(\beta_m, \beta_n) \in \mathbb{R}, \forall m, n$	$\boldsymbol{\mu}, \boldsymbol{\Sigma}_\beta$		Bayes [100]	
			$\mathrm{Cov}(\beta_m, \beta_n) = 0, \forall m, n$	$\boldsymbol{\mu}, \boldsymbol{\sigma}^\beta$	MLE [55] Bayes [102]		
Scaled MNL[1]	$\boldsymbol{\beta}_i = \sigma_i^\varepsilon \boldsymbol{\mu}$	$EV1(0, \sigma_i^\varepsilon)$ $\sigma_i^\varepsilon \sim LN(1, \tau^2)$	$\mathrm{Cov}(\beta_m, \beta_n) = 0, \forall m, n$	$\tau^2, \boldsymbol{\mu}$			
Nested logit	$\boldsymbol{\beta}_i = \boldsymbol{\mu} / \lambda_q$	EV1(0,1) $\mathrm{Cov}(\varepsilon_{ijt}, \varepsilon_{ikt}) = 0,$ $\forall j \in G_q, k \in G_r, q \neq r$	$\sigma_m^\beta = 0, \forall m$	$\boldsymbol{\mu}, \boldsymbol{\lambda}$	♦	MLE [46] Other [50,62]	Other [34,45]
L	$\boldsymbol{\beta}_i = \boldsymbol{\mu}$	EV1(0,1) (iid)	$\sigma_m^\beta = 0, \forall m$	$\boldsymbol{\mu}$	MLE [40,55,57,59] Bayes [99] ♦	MLE [31,60] Other [96]	Other [39]

Notes: boxed cells indicate model, ξ assumption, and estimation method are included in a thesis study (♦ study 1, ♣ study 2, ♠ study 3); some DCM assumptions and models are excluded; "Other" estimation category may include expert elicitation, canned software, or two-staged least squares; Notation: $\boldsymbol{\mu}$ is a vector of mean taste parameters, $\boldsymbol{\eta} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\beta)$ is the residual taste heterogeneity controlling for scale, and $\boldsymbol{\sigma}^\beta$ is the vector of diagonal elements of $\boldsymbol{\Sigma}_\beta$; ξ is the ASC, G represents a set of vehicles or nest, and $\boldsymbol{\lambda}$ is a vector of nest-specific parameters; Abbreviations: EV1=extreme value type 1, LN=lognormal, N=normal, iid=independently and identically distributed, MLE=Maximum Likelihood Estimation, GMM=Generalized Method of Moments

[1] The distribution of σ_i^ε need not be lognormal but cannot permit negative values of σ_i^ε. Setting the mean of σ_i^ε equal to 1 is a standard normalization. [18]

[2] Δ is a matrix relating consumer variables $\mathbf{z}_i$ to $\boldsymbol{\beta}_i$, when an individual's taste parameters follow a general individual-specific distribution (like $\Delta\mathbf{z}_i$) the model is referred to as a mixture model

[3] When η is distributed according to a probability mass function rather than a continuous density function the mixed logit model is equivalent to the latent class model

2. Sensitivity of Vehicle Market Share Predictions to Discrete Choice Model Specification

This first study covers all three major research themes: defining how to measure forecast accuracy, comparing model specifications and forecasting methods in terms of prediction accuracy, and comparing design implications of different model specifications and forecasting methods. We ask specific questions: How should we measure prediction accuracy, and do different measures lead to different conclusions about which models predict best? How widely do predictions vary for alternative model specifications? Which specifications have the best predictions, and how good are they? What are the implications for using choice models in design, particularly of new products? Of the assumptions and models in Table 1 this study assumes the alternative-specific constant (ASC) is equal to zero and specifies logit, nested logit, and mixed logit with independent coefficients models[3].

When design decisions are informed by consumer choice models, uncertainty in choice model predictions creates uncertainty for the designer. We investigate the variation and accuracy of market share predictions by characterizing fit and forecast accuracy of discrete choice models (DCMs) for the US light duty new vehicle market. Specifically, we estimate multinomial logit models for 9,000 utility functions representative of a large literature in vehicle choice modeling using sales data for years 2004-2006. Each model predicts shares for the 2007 and 2010 markets, and we compare several quantitative measures of model fit and predictive accuracy. We find that (1) our

[3] The study presented in this chapter has been completed and has appeared in the Journal of Mechanical Design [93]. In this chapter the use of first person plural includes coauthors Jeremy Michalek, Ross Morrow, and Yimin Liu.

accuracy measures are concordant: model specifications that perform well on one measure tend to also perform well on other measures for both fit and prediction. (2) Even the best DCMs exhibit substantial prediction error, stemming largely from limited model fit due to unobserved attributes. A naïve "static" model, assuming share for each vehicle design in the forecast year = share in the last available year, outperforms all 9,000 attribute-based models when predicting the full market 1-year-forward, but attribute-based models can predict better for 4-year-forward forecasts or new vehicle designs. (3) Share predictions are sensitive to the presence of utility covariates but less sensitive to covariate form (e.g. miles per gallons versus gallons per mile), and nested and mixed logit specifications do not produce significantly more accurate forecasts. This suggests ambiguity in identifying a unique model form best for design. Furthermore, the models with best predictions do not necessarily have expected coefficient signs, and biased coefficients could misguide design efforts even when overall prediction accuracy for existing markets is maximized.

2.1 Introduction

Design researchers have proposed a variety of methods to predict the influence of design decisions on firm profit as part of a broader effort to base design decisions explicitly on predictions of downstream consequences for the firm [22]. The majority of these methods apply discrete choice methods [23] to predict consumer choice as a function of product attributes and price. Such predictions are proposed as a way to guide or even optimize design decisions [20,24–31]. Application of choice models within design implicitly relies on accurate choice predictions [3,26]. Given the many sources of uncertainty in such models, however, Frischknecht *et al.* [20] question the suitability of using choice models in a design context. At a minimum, researchers must be aware of the degree of prediction error and uncertainty when employing market models in design.

Prediction error can arise from many sources, including noisy data, finite data, omitted variables, changes in preferences or market conditions between estimation and prediction, and misspecification of the choice process [19]. Recent design research has modeled some aspects of model uncertainty by posing distributions over model coefficients [3,26]. Following standard asymptotic results, coefficient distributions are most often assumed to be normal with mean vector and covariance matrix determined by properties of the log-likelihood function. However, model misspecification is virtually guaranteed in most revealed-preference contexts, given the complexity of human choice behavior for difficult decisions [32], and standard statistical results do not apply in such settings, nor are they comprehensive. Moreover, few applications of choice modeling in any field carefully analyze sensitivity of model fit or forecast accuracy using alternative utility specifications or error structures that might imply different design decisions. A realistic portrait of these aspects of predictive error cannot be captured in a fully generalizable way across product domains or contexts but can nevertheless be better understood via data-driven examination in the specific market of interest.

We focus on the effect of model specification and characterize share prediction accuracy of multinomial logit models in an empirical study of recent new vehicle markets using revealed preference sales data. The automotive sector is among the most popular product domains for application of choice modeling in general [17,20,25,28–31,33–60] and in the design literature specifically [17,20,25,28–30,38,43,44,48,51]. Logit models, along with variants including nested and mixed logit models, represent the most popular modeling approach by far. While stated choice methods fit to conjoint survey data are common [17,24,29,43,55–57], they measure hypothetical choices and generally must be calibrated to achieve a match with market sales data [41,61]. We focus here on choice models fit to aggregate market sales data [20,25,28,30,33–37,39,44,45,48–54,56,59,62].

Given the importance of the vehicle choice application in the design literature and beyond, a better understanding and characterization of prediction accuracy in this domain and its implications for design is needed. We aim to address this need with an automotive case study by fitting a set of models representative of those in the literature to past vehicle sales data, using the resulting models to predict sales in later years, and assessing prediction accuracy.

Our analysis is focused on the following research questions:

(Q2.1) How should we measure prediction accuracy, and do different measures lead to different conclusions about which models predict best?

(Q2.2) How widely do predictions vary for alternative model specifications? Which specifications have the best predictions, and how good are they?

(Q2.3) What are the implications for using choice models in design, particularly of new products?

The design literature has not yet investigated what measures of forecast accuracy exist or compared these measures to understand how they differ in characterizing accuracy, thus Q2.1. Q2.2 applies appropriate measures to the specific task in our case study. Q2.3 focuses on prediction accuracy for new vehicle designs, and we examine the relationship between accurate prediction in existing markets versus potential to predict response to new designs that deviate from market patterns (e.g. correlations with unobserved attributes). We view design as primarily interested in the introduction of new products or (large) changes to product features, motivating a focus on new vehicles.

2.2 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 A compares publications of each type.

Forecasting studies are conducted by private or government research entities or issued in report format from an academic research institute (see Appendix A). Reports are typically not 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 [5] 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 forecasting literature is typically not used in engineering design models due to lack of transparency and documentation of data and modeling assumptions and lack of models that make predictions as a function of design variables. Rather, models from the explanatory literature are applied in a predictive context.

The bulk of the new vehicle purchase demand literature is explanatory, conducted by academic researchers and published in peer-reviewed academic journals (see Appendix A). 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.* [20] is a rare exception). In general, models that fit the existing data best may not necessarily be the best at predicting counterfactuals: statistical models may be misspecified, containing systematic difference in prediction from true process ("bias"), or may be sensitive to overfitting noise in the data instead of signal ("variance") [63].

The earliest applications of economic models for overall automotive demand focused on macroeconomic variables and, as Train [64] highlights, only included price. These studies are referred to as aggregate studies because the level of granularity of predictions is at the whole market or vehicle class level as opposed to individual vehicle designs[4]. Disaggregate studies evolved to predict the number of vehicles an individual household would choose to own [64]. For example, Lave and Train [60] advanced this

[4] We use the term "vehicle design" to refer to vehicle make-model.

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: Boyd and Mellman [59], who propose a random coefficient logit model adopted by others [31,44,51,65,66]; Berry *et al.* [58], who include an ASC in the utility function of a random coefficient demand model adopted by others [33,35,67,68];

Brownstone and Train [55], who propose several choice model specifications using the results of a California conjoint study described in Bunch *et al.* [69] and adopted by others [70,71]; and Whitefoot and Skerlos [31], who investigate the effect of fuel economy standards on vehicle size and employ a logit model with coefficients drawn directly from the literature. Other new-vehicle purchase models include [40,42,45,47,50,56,57,72].

We use the preceding literature to inform comparison models of our creation; we do not recreate prior models exactly due to limited availability of data or specifics about estimation methods. Instead, we form a combinatorial set of utility specifications using covariate forms from these prior models, fit them all to a common data set, and test them all on a common prediction set. Table 2 summarizes the covariates used in past models and those adopted for our tests.

2.3 Methods

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 brevity we refer to the multinomial logit model as "logit"). 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. We then estimate the logit coefficients on US consumer vehicle purchase data from 2004-2006 and predict market share for each of the vehicles in the US purchase data from 2007 and 2010.

Using the measures described in Section 2.3.4, we rank the predictive accuracy across utility function specification for each of the measures.

2.3.1 The Data Set

Our data set draws vehicle attribute information from Ward's Automotive Index [73] and aggregate US sales data from Polk [74] for vehicle sales during 2004-2007 and 2010. Other studies have used a variety of data sources (including these) as well as stated preference surveys. We use 2004-2006 data for estimation because we expect three years of data to be sufficient to predict a successive year, and we predict 2007 and 2010 sales to examine the effects of different time horizons. We implicitly assume that all individuals who purchased a vehicle considered all of the other vehicles available in the same year and made a compensatory decision based on vehicle attributes.

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, which is likely less sensitive to macroeconomic factors. There are many factors that drive share and 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.

2.3.2 Model Specification

Each model uses the utility function:

$$u_{ij} = \mathbf{x}_j^{'}\boldsymbol{\beta} + \varepsilon_{ij} \tag{3}$$

where u_{ij} is the utility of vehicle design 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 randomly selected consumer will choose vehicle j can be expressed as:

$$P_j = \frac{\exp\left(\mathbf{x}_j^{'}\boldsymbol{\beta}\right)}{\sum_{k=1}^{J}\exp\left(\mathbf{x}_k^{'}\boldsymbol{\beta}\right)} \tag{4}$$

where J is the number of vehicle design options. This is the (multinomial) logit formula.

While any choice of covariates $\mathbf{x}$ is possible in principle, we focus on combinations of covariates used in the prior literature. We survey the automotive demand literature to identify the universe of independent variables historically used in automotive DCMs (Appendix B). From this list of candidate covariates, we select a subset to define a manageable set of models. Many of the models in the literature include demographic or consumer usage covariates, but because Ward's Automotive Index data [73] does not include individual-level choices, we ignore demographics. For some 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. We omit several variables 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 0-60mph 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.
- *Attributes specific to alternative-vehicles* (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 powertrains.

Table 2 — Covariate forms tested in utility function specifications

Covariate	Functional form options				
	Option 0	Option 1	Option 2	Option 3	Option 4
Price		price ($)	price + op cost	ln(price)	
Operating cost[1]	Excl.	fuel cost/mile	miles/fuel cost	miles/gallon	gallons/mile
Acceleration[2]	Excl.	horsepower/weight (hp/wt)	wt/hp	$\exp(c_1*(hp/wt)^\wedge c_2)$	hp
Size	Excl.	length	width	length-width	length*width
Style	Excl.	(length*width)/height			
Luxury	Excl.	dummy if air-conditioning is standard			
Transmission	Excl.	dummy if auto. transmission is standard			
Brand	Excl.	dummy for country of origin[3]	dummy for brand[4]		
Vehicle class		dummies for vehicle class[5]			

[1] Fuel cost is average annual gas price [115] in 2004 dollars, adjustment based on the Consumer Price Index [116]

[2] c_1= -0.00275 and c_2= -0.776 as in the EIA Annual Energy Outlook 2011 [5]

[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.

The highlighted covariates in Table 2 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 well into a small number of discrete levels. Price is always included as a covariate and can take any of the forms listed in Table 2; vehicle class dummies are also always included. The other highlighted covariates in Appendix B can take one of the forms listed in Table 2 or can be excluded from the utility function entirely ("excluded" option). Given these covariate options, there are 9,000 possible utility specifications for the logit model outlined in Table 2. Operating cost includes the macroeconomic variable of retail gas price. Though we aim to exclude non-vehicle attributes, this covariate was particularly prevalent in the literature. Furthermore, while having more covariates cannot decrease best model fit on a given data set, that does not imply that more covariates will improve model forecast accuracy. In general, introducing more covariates introduces the risk of overfitting the estimation data.

From the selected covariates, we assume that the utility function is linear in parameters (a standard assumption in the vast majority of logit model applications because it ensures that the log-likelihood function is concave [23]) and construct models using all possible linear combinations of covariates.

Many of these covariates are correlated. Such correlations can induce bias in the estimated coefficients if not corrected [75]. However, while this presents difficulties in drawing inferences from the coefficients (e.g. willingness-to-pay) it does not necessarily 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 likely to hold for near-term predictions, though it may not hold for new designs that do not follow prior patterns in the marketplace.

For illustration of this concept, suppose the true choice generator uses the utility function $u(\mathbf{x}|\boldsymbol{\beta}_0) = \boldsymbol{\beta}_0'\mathbf{x} + \varepsilon$, and the designs in the market follow a pattern: $\mathbf{x} = \mathbf{A}'\mathbf{y}$ for $\mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, m < n$. Then for any coefficient vector $\{\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \boldsymbol{\Delta}: \mathbf{A}\boldsymbol{\Delta} = \mathbf{0}\}$, $u(\mathbf{x}|\boldsymbol{\beta}) = \boldsymbol{\beta}'\mathbf{A}'\mathbf{y} + \varepsilon = (\mathbf{A}\boldsymbol{\beta}_0 + \mathbf{A}\boldsymbol{\Delta})'\mathbf{y} + \varepsilon = \boldsymbol{\beta}_0'\mathbf{x} + \varepsilon = u(\mathbf{x}|\boldsymbol{\beta}_0)$. Therefore, choice probabilities are identical for any $\boldsymbol{\Delta}$ in the null space of $\mathbf{A}$, and $\boldsymbol{\beta}_0$ is not identifiable: coefficient estimates $\boldsymbol{\beta}$ could be arbitrarily far from their true value $\boldsymbol{\beta}_0$. Nevertheless, $u(\mathbf{x}|\boldsymbol{\beta}) = u(\mathbf{x}|\boldsymbol{\beta}_0)$, so utility estimates (and therefore choice probabilities) can be correct even for arbitrarily biased coefficients as long as the new designs follow the pattern in the marketplace $\mathbf{x} = \mathbf{A}'\mathbf{y}$. If a new design deviates from the prior pattern $\tilde{\mathbf{x}} = \mathbf{A}'\mathbf{y} + \mathbf{z}$, utility (and therefore choice probabilities) may be biased: $u(\tilde{\mathbf{x}}|\boldsymbol{\beta}) = (\boldsymbol{\beta}_0 + \boldsymbol{\Delta})'(\mathbf{A}'\mathbf{y} + \mathbf{z}) + \varepsilon = (\mathbf{A}\boldsymbol{\beta}_0 + \mathbf{A}\boldsymbol{\Delta})'\mathbf{y} + (\boldsymbol{\beta}_0 + \boldsymbol{\Delta})'\mathbf{z} + \varepsilon = \boldsymbol{\beta}'_0\mathbf{A}'\mathbf{y} + \boldsymbol{\beta}'_0\mathbf{z} + \boldsymbol{\Delta}'\mathbf{z} + \varepsilon = u(\tilde{\mathbf{x}}|\boldsymbol{\beta}_0) + \boldsymbol{\Delta}'\mathbf{z}$. Therefore, models that predict well overall may nevertheless have biased coefficients that predict poorly for new designs that deviate from the market pattern. We assess predictive accuracy for products in the marketplace and also examine variation in implications of coefficient estimates for new designs.

2.3.3 Model Estimation

The likelihood of the estimated parameters L is defined as the probability of generating the observed data given the estimated parameter values:

$$L\left(\hat{\boldsymbol{\beta}}\middle|\mathbf{x}\right) = \prod_{k=1}^{J}\left(P_k\right)^{n_k} \tag{5}$$

where n_k are the sales of vehicle k. The maximum likelihood estimator of the parameters $\hat{\boldsymbol{\beta}}$ is the value of the vector that maximizes L. The monotonic transformation ln(L) is typically used as the objective function for computational benefit. For more detail on logit models and their estimation see Train [23].

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:

$$\boldsymbol{\beta} \sim N\left(\boldsymbol{\mu},\boldsymbol{\Sigma}\right) \tag{6}$$

where $\mathbf{\Sigma}$ is a diagonal matrix, and the maximum likelihood estimation (MLE) procedure estimates the elements of $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ using numerical integration [23]. This specification relaxes the independence from irrelevant alternatives (IIA) restriction for substitution patterns [23].

Our nested logit specification divides the vehicles into groups or nests by vehicle class and fits a logit model to each of the nests. We assume that the utility functional form is the same for each nest, but coefficients may differ across nests. 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.

As generalizations of the logit model, nested and mixed logit models will necessarily fit any set of *estimation* data at least as well as the logit. The mixed logit generalization of the logit model is even flexible enough to represent most random utility maximization models, given enough flexibility over the coefficient distribution [23]. However, nested and mixed logit models need not *predict* as well as logit models due to the potential for overfitting.

2.3.4 Evaluation Measures

After fitting each of the model specifications, we evaluate prediction error using likelihood measures, the Kullback-Leibler divergence (KL) [76], a cumulative distribution of error tolerance (CDFET), and the average share error (ASE), and we compare the goodness-of-fit using the above measures as well as the Akaike Information Criterion (AIC) [77] and the Bayesian Information Criterion (BIC) [78]. Each of these measures is described below. We compare models selected as best by these measures to one another and to literature-informed benchmark models.

Likelihood: Likelihood, defined in Eq. 5, and monotonic transformations of likelihood, such as log-likelihood ln(L) and average likelihood ($L^{1/J}$, where J is the number of choices observed) measure the probability that the model would generate the data observed. When comparing two models for the same data set, the model with larger L is more likely to generate the data observed.

Kullback-Leibler divergence: The KL divergence measures the difference between a predicted distribution and the true distribution [79].

$$KL\left(s_j \| p_j\right) = \sum_{j=1}^{J} ln\left(\frac{s_j}{P_j}\right) s_j \tag{7}$$

where $s_j = n_j/J$ is the market share of vehicle design j. The KL measure is also a monotonic transformation of L, thus L and KL will rank models identically, and maximizing likelihood is equivalent to minimizing KL (see Appendix C for proof).

Average share error: ASE measures the average error in share predictions across the vehicle designs.

$$ASE = \frac{1}{J} \sum_{j=1}^{J} \left| s_j - P_j \right| \tag{8}$$

We report ASE as a summary statistic in Appendix D but do not use it as a basis for model selection because it does not holistically capture distribution divergence: It will not distinguish between models with large error for one vehicle alternative versus the same degree of error spread out among many vehicle alternatives.

Error tolerance CDF: The cumulative distribution function (CDF) of Error Tolerance (CDFET) graphs the fraction of vehicles with absolute share prediction error,

$|s_j - P_j|$ for vehicle design j, less than a specified value. This measure, to our knowledge proposed here, evaluates a model in terms of error tolerance levels. We use absolute share error rather than relative error because relative error overemphasizes small prediction errors for vehicles with small market share. A CDFET is a more comprehensive description of model prediction error than likelihood measures because it characterizes the distribution of accuracy across the vehicle share predictions, rather than just how well a model predicts "on average".

Two additional measures apply only to assess fit with estimation data, not predictive accuracy [80].

Akaike information criterion: AIC is a variation of likelihood that attempts to penalize overfitting.

$$AIC = 2ln(L) - 2k \tag{9}$$

where k is the number of model parameters.

Bayesian information criterion: BIC is similar to AIC but with a stronger penalty for an increasing number of covariates.

$$BIC = 2ln(L) - ln(J)k \tag{10}$$

AIC and BIC can take on the value of any negative real number, have no standalone meaning, and are only useful as compared to other candidate models fit to the same data set. Larger values are preferred. Derivations and consistency proofs for the KL, AIC and BIC measures can be found in [80].

2.4 Results

Of the 9,000 tested utility function specifications, for 8,993 (99.9%) the Knitro optimization algorithm for Matlab converged to likelihood-maximizing coefficients, and the other seven failed to converge. Only the 8,993 models that successfully converged were considered as candidate models. The candidate models were ranked from best to worst on each measure. There were no two models with identical values for any measure (no ties). In the following results "best models" refer to the models ranked as number one for a given measure.

2.4.1 Q2.1: Model and Evaluation Measure Comparison

We refer to a model that most accurately predicts the in-sample estimation data according to a given measure as the "best estimative model", and we refer to a model that most accurately predicts the out-of-sample prediction data as the "best predictive model". The traditional goodness-of-fit measures— likelihood/KL and AIC/BIC— select the same best estimative model, and they also agree upon the specification of the best predictive model. The CDFET goodness-of-prediction measure selects distinct model specifications as the best predictive models dependent upon the desired error tolerance level (we test error tolerance levels of 25%, 50% and 75%). The three CDFET best predictive models are also distinct from the best estimative and predictive models under the AIC, BIC, and likelihood criteria. See Appendix D for selected model measure comparisons and coefficient estimates.

Though the best likelihood/AIC/BIC estimative model is distinct from the best predictive model, the difference in form is small. They include the same covariates but in different forms (e.g. operating cost as miles/dollar as opposed to gallons/mile) with the exception of luxury and transmission which contribute little to utility relative to the contribution of the other attributes.

Table 3 — Relative Average Likelihood (RAL) calculated on the prediction data set for select model specifications and data sets

Scenario	**1**	**2**	**3**	**4**	**5**	**6**
Estimation data	2004-2006	2006	2007	2004-2006	2004-2006	2004-2006
Prediction data	2007	2007	2007	2010	2007	2007
Market	Full market	Full market	Full market	Full market	Luxury sedan[1]	New designs[2]
AL of ideal model (predicted shares=actual shares)	*0.0076*	*0.0076*	*0.0076*	*0.0080*	*0.0384*	*0.4610*
RAL of no info model	55.3%	55.3%	55.3%	43.6%	63.6%	93.2%
RAL of static model	**88.3%**	**88.3%**	**88.3%**	23.7%	73.3%	95.9%
RAL of class dummies only logit	65.9%	65.9%	65.9%	53.6%	NA	95.0%
RAL of best fit logit model for L/AIC/BIC of estimation data	76.4%	77.7%	81.7%	67.3%	73.3%	96.5%
RAL of logit model with greatest likelihood for prediction data	79.0%	79.0%	81.7%	**68.5%**	87.9%	**97.5%**
RAL of mixed logit with best logit estimation fit covariates	79.0%	79.0%	85.6%	67.3%	**89.2%**	97.0%
RAL of nested logit with best logit estimation fit covariates	73.8%	72.4%	80.3%	64.8%	NA	95.3%

[1] Luxury sedan vehicles used for estimation and prediction
[2] Full market used for estimation, evaluation measures assessed for prediction of new vehicles only

2.4.2 Q2.2: Model Accuracy

Table 3 summarizes the average likelihood (AL) calculated on the prediction data set for select combinations of model specification (rows) and estimation/prediction data set scenarios (columns). We report the *relative* average likelihood (RAL) in Table 3 defined as the average likelihood of the model divided by the average likelihood of an ideal aggregate model that predicts shares perfectly. The reason we report RAL instead of simply AL is because choice diversity in the data necessarily lowers the maximum attainable value of AL with any model. Thus RAL describes the amount of predictive power obtained by a particular model relative to the best possible predictive power that could be obtained with any aggregate model.

The rows compare the predictive performance of the model that has the best predictions and the model that fit the estimation data best. Using each of the utility functions from the best estimative logit models, we fit additional mixed and nested logit models. 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 the selection of covariate form for the mixed and nested logit models. The "no info" row is calculated by assigning an equal share to all vehicles. The static model row assumes that shares in the prediction year are identical to the most recent share of the vehicle design available in the estimation data for all vehicle designs present in both the estimation and prediction data, and all new vehicle designs receive an equal proportion of the remaining share.

Scenario 1 is our base case, where models are fit to sales in years 2004-2006 and used to predict 2007 sales. Scenario 2 uses only 2006 data to predict 2007, assessing sensitivity of predictions to the amount of data used for estimation. Scenario 3 fits the models directly to 2007 data, helping to identify the portion of prediction error that stems from model fit, rather than from changes over time. Scenario 4 uses 2004-2006 data to predict 2010 sales, assessing differences when predictions are made farther into the future. Scenario 5 assesses predicative accuracy for a single vehicle class[5], rather than the entire market, and scenario 6 assesses only the predictive accuracy for new vehicle designs introduced in 2007. Comparisons can be made within each column to evaluate the prediction accuracy across model specifications for a given estimation/prediction data set.

[5] This is distinct from the "class dummies only logit" which includes data for the entire market but uses only dummies representing each class as covariates

In scenarios 1-3, which predict the full 2007 market, the best predictive logit model predicts better than the best estimative model, the class dummies model does not predict as well as the models which contain vehicle attributes, and the no info model predicts worst, as expected. Nested logit predictions have lower average likelihood than logit, but mixed logit predictions have higher average likelihood[6]. That the nested logit does not predict better than the logit suggests that the relaxation of the IIA property among the nests selected does not improve prediction. Model predictions could potentially be improved further by exploring alternative parameter distributional forms such as multivariate normal with a full covariance matrix [18], although that introduces more potential for overfitting with aggregate sales data. We leave such explorations for future work. See Appendix E for mixed and nested logit coefficient estimates and Appendix F for actual versus predicted shares.

In all three scenarios the static model outperforms all other models. Additionally, we see little difference in prediction quality between scenarios 1-3 when using the same model (compare across columns) compared to the difference due to model specification (compare down rows), even though the prediction set and estimation set are identical in scenario 3. Together, these results indicate that residual error in model fit is a major source of prediction error, and there is too much missing data or model misspecification in the attribute-based models to fit or predict the full market as well as the static model. Without data on missing covariates that influence choice, such as vehicle aesthetics, it is

[6] A likelihood ratio test of the best logit and mixed logit models calculated on 2007 data suggests that there is sufficient evidence to reject the null hypothesis that the mixed logit model predicts significantly better at the α=0.10 level.

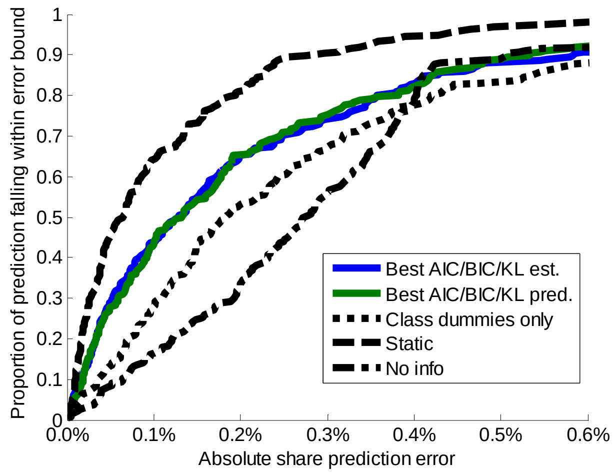
difficult to fully explain choice behavior at the vehicle design level with only the available covariates.

However, Scenario 4 examines a longer time horizon and reveals that the static model has poor predictive capability when forecasting farther into the future. The attribute-based models attempt to capture consumer choice as a function of observable attributes plus random noise, but since not all attributes are observed, share is not fit perfectly. In contrast, the static model does not attempt to explain the reason for consumers' choices but instead simply assumes consumers will make the same choices year after year. The static model does well for the 2007 forecasts because share for each vehicle model changes little from year to year, but over a longer time horizon vehicle designs change and new designs are added to the market (~37% of the vehicle designs sold 2010 did not appear in the 2004-2006 data). The static model has no information about these new designs, so it loses predictive capability, and over a longer prediction horizon the attribute-based models perform substantially better than the static model.

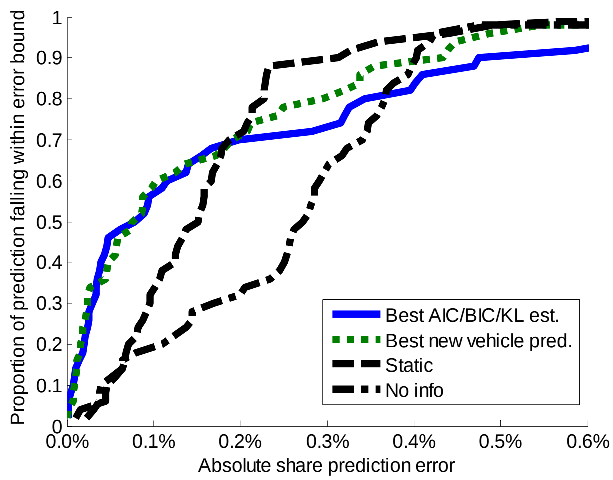
Scenario 5 indicates that the attribute-based models also perform better than the static model in the luxury sedan class. The best class model is distinct for each class, though all class models include some form of all covariates with the exception of style and automatic transmission as standard. The average likelihood of 2007 class predictions increases when the best estimative class level model is fit to class data as opposed to the best estimative full market model fit to the full market data with the exception of midsize and sports cars (see Appendix G for table of class model specifications and model average likelihood comparison by class).

Figure 1a shows the CDFET for selected models of scenario 1. The x-axis is the absolute difference between the predicted share and the actual share, and the y-axis is the proportion of vehicle designs whose share prediction error is less than the corresponding value on the x-axis. For example, in Figure 1a point (0.25%, 0.7) indicates that 70% of the share predictions made by the best AIC/BIC/KL models deviate from the observed share by less than 0.25% (the average vehicle design share in this market is 0.42%).

The worst models all perform similarly to one another in scenario 1 and lie on top of the class-only curve in Figure 1a (and are thus omitted for readability). While a model could plausibly be posed that predicts worse than the no-info model, we do not observe it in our utility specifications. 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 likelihood and AIC measures 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. No one covariate in isolation sets the best models apart from the worst models. A model's predictive power thus appears to be robust to covariate form but sensitive to the exclusion of attributes.



(A)



(B)

Figure 1 — CDF of error tolerance for the best logit model specifications as measured by likelihood/KL and AIC/BIC measures on 2004-2006 sales estimation data and 2007 sales prediction data compared to alternative models (A) full market (B) new vehicle designs only

2.4.3 Q2.3: Implications for Design

Scenario 6 compares the best-predictive logit model for all vehicles to the model that best predicts the shares of the new vehicle designs introduced in 2007. The best new vehicle model is determined similarly to the best predictive logit model of scenarios 1-3 by ranking the models on each of the measures; however, the measures in this case were calculated by treating each of the new vehicles individually and the holdover vehicles as an aggregated “other” share. (The “other” share is calculated as the sum of all holdover vehicle shares.) In contrast to scenario 1, the attribute-driven logit models of scenario 6 have a higher likelihood than the static model, since the static model has no information about new designs.

The CDFET of Figure 1b shows that at lower values of error tolerance the attribute-driven models are superior to the static model and that there is some difference in prediction quality between models that predict best for the whole market versus the new vehicle market. Overall, while the static model outperforms attribute-based models for near-term predictions, attribute-based models are needed for predicting the performance of new vehicle designs and for making longer-term predictions. Still, the degree of uncertainty and error in predictions for new designs may be too large to guide design choices appropriately in some contexts.

Appendix D summarizes model coefficients for several specifications including the representative of models in the literature as well as best-fit and best estimative models. It is clear that different specifications lead to different inferences about the effect of attribute changes on choice. For instance, the utility function specifications based on Boyd and Mellman [59], Berry *et al.* [58], and Whitefoot and Skerlos [31] result in a coefficient for operating cost that suggests consumers prefer higher efficiency (longer range per unit cost or lower fuel consumption per unit distance) all else being equal, as expected. But the best-fit and best predictive models suggest that consumers prefer lower efficiency. This can happen because efficiency may serve as a proxy for unobserved variables (e.g. size, performance, or styling variables not captured in the data). While the latter models make better predictions for existing vehicle markets that follow established patterns (attribute correlations), they could misguide design efforts that divert from established market patterns.

2.5 Limitations

Our investigation is a first step in a larger goal of characterizing the design impacts of choice prediction uncertainty. All of our models have error resulting from misspecification and missing information (as do all similar models in the literature that are based on market sales data rather than controlled experiments). For example, we do not have 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 [29], which can help explain choice behavior and improve predictions 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. Our study suggests that if models lack transparent quantifications of important determinants of product choices, designers should be cautious about basing design decisions on choice models.

More research is needed to assess a wider scope of modeling alternatives. We did not consider ASCs – product-specific factors that proxy for omitted variables – and their use in prediction or design. ASCs can generate models that match estimation data shares exactly; however, they contain no information about specific unobserved product features, and they are unknown for any new product designs. We also ignore a major component of the new vehicle modeling literature: covariate endogeneity— a correlation between model covariates and the unobserved terms like error. Endogeneity implies that coefficients are biased and inconsistent if not properly estimated, typically requiring instrumental variables techniques [23]. We also did not consider alternative estimation methods (e.g. Bayesian methods) and alternative heterogeneity specifications (e.g. latent class models, mixed logit model with joint parameter distributions, mixture models, and generalized logit models that account for scale and coefficient heterogeneity [18]).

Our study uses random utility DCMs that treat consumers as observant rational utility maximizers with consistent preferences. While this is a popular approach to modeling consumer choice, important criticisms exist. For instance, preferences can evolve over time [41], changing with cultural symbolism [81] and/or social interactions [82]. The theory of construction of preference adapted to design by MacDonald *et al.* suggests that consumers' preferences for attributes do not exist a-priori but are rather evaluated on a case-by-case basis [32]. Morrow *et al.* [30] suggest that vehicle choice behavior may be better represented by a "consider-then-choose" model [83] where consumers first screen out most alternatives using simple rules, subsequently maximizing utility over a smaller "consideration set" [84]. The potential value of this type of model is suggested here by the better performance of class-only models, a special case of the consider-then-choose model. More broadly, the Lucas critique warns against use of aggregated historical data to predict outcomes in counterfactual future scenarios [85].

2.6 Conclusions

While the topic of uncertainty associated with choice predictions is widely discussed in the design community (e.g. [3,24–26,32,44]), there is no current consensus as to what processes and measures best quantify model uncertainty. This gap motivated our first research question, Q2.1. We investigated several well-known measures of model performance evaluated on a prediction set. For the automotive case study examined, likelihood measures (and the rank-equivalent Kullback-Leibler divergence measure) tend to identify the same top-ranked model as the penalized likelihood measures AIC and BIC do. While CDFET measures identify different top-ranked models, depending on the error tolerance selected, the resulting models share most covariates. Models that perform well on one measure tend to perform well on the other measures, and models that perform poorly on one measure also tend to perform poorly on the other measures. In other words, determination of the best models in our study did not depend strongly on potentially arbitrary selection of the measure used to evaluate predictive accuracy.

Overall our results confirm several intuitive features of this application: attribute-based models predict better than models with no information; models of a particular vehicle class typically make better predictions than models of the full market; including more covariates generally improves predictive accuracy; and better model fit correlates well with better predictive accuracy. The match between fit and predictive accuracy, suggesting no major overfitting issues, is particularly encouraging, since the modeler has access to choice data for estimation but not choice data in the counterfactual predictive context. These findings would have to be validated in other product domains on a case-by-case basis.

We also observe a number of less intuitive results that are relevant to design. First, the models we construct are fairly poor predictors of future shares. In our base scenario, our best predictive model has an average error of 0.24% (the average share of a vehicle design is 0.42%), which translates to an error of approximately 37,500 vehicles sold for the 2007 market. The limited predictive power of standard models on real data in a canonical product category suggests designers should apply DCMs cautiously, though predictions may be substantially better in domains with fewer unobserved attributes or with conjoint data (where all attributes are observed).

Second, we find that attribute-based models do not furnish the best predictions for short-run forecasts in stable market conditions: Attribute-based models estimated on 2004-2006 data were outperformed in predicting 2007 shares by the "static" model that assumes no changes in shares. However, attribute-based models are superior to the static model when predicting new vehicles only, since the static model lacks information about new entrants. There are some intuitive reasons why the static model might perform better than attribute-based models for short term predictions of existing designs given relatively stable market conditions. First, the static model may implicitly capture effects related to omitted vehicle attributes neglected by attribute-based models. Second, the static model may predict well in the short-run simply because of "inertial" conditions specific to the automotive market, particularly multi-period production schedules and inventory build-up that must ultimately be cleared over the short run using unobserved advertising and/or purchasing incentives.

Third, while including an appropriate set of product attributes as model covariates is important to improving predictive accuracy, the form those covariates take in the utility function is less important in this application. This implies that it may be less important to test many variations of utility function covariate form when constructing a model, but it also means that any design decisions (e.g. design optimization results) that are not robust to variation in utility function covariate form may not be justified given the near equivalence of alternative covariate form in fit and prediction error with market data. If different utility specifications lead to different design decisions but the data cannot discern which form best represents choices, then design decisions cannot be reliably based on any single specification.

Finally, we observe that some of the models with the best predictive accuracy have coefficients with unexpected signs – likely biased due to correlation with unobserved attributes. Despite good prediction accuracy in existing markets, where attribute correlations are similar from year to year, these models may misguide design efforts if the designer makes changes that do not follow correlations in the marketplace. For example, the sign of the coefficient for the gallons per mile (gpm) attribute of the best predictive logit model is negative[7], suggesting that consumers prefer lower fuel economy, all other attributes being equal. In fact, consumers may purchase vehicles with lower fuel economy because of other features of those vehicles unobserved by the modeler (e.g. size, performance, or styling attributes not captured in the model). The model predicts well if the new market retains such correlations, but a designer who lowers fuel economy alone is not likely to obtain the outcome predicted by the model. Thus, accuracy of predictions in existing markets is not a sufficient condition for use in design.

To verify that our results are not specific to the 2004-2006 timeframe, we conducted a similar analysis with estimation data from years 1971-1973 and 1981-1983 with prediction data from the respective 1 and 4 year forward markets. We find that our conclusions are generally robust to alternate timeframes: our accuracy measures are concordant; the best models exhibit substantial prediction error stemming from limited model fit; the static model outperforms the attribute-based models when predicting the

[7] We reject the null hypothesis that the coefficient is equal to zero at the α=0.01 level for a two-sided t-test.

full market 1-year-forward but attribute-based models can predict better for 4-year-forward forecasts or new vehicle designs; share predictions are sensitive to the presence of utility covariates but less sensitive to covariate form; nested and mixed logit specifications do not produce significantly more accurate forecasts; and the 1971-1973 models with best predictions do not necessarily have expected coefficient signs (though 1981-1983 models do). See Appendix H of the supplemental material for additional detail.

3. Improving Forecasts for Light-Duty Vehicle Demand: Assessing the Use of Alternative-Specific Constants for Endogeneity Correction versus Calibration

This second study covers two of the major thesis research themes: comparing model specifications and forecasting methods in terms of prediction accuracy and comparing design implications of different model specifications and forecasting methods. We investigate the following research questions: How should modelers address the potential endogeneity between price and omitted variables when forecasting new vehicle market shares? Can alternative-specific constants (ASCs) improve forecast accuracy? How should estimates of past ASCs be used in future share forecasts? Of the assumptions and models in Table 1 this study examines the effects on estimated coefficients and share predictions when the ASC is assumed to be correlated with price (endogenous) as opposed to when it is treated as independent. We estimate mixed logit models with independent coefficients[8].

We investigate the implications of different approaches to incorporating ASCs in the product utility function on parameter recovery and forecast accuracy. We do so for a specific application, i.e. light-duty vehicle choices in the United States. Specifically, we estimate mixed logit models using both synthetic data and 2002-2006 US midsize sedan automotive sales data. We test two methods of incorporating ASCs: (1) a maximum likelihood estimator that computes ASCs post-hoc as model calibration constants (MLE-C) and (2) a generalized method of moments estimator with instrumental variables (GMM-IV) that uses ASCs to account for price endogeneity.

---

[8] This chapter is included in a working paper. In this chapter the use of first person plural includes coauthors Ross Morrow, Inês Azevedo, Elea Feit, and Jeremy Michalek.

From the synthetic data study we observe that MLE-C coefficient bias increases as the price-ASC correlation (endogeneity) increases, as expected. Given valid instruments, GMM-IV successfully mitigates this bias at the cost of larger parameter variance; but, given invalid instruments, GMM-IV exacerbates bias. One- and five-year MLE-C forecasts are as good as or better than GMM-IV forecasts regardless of instrument validity except when a high price-ASC correlation present in the estimation data is absent in the forecast data.

In the market data study the GMM-IV model better predicts the 1-year-forward market (4% accuracy improvement over MLE-C), but the MLE-C model better predicts the 5-year-forward market (15% accuracy improvement over GMM-IV). Including an ASC in predictions by any of the methods proposed improves share forecasts, and assuming that the ASC of a new vehicle is most similar to the ASC of its closest competitor vehicle yields the best long term forecasts. We find evidence that the instruments most frequently used in the automotive demand literature are likely invalid.

3.1 Introduction

Discrete choice models (DCMs) are used to interpret and forecast product demand in a variety of contexts, and a popular application is the new vehicle sales market. In particular, the automotive literature employs DCMs to understand drivers of purchase behavior [34,37,39,46,53,60] and predict future vehicle market shares [4,5,14,15,50,86]. DCM specifications include popular multinomial logit, nested logit, mixed logit, and probit models [23], as well as variants of these models, such as the generalized multinomial logit model [18].

DCMs of product purchases are generally estimated using either stated choice data or revealed preference data. Stated choice data can be obtained from choice-based conjoint experiments where the modeler selects a set of attributes that the respondent observes. These studies can avoid issues such as omitted variables, endogeneity, and multicollinearity. However, such studies typically rely on surveys that do not directly measure real purchase choices in a market context. In contrast, revealed preference data, often aggregate market sales data, measures real market purchases. Revealed preference studies have the limitation that buyers evaluate factors that are unobserved by the modeler or are difficult to represent mathematically (e.g. aesthetics). Also, attributes tend to be correlated among product alternatives in the marketplace (e.g. vehicles with luxury features routinely have higher prices), sometimes introducing multicollinearity [87]. Different sets of econometric assumptions are needed for the stated and revealed preference modeling approaches. In this work we focus on models constructed from revealed preference (sales) data.

The number and nature of attributes considered by consumers in a vehicle purchase decision is sufficiently large and complex that any DCM posed will be missing information about some of the attributes that inform consumer choice. In order to address the utility not captured by explanatory variables, modelers may include an alternative-specific constant (ASC) in the utility function. For example, following Train [60]:

$$u_{ijt} = \mathbf{x}'_{jt}\boldsymbol{\beta}_i + \xi_{jt} + \varepsilon_{ijt} \tag{11}$$

where u_{ijt} is the utility consumer i derives from product j in market (year) t, $\mathbf{x}_{jt}$ is a vector of attributes specific to product j in market t, $\boldsymbol{\beta}_i$ is a vector of taste parameters for consumer i, ξ_{jt} is the ASC for product j in market t, and ε_{ijt} is an idiosyncratic error term. Consumer-specific attributes like income or family size can also be included in the utility function, but they are not in this study since we use aggregate data sets for which this information is not available.

An interpretation of the ASC introduced into the automotive demand context by Lave and Train [60][9] and popularized by Berry *et al.* [58] is that it represents the mean cumulative effect of all product attributes that consumers use to evaluate a product but that are unknown to researchers. Alternative terms for the ASC when it is used to represent omitted variables include the unobservable [34,53,54,58,88], the unobserved product characteristic or attribute [36,49], market-level disturbance [52], and demand shock [33,35]. However, the ASC need not necessarily be viewed as a representation of unobserved attributes but rather can be included as purely mathematical construct to improve model fit [15,50] and sometimes referred to as a calibration constant [62].

In the literature, the treatment and interpretation of the ASC differs depending on whether the focus of the research is to *forecast* future vehicle demand shares (i.e. the "predictive" literature) or the focus is on measuring the importance of attributes to consumers (i.e. the "explanatory" literature), especially as it pertains to willingness-to-pay and price elasticities of demand. The predictive literature generally obtains ASCs by estimating coefficients in a model that excludes the ASCs and then "calibrating" the model post hoc by choosing values for the ASCs so that the modified model-predicted shares of the estimation data match observed shares. In contrast, the explanatory literature is primarily concerned with coefficient estimation and thus views it as imperative to

[9] Lave and Train [60] estimate a disaggregate model of vehicle choice in which all observed vehicle attributes are interacted with consumer attributes, e.g. income, so that the ASC is identical to the unobserved vehicle-specific utility. However, as in Train and Winston [47], the ASC refers to the mean utility derived from both the observed and unobserved vehicle attributes. In the aggregate demand model context here the ASC refers only to the unobserved portion of utility.

address potential sources of coefficient inconsistency— specifically price endogeneity [58]. Inconsistency arises if the ASC is correlated with an observed attribute. If the ASC is interpreted as a representation of aggregate utility from unobserved attributes, then it is plausible that observed and unobserved vehicle attributes (e.g. price and aesthetics) are correlated for markets in which prices are set by strategic firms, which would asymptotically bias[10] the coefficient of the observed attribute away from the true value. Though the true population taste parameters are unknowable for real data, researchers have demonstrated that for models estimated on actual market data the price coefficient bias, measured as the difference between estimates when endogeneity is ignored versus when it is corrected for, is in expected directions [58,89,90]. The explanatory literature implements estimation techniques that mitigate endogeneity— typically using instrumental variables (IVs) and estimating the ASC simultaneously with the coefficients.

There are drawbacks to mitigating coefficient bias with IVs. Model estimation is challenging in part because *valid* instruments are difficult to specify and impossible to verify as demonstrated by Rossi [91]. Valid instruments require that they are correlated with the endogenous observed vehicle attribute, uncorrelated with the unobserved attribute(s), and do not affect the dependent model variable (market share) except through the observed attributes [92]. Instruments that do not meet these conditions are termed invalid. Because these properties are difficult to satisfy in many situations, instrument selection is somewhat subjective and ad hoc, and, as we show, the "wrong" choices can

[10] Methods that incorporate (valid) IVs result in estimators that are consistent— as the data sample size goes to infinity the expected value of the estimator converges to the true value of the parameters should they exist— but not unbiased in the sense that the sampling distribution of the estimator is centered on the true value of the parameters (also termed "finite sample bias"). In the literature discussed here "bias" is shorthand for "asymptotic bias" and is used interchangeably with "inconsistency." [37]

generate worse models than ignoring endogeneity all together. Specifically, Rossi shows that invalid instruments can lead to coefficient estimates exhibiting larger bias than those obtained *ignoring* endogeneity [91]. Furthermore, a common estimation technique used to incorporate IVs (generalized method of moments or GMM) is known to be inefficient so that a large number of instruments are needed to obtain statistically significant coefficient estimates [92].

Even if valid instruments can be specified, asymptotically biased coefficients do not necessarily mean a model will predict poorly [63]. Particularly, biased models may predict future choices well in markets that have persistent patterns of endogeneity because bias may carry forward information about the connection between observed and omitted variables— a useful feature when the relationships that hold in the estimation data are also likely to be hold in the future [93]. In the case of automotive demand modeling, endogeneity comes from strategic pricing by firms that observe and account for demand for vehicle aspects not observed by the modeler. It is thus plausible that price-ASC correlation persists in future markets absent dramatic changes in market structure, raising doubts as to whether correcting for endogeneity is needed.

This work is motivated by the preceding discussion and aims to answer the following three questions:

(Q3.1) Should modelers address the potential endogeneity between price and omitted variables when forecasting new vehicle market shares?

For contexts in which the source of endogeneity is likely to persist— e.g. mature product markets with relatively stable consumer preferences like the automotive market— models may be well served by calibration constants free from the problems with specifying valid IVs. The main drawback of post-hoc ASC calibration is that it lacks the coherent theoretical grounding of the IV approach: the effect of missing attributes on choice is attributed to observed attributes when fitting the model, and the ASCs are added post hoc to “artificially” match observed shares. Forcing model predictions to match observations exactly guarantees overfitting— fitting the noise in the data— which should degrade forecast quality. This suggests that use of ASCs as calibration constants in forecasts should be approached with caution and motivates our second research question:

(Q3.2) Can ASCs improve forecast accuracy?

Several studies implement methods to predict ASCs for out-of-sample alternatives [15,47,49,62], but neither the predictive nor explanatory literature suggests how to predict future values of the ASCs for products that do not appear in the estimation data. Furthermore, these prior studies represent predicted ASCs as point estimates without characterizing the implications of uncertainty in the value of ASCs for future alternatives. We propose four methods for forecasting ASCs when they are interpreted as representing the utility of unobserved vehicle characteristics and evaluate the resulting accuracy and uncertainty of predicted shares when ASCs are included in the utility function in order to investigate our third research question:

(Q3.3) Should estimates of past ASCs be used in future share forecasts?

Our work examines the implications of including ASCs in DCMs in an automotive demand modeling context. Specifically we are interested in the implications of estimation and prediction techniques employed in the literature on forecast accuracy and uncertainty. Much of the literature relies on these models to inform analysis of vehicle design and policy, and these types of analyses often require forecasting.

The remainder of this paper is structured as follows: Section 3.2 reviews the literature on DCMs that employ ASCs with a focus on the automotive demand literature. Section 3.3 describes model specifications and estimation techniques. Section 3.4 outlines the generation of a synthetic data set and the methods used to predict future ASCs as well as estimation and prediction results for the synthetic data set. Section 3.5 extends the analyses to a case study on midsize sedan purchases for 2002-2006, 2007, and 2011. Section 3.6 discusses the studies' results and challenges posed by estimating a model with IVs. Section 3.7 addresses the limitations, and section 3.8 concludes.

3.2 Literature Review

Table 4 compares automotive demand studies that use DCMs with ASCs and are estimated in a classical framework (as opposed to Bayesian). The explanatory literature primarily introduces or evaluates model estimation techniques, estimates coefficients to describe consumer preferences or firm behavior, or tests counterfactual policy scenarios and often reports willingness-to-pay or willingness-to-accept for different vehicle attributes. For some of these purposes no predictions are made, but counterfactuals are simulated; such simulations are typically in-sample. For such analyses there is often no need to determine ASCs, which are assumed identical to their estimated values. (Berry *et al.* [49] and Train and Winston [47] are exceptions.) The forecast literature primarily forecasts future market shares or tests counterfactual scenarios. Reviewing the literature summarized in Table 4 a major distinction between the two bodies of literature emerges. The explanatory literature estimates models by formal, econometric methods that often use IVs to obtain consistent coefficients, reducing or eliminating (asymptotic) coefficient bias and estimating ASCs concurrently with observed variable coefficients. The forecast literature relies on expert opinion or historical coefficient estimates, calibrates ASCs post model estimation, and does not take additional steps to control for endogeneity.

While there are many reasons to include ASCs in DCMs, they do present a practical problem in forecasting. Typically, researchers will assume that estimated values of the ASCs for a particular vehicle model carry forward from year to year, but when the forecast market includes new vehicles, an assumption must be made about the value of the ASC for that new vehicle. We review new vehicle ASC forecasting methods from the literature and use these to inform our proposed methods in sections 3.4 and 3.5. Four of the studies in Table 4 predict out-of-sample or new vehicle shares for which unknown ASCs must be generated or assumed. Berry *et al.* 2004 [49] predict in-sample shares for a counterfactual scenario, but they introduce two new vehicles into the data set. For the new vehicles, the ASCs are generated by averaging the estimated ASCs of the respective brand and class of the new vehicle. Train and Winston [47] forecast out-of-sample future market shares. They hold the ASCs of the observed products constant from the calibrated value and similarly to Berry *et al.* [49] generate new product ASCs by averaging over the estimated ASCs of the same class. Greene *et al.* [15] forecast out-of-sample shares for future markets in which the vehicle set is modified by introducing diesel and hybrid versions of vehicles in the estimation data set. They assume diesel and hybrid vehicles have identical ASCs to their conventional counterparts and hold the ASCs constant from the estimated values. Bunch *et al.* [62] forecast out-of-sample future market shares under alternative policies and assume that the vehicle set is unchanged from the estimation data. Similarly to Greene *et al.* [15], they hold future ASCs constant from the estimated values.

One of our model estimation techniques described in section 3.3 is based on the two-stage generalized method of moments with instrumental variables (GMM-IV) estimation implemented by Berry *et al.* [58] that is commonly referred to as "BLP". Berry *et al.* [58] develop a technique to estimate joint models of supply and demand that include ASCs and IVs in a nonlinear DCM framework. This canonical study informs many of the studies in Table 4 [33–37,39,45,47,49,52–54] and Knittel and Metaxoglou [33] list BLP-type demand models in addition to the vehicle demand focused studies included here. We isolate the demand side model and ignore the supply side as described in Nevo [67] who provides a detailed econometric guide on the technical details of estimating a BLP-like model for interested readers. However, several recent studies have focused on the difficulties of estimating models with the BLP method. Knittel and Metaxoglou [33] find that for several combinations of algorithms and starting values the estimation routine results in specious convergences and convergences to several local minima; moreover they find that different local GMM-IV solutions have significantly different economic implications. Dubé *et al.* [35] find that loose convergence tolerance criteria on BLP's nested fixed point iteration exacerbate coefficient bias and that reasonable estimation times often require a convergence criterion too loose for good estimates. They propose an alternative nonlinear programming formulation that we adopt here. Similarly Su and Judd [94] propose an alternative formulation of the BLP-style estimation procedure for more general strategic market models to avoid nested fixed point iteration. We discuss some other difficulties vehicle demand researchers are likely to encounter in section 3.6.3.

No studies in Table 4 conduct a formal statistical test regarding the appropriateness of the IVs, but two of the studies qualitatively address it. Allcott and Wozny [34] verify that their instruments should not be included as explanatory variables in the utility function, and Li *et al.* [39] estimate a model that implies inelastic price elasticities, suggesting that their BLP-like instruments are invalid. Though not included in Table 4 because the model is a linear regression (as opposed to a DCM), Jenn *et al*. [95] examine the effects of federal policies on hybrid vehicle sales and conduct a J Hansen test to verify that the model is not overspecified.

Most of the studies in Table 4 estimate models on aggregate data only. As exceptions, [36,49,52,62] combine aggregate and disaggregate data, and [46,47,60,96] use disaggregate, household data only.

Table 4 — Automotive demand literature that includes an alternative-specific constant in the discrete choice model

Study	Year	Specification[1,2]	Estimation[3]	Instruments[4]	Model purpose: Estimation technique or model proposal / investigation	Model purpose: Attribute valuation / market description	Model purpose: Counter-factual / policy evaluation	Model purpose: Forecast future market shares
Explanatory literature— ASC is estimated simultaneously with taste parameters								
Knittel and Metaxoglou [33]	2013	Mixed logit	Two stage GMM	BLP	x			
Allcott and Wozny [34]	2012	Nested logit	NFP + 2SLS	Fuel price[5]		x		
Dubé *et al.* [35]	2012	Mixed logit	Two stage GMM, GMM MPEC	Synthetic[6]	x			
Beresteanu and Li [36]	2011	Mixed logit	Two stage GMM	Fuel price[5]			x	
Copeland *et al.* [37]	2011	Mixed logit	Two stage GMM	BLP		x		
Li *et al.* [39]	2011	Logit	NFP + 2SLS	BLP		x		
Vance and Mehlin [45]	2009	Nested logit	NFP + 2SLS	BLP			x	
Dasgupta *et al.* [46]	2007	Nested logit	MLE	None		x		
Train and Winston [47]	2007	Mixed logit	Two stage MLE	BLP		x		
Berry *et al.* [49]	2004	Mixed logit	Two stage GMM[7], Expert elicitation	Price makeup/None	x			
Petrin [52]	2002	Mixed logit	Two stage GMM	BLP	x			
Sudhir [53]	2001	Mixed logit	Two stage GMM	BLP		x		
Berry *et al.* [54]	1999	Mixed logit	Two stage GMM	BLP			x	
Berry *et al.* [58]	1995	Mixed logit	Two stage GMM	BLP	x			
Lave and Train [60]	1979	Logit	MLE	None		x		
Predictive literature— ASC is calibrated post-estimation of model taste parameters								
Whitefoot and Skerlos [49]	2012	Logit	Literature informed	None			x	
Bunch *et al.* [62]	2011	Nested logit	Canned software	None			x	
US EIA [5]	2011	Nested logit	Expert elicitation	None				x
Greene *et al.* [50]	2005	Nested logit	Expert elicitation	None				x
Greene *et al.* [15]	2004	Nested logit	Expert elicitation	None				x
Unknown/other								
Choo and Mokhtarian [96]	2004	Logit	Canned software	None		x		

[1]More than one model may be specified, model listed is the discrete choice model or study focus relevant to this work; [2]Mixed logits are independent mixed logits for all studies listed, logit is multinomial logit; [3]GMM=generalized method of moments, MLE=maximum likelihood estimation, 2SLS=two staged least squares, NFP=nested fixed point, IV=instrumental variable; [4]"BLP" refers to the instruments used in Berry et al. [58] or a similar variant; [5]Allcott and Wozny [34] use the vehicle's expected lifetime fuel costs (applicable when used vehicle sales are included in the model) and Beresteanu and Li [36] use fuel costs in other Metropolitan Statistical Areas (MSAs); [6]Synthetic data study, instruments are generated; [7]First stage of GMM inverts shares to obtain mean utility, three methods for mean utility parameters: expert elicitation, IV regression, regression assuming no endogeneity

3.3 Methods

We are primarily interested in comparing the accuracy and uncertainty associated with forecasting using a random coefficients logit model with independent and normally distributed coefficients ("mixed logit" here for brevity) when potential endogeneity due to the correlation of the ASC with price is mitigated using IVs as opposed to when it is ignored. We estimate choice models using both maximum likelihood estimation with calibration (MLE-C) and GMM-IV methods, and we compare the out-of-sample market share predictions resulting from the two estimation methods under several proposed techniques for forecasting the new market ASCs.

3.3.1 Choice Model

The assumed utility function is linear in coefficients and includes an ASC ξ:

$$u_{ijt} = \mathbf{x}'_{jt}\boldsymbol{\beta}_i + \xi_{jt} + \varepsilon_{ijt} \tag{11}$$

where u_{ijt} is the utility consumer i derives from product j in market (year) t, $\mathbf{x}_{jt}$ is a vector of attributes specific to product j in market t, $\boldsymbol{\beta}_i$ is a vector of taste parameters for consumer i, ξ_{jt} is the ASC for product j in market t, and ε_{ijt} is an idiosyncratic error term. Note that we use the terms market and year interchangeably, although they have different implications in some contexts. We assume the attribute vector $\mathbf{x}_{jt}$ for a given product to be similar but not necessarily constant over time or across markets. For example, the price or weight of a Ford Focus may vary slightly from year to year though it is still considered the same product[11].

---

[11] In the case study data the greatest year over year non-price attribute change is less than 15%.

We assume that the distribution of preferences is constant in time or across markets such that β is indexed only by i and not t. (This is a standard assumption in the vehicle demand literature. See Axsen *et al.* [36] for an exception.) If the coefficients are assumed to be independently and normally distributed:

$$\boldsymbol{\beta} \sim N\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\beta}\right) \tag{12}$$

where $\boldsymbol{\mu}$ is a (K×1) mean vector and $\boldsymbol{\Sigma}_{\beta}$ is a (K×K) diagonal covariance matrix with (K×1) diagonal element vector $\boldsymbol{\sigma}^2$, then the portion of the utility function dependent on product attributes can be expressed as the sum of the deterministic mean utility common to all consumers and the stochastic consumer-specific utility:

$$u_{ijt} = \left(\mathbf{x}'_{jt}\boldsymbol{\mu} + \xi_{jt}\right) + \left(\mathbf{x}'_{jt}\left(\boldsymbol{\sigma} \circ \mathbf{v}_{i}\right) + \varepsilon_{ijt}\right) \tag{13}$$

where $\mathbf{v}_i$ is a (K×1) vector of independent standard normal random variables and the open circle ∘ is the Hadamard product (element-wise product). If ε_{ijt} is assumed to follow an independent and identically distributed (iid) extreme value type I distribution, the probability P_{ijt} of individual i selecting a product j in market t is then given by the logit probability:

$$P_{ijt} = \frac{\exp\left(\mathbf{x}'_{jt}\boldsymbol{\mu} + \xi_{jt} + \mathbf{x}'_{jt}\left(\boldsymbol{\sigma} \circ \mathbf{v}_{i}\right)\right)}{\sum_{k \in J_t} \exp\left(\mathbf{x}'_{kt}\boldsymbol{\mu} + \xi_{kt} + \mathbf{x}'_{kt}\left(\boldsymbol{\sigma} \circ \mathbf{v}_{i}\right)\right)} \tag{14}$$

where J is the set of V distinct products observed across all markets and J_t is a subset of J containing the products that appear in market t. Our model considers only the set of consumers who purchase products, thus there is no outside good (option to not purchase any product). When an outside good is excluded, the choice probabilities (Eq. 14) are invariant over uniform shifts in the ASCs. In order to enforce uniqueness of the ASCs we assume throughout (and constrain in estimation) $\sum_{k \in J_t} \xi_{kt} = 0 \ \forall t$.

The share P_{jt} of product j in year t can be obtained by integrating over the consumer-specific stochastic utility:

$$P_{jt} = \int_{\mathbf{y}} \frac{\exp\left(\mathbf{x}'_{jt}\boldsymbol{\mu} + \xi_{jt} + \mathbf{x}'_{jt}\left(\boldsymbol{\sigma} \circ \mathbf{y}\right)\right)}{\sum_{k \in J_t} \exp\left(\mathbf{x}'_{kt}\boldsymbol{\mu} + \xi_{kt} + \mathbf{x}'_{kt}\left(\boldsymbol{\sigma} \circ \mathbf{y}\right)\right)} f_{\mathbf{v}}\left(\mathbf{y}\right) d\mathbf{y} \tag{15}$$

where $f_{\mathbf{y}}(\mathbf{y})$ is the multivariate standard normal density function, all v are iid, and the integral is a K-dimensional integral. We approximate this integral using numerical integration with sample averages [23]. One hundred Halston draws of **v** are used and held constant throughout estimation. As discussed in Dubé *et al*. [35], other integral approximation methods are available, but they are not the focus of this work.

3.3.2 Estimation

The choice of estimator is determined almost entirely by the assumptions regarding the endogeneity of price and the ASC. If it is assumed that price exogenous, then estimating the model in Eq. 15 *excluding* the ASC by MLE and then calibrating the ASCs post-hoc (MLE-C) should yield a consistent price coefficient estimate (assuming all other attributes are also exogenous and there is no model misspecification). However, if it is assumed that price is endogenous, then GMM-IV is the estimation strategy preferred by the literature for incorporating IVs into the estimation routine in order to mitigate the potential price coefficient bias[12]. The presence of endogeneity (or

[12] MLE methods can be used to incorporate IVs, however, GMM-IV is the most common estimation approach. Train and Winston [47] and is an exception. Park and Gupta [48] propose a method of estimating

assumption thereof) does not itself determine the estimator, but rather the techniques used to address endogeneity— in this case IVs— suggest one estimation method over the other.

Maximum Likelihood with Post Hoc Calibration

The likelihood at the estimated parameters L is defined as the probability of generating the observed data given the estimated parameter values:

$$L\left(\hat{\boldsymbol{\beta}} \mid \mathbf{X}\right) = \prod_{t=1}^{T}\left(\prod_{k \in J_t}\left(P_{kt}\right)^{n_{kt}}\right) \tag{16}$$

where $\mathbf{X}$ is the (V×K) stacked matrix of transposed attribute vectors $\mathbf{x}_{jt}$ for all products in all markets, n_{kt} is the observed sales of product k in time t, T is the total number of markets, and P_{jt} in Eq. 15 is modified to exclude ξ_{jt}. The MLE estimator of the parameters $\widehat{\boldsymbol{\beta}} = [\widehat{\boldsymbol{\mu}}', \widehat{\boldsymbol{\sigma}}']'$ is the value of the (2K×1) vector that maximizes L. The monotonic transformation ln(L) is typically used as the objective function for computational benefit. The ASCs are "calibrated" [23] post-hoc by solving the system of equations:

$$\begin{aligned} &\ln\left(P_{jt}\left(\boldsymbol{\xi}_t \middle| \mathbf{X}_t, \hat{\boldsymbol{\beta}}\right)\right) = \ln\left(s_{jt}\right), \forall j \in J_t^-, t \\ &\sum_{k \in J_t} \xi_{kt} = 0, \forall t \end{aligned} \tag{17}$$

where $\boldsymbol{\xi}_t$ is the stacked vector of all ξ_{jt} in market t, $\mathbf{X}_t$ is the matrix of product attributes for all products in market t, s_{jt} is the observed share of product j in market t, P_{jt} includes

observed coefficient parameters and the ASC simultaneously using MLE but do not use automotive data sets.

the ASC as written in Eq. 15, and J_t^- is the set of products in market t excluding one[13]. For more detail on mixed logit models and MLE-C see Train [23].

Generalized Method of Moments with Instrumental Variables

For GMM-IV, following BLP [58] we specify the moment conditions:

$$E\left[\xi_{jt} \middle| \mathbf{z}_{jt}\right] = 0, \forall j, t \tag{18}$$

where $\mathbf{z}_{jt}$ is an ($L\times 1$) vector of instruments for product j in market t and $L \geq 2K$ as a necessary condition for identification [37]. The choice of these instruments is determined by the modeler and is specific to the data. We describe the instruments used for the synthetic data example and the case study in their respective sections.

Define $\tilde{\boldsymbol{\beta}} = [\tilde{\boldsymbol{\mu}}', \tilde{\boldsymbol{\sigma}}']'$ and $\boldsymbol{\xi}$ as the stacked vector of ξ_{jt} for all products and markets. The GMM-IV estimator of the parameters is the value of the vector $[\tilde{\boldsymbol{\beta}}', \tilde{\boldsymbol{\xi}}']'$ that solves:

$$\begin{aligned} &\underset{\beta,\xi}{\text{minimize}}: \left(\mathbf{Z}'\xi\right)' \mathbf{W}\left(\mathbf{Z}'\xi\right)/T \\ &\text{subject to: } \ln\left(P_{jt}\left(\beta,\xi_t\right)\right) = \ln\left(s_{jt}\right), \quad \forall j \in J_t^-, t \\ &\qquad\qquad \sum_{k\in J_t} \xi_{kt} = 0, \forall t \end{aligned} \tag{19}$$

where $\mathbf{Z}$ is the ($V\times L$) matrix of stacked vectors for all $\mathbf{z}_{jt}'$, $\mathbf{W}$ is an ($L\times L$) weighting matrix, T is the total number of markets, and P_{jt} is defined in Eq. 15. We set $\mathbf{W} = (\mathbf{Z}'\mathbf{Z})^{-1}$ as used by Dubé *et al.* [35][14]. Eq. 19 was not literally proposed by BLP [58], but it is a more recent interpretation of the model that has better statistical and computational

---

[13] The exclusion of one ASC from each market is for computational purposes. The selection of which ASC to exclude (set to zero) in each market is arbitrary.

[14] Other choices of **W** or more instruments can improve the statistical efficiency of the estimator [37], though alternative formulations may be more computationally burdensome [35]. Nevo [67] argues in favor of the less efficient but less statistically burdensome weighting matrix used here for BLP type models. For logit and nested logit models with homoskedastic errors $\mathbf{W}=(\mathbf{Z}'\mathbf{Z})^{-1}$ is the optimal weighting matrix [67].

properties [35,94]. We further transform Eq. 19 to improve computational properties. See Appendix I for explicit formulation of the optimization problem provided to the solver. For more information on GMM estimation and IVs see Wooldridge [75].

3.4 Simulation Study

We begin our investigation of forecasting with ASCs using synthetic data sets. This allows us to evaluate the statistical properties of the predictions in a controlled setting for comparing model coefficients and predictions. We generate five years of estimation data and two years of prediction data representing 1-year- and 5-year-forward forecast horizons. Each year of data has 70 products with 40% year over year turnover (40% of products each year are newly introduced replacing 40% from the previous year). These parameters were chosen to closely mirror the structure of the real data in the case study. The estimation data is generated with low, base, and high price-ASC endogeneity and a mixed logit model is estimated by MLE-C and GMM-IV using valid and invalid instruments. The potential level of endogeneity and the validity of the instruments are both unknowns for real forecasting exercises, and we would like to know their possible implications. The prediction data is generated under two market futures— one for which the source of price-ASC endogeneity persists and one in which it does not. This is again unknown in forecasting. We compare the accuracy and uncertainty of predictions made by the MLE-C and GMM-IV estimated models.

3.4.1 Synthetic Data Generation

An initial market is assigned 70 products, each with a randomly generated price attribute p, a single non-price or "technology" attribute x (for simplicity), and an unobserved contribution to utility ξ that is correlated with prices. Instruments z are generated simultaneously with the price and technology attributes so that we are able to specify any arbitrary correlation between them.

The product attributes and instruments for year t=1 are generated by drawing from a multivariate normal distribution:

$$\left[p_{jt}, x_{jt}, \xi_{jt}, z_{jt}^{(1)}, z_{jt}^{(2)}, z_{jt}^{(3)}\right]' \sim N\left(\mathbf{0}, \mathbf{\Sigma}_x\right), \forall j, t \tag{20}$$

Exogenous attributes (meaning those uncorrelated with the ASC) and functions of the instruments in the case of nonlinear models can also be used as instruments [91]. We adopt this approach and generate additional instruments:

$$\begin{aligned} \mathbf{z}_{jt}^{*} &= \left[z_{jt}^{(1)}, z_{jt}^{(2)}, z_{jt}^{(3)}\right] \\ \mathbf{z}_{jt} &= \left[\mathbf{z}_{jt}^{*}, \left(\mathbf{z}_{jt}^{*} \circ \mathbf{z}_{jt}^{*}\right), \left(\mathbf{z}_{jt}^{*} \circ \mathbf{z}_{jt}^{*} \circ \mathbf{z}_{jt}^{*}\right), \mathbf{z}_{jt}^{*} x_{jt}, \mathbf{z}_{jt}^{*} x_{jt}^{2}, \left(\mathbf{z}_{jt}^{*} x_{jt}\right) \circ \left(\mathbf{z}_{jt}^{*} x_{jt}\right)\right] \end{aligned} \tag{21}$$

Increasing numbers of instruments improve the efficiency of the estimator so that coefficient estimates are more likely to be statistically significant for a given data set. We generate the additional instruments as in Eq. 21 rather than drawing from the distribution of Eq. 20 for numerical reasons discussed following Eq. 23.

Negative values of price are drawn under the specification in Eq. 20, but since only differences in utility between products (as opposed to absolute utility) affect the (mixed) logit probabilities when no outside good is included in the model, all prices could be trivially shifted upward by a constant so long as the population taste parameters are independent. We specify the data covariance matrix $\mathbf{\Sigma}_x$ to have the structure:

$$
\boldsymbol{\Sigma}_x = \begin{bmatrix} 1 & \rho_x & \rho_\xi & \rho_i & \rho_i & \rho_i \\ \rho_x & 1 & 0 & 0 & 0 & 0 \\ \rho_\xi & 0 & 1 & \rho_z & \rho_z & \rho_z \\ \rho_i & 0 & \rho_z & 1 & 0 & 0 \\ \rho_i & 0 & \rho_z & 0 & 1 & 0 \\ \rho_i & 0 & \rho_z & 0 & 0 & 1 \end{bmatrix} \tag{22}
$$

where ρ_ξ is the price-ASC correlation that determines the presence and magnitude of price endogeneity, ρ_x is the correlation of price with the technology attribute, ρ_i is the correlation of price with each of the instruments, and ρ_z is the correlation of the ASC with each of the instruments ("IV-ASC correlation"). The covariance matrix $\boldsymbol{\Sigma}_x$ is by definition positive semi-definite and symmetric. We examine cases under the following correlations:

$$
\begin{aligned}
&\rho_\xi = 0.1 \text{ } or \text{ } 0.4 \text{ } or \text{ } 0.7 \\
&\rho_z = 0 \text{ } or \text{ } 0.4 \\
&\rho_x = 0.1,\ \rho_i = 0.4
\end{aligned} \tag{23}
$$

The three levels of ρ_ξ represent the low, base, and high endogeneity cases respectively, and the two levels of ρ_z represent the valid and invalid instrument cases respectively. Specifying ρ_z equal to 0.4 results in invalid instruments because valid instruments must be uncorrelated with the ASC. The correlation levels were determined in large part by the restriction that the covariance matrix must be positive semi-definite and larger off-diagonal elements make this property more difficult to satisfy. Given this restriction, we jointly choose values for the ρ's that allow for the greatest difference in data scenarios. Setting ρ_ξ equal to 0.4 versus 0.5, e.g., does affect the results of the estimation, but the general conclusions of the synthetic data study hold. The positive semi-definite restriction on the covariance matrix is also the motivation for drawing only three instruments and generating others as functions of these basis instruments in Eq. 21; increasing the dimensions of the sparse matrix while adding off-diagonal non-zero elements further limits the maximum allowable magnitudes of the data correlations.

Market shares are simulated for t=1 according to Eq. 15 with the specified values of the taste parameters:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_p \\ \mu_x \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \ \boldsymbol{\sigma} = \begin{bmatrix} \sigma_p \\ \sigma_x \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{24}$$

where μ_p and σ_p are the population mean and standard deviation of the price coefficient and μ_x and σ_x are the population technology attribute coefficients.

To generate data for future markets in years *t=2,3,4,5,6,11* we randomly select 28 (40%) of the products from the prior year to be replaced and generate new product attribute, instrument, and ASC values as described for the base year *t=1*. Years *t=1,2,3,4,5* comprise the estimation data set and years *t=6,11* are the 1-year-forward and 5-year-forward prediction data sets, respectively. Note that the generated ASC for a given product is held constant over all estimation and prediction data, consistent with the idea that it is an aggregate measure of unobserved attributes, but in our estimation procedures a year-specific ASC is estimated so that there are sufficient degrees of freedom to constrain predicted shares of the estimation data to equal observed shares exactly. The price and technology attribute for a given product are also held constant year over year.

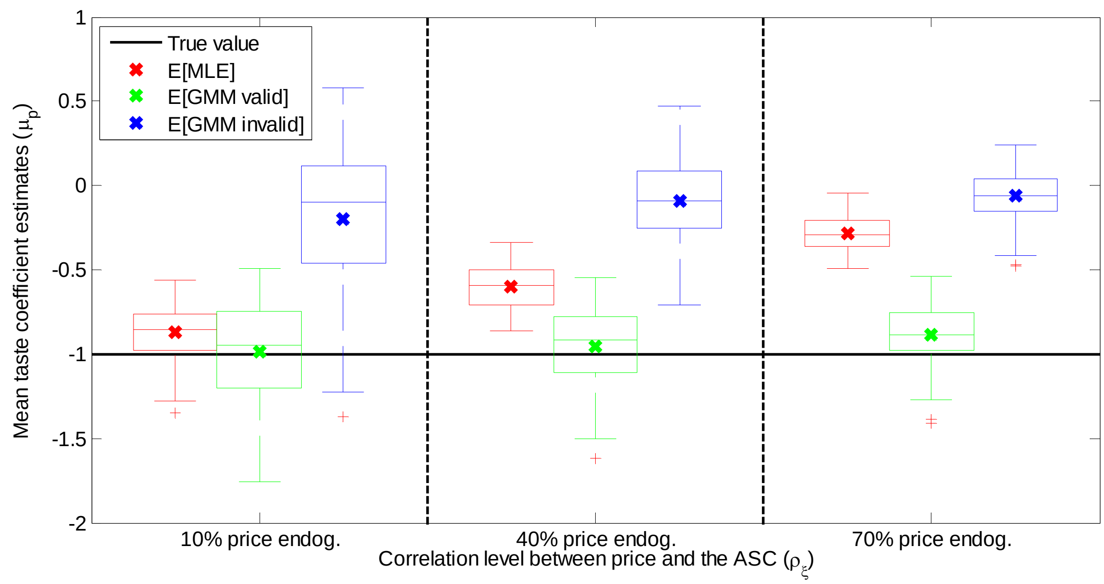
3.4.2 Estimation

We estimate independent mixed logit models for 125 synthetic data sets generated by the process described in section 3.4.1 for each pairing of price-ASC and IV-ASC correlation levels. There are a total of six estimation cases (price correlation=10% (low), 40% (base), 70% (high) and instrument correlation=0% (valid) and 40% (invalid). For a given estimation case, the expected value of the coefficients is approximated by averaging over the 125 estimates. A coefficient's bias is the deviation of its expected value (obtained here by estimating the model repeatedly for different simulated data sets) from the population parameters. The estimation routine successfully converged (exit flag of 0 using the Knitro solver for Matlab) for all 125 data sets for both the MLE-C and GMM-IV estimation methods.

Figure 2 compares the distributions of estimates of the mean taste parameter price μ_p obtained by MLE-C estimation and GMM-IV estimation with valid and invalid instruments for the three levels of price-ASC endogeneity. The distributions represent the coefficients resulting from models fit to the 125 data sets excluding the least and greatest 2.5% of estimates of each coefficient (120 total coefficient estimates are included in each box plot). The expected values of the estimated coefficient values are indicated by x's, and the coefficient values from Eq. 24 used to generate the data ("true coefficients") are represented by horizontal black lines. The distance between an x and a black line is the finite sample coefficient bias. The estimates of the mean taste parameter μ_x and the standard deviation of the taste parameters $\boldsymbol{\sigma}$ are presented in Appendix J.

The bias of the price coefficient increases with endogeneity for the MLE-C estimator. The valid GMM-IV estimator successfully corrects for endogeneity (Figure 2) as expected, but the invalid GMM-IV estimator is more biased than the MLE-C estimator. The MLE-C estimates on average have a standard error on the order of 1e-4 and the invalid GMM-IV estimates on average have standard error on the order of 1. So while the biased MLE-C price parameter estimates are less biased than those of the

GMM-IV estimates, the confidence intervals from MLE-C may be less likely to contain the truth (i.e. have worse coverage).



(A)

Figure 2 — The MLE-C estimator and invalid GMM-IV estimator exhibit bias that increases with price correlation whereas the valid GMM-IV estimator mitigates the endogeneity

We are interested in the accuracy of the ASCs in addition to the accuracy of the observed attribute coefficients because the ASCs may be used in counterfactuals or for forecasting. Figure 3 shows the distribution of error as the difference between the value of the ASCs obtained from MLE-C and valid GMM-IV estimation (A) and invalid GMM-IV estimation (B) and the actual ASCs, which are known from the synthetic data generating process. The box plots in Figure 3 include the ASC error from 125 estimated data sets for the valid instrument case excluding the least and greatest 2.5% of error differences. While both methods produce unbiased estimates of the ASC, the MLE estimator is more efficient, meaning that for any individual data set, the MLE estimator is likely to be closer to the truth than the GMM-IV estimator[15]. The relative efficiency of the MLE estimator as compared to the GMM-IV estimator diminishes as the endogeneity increases, but it is still more efficient even at 70% price endogeneity. This is anticipated as MLE-C utilizes all of the variation in the data, whereas GMM-IV does not [91].

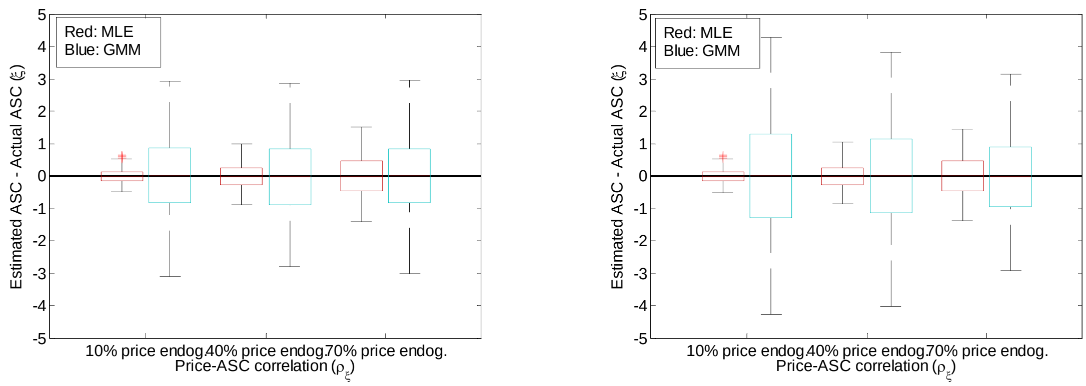


Figure 3 — The GMM-IV estimator exhibits greater ASC error variance than the MLE-C estimator at low levels of correlation but the difference shrinks at greater price correlation levels when both valid (A) and invalid (B) instruments are used in GMM-IV estimation

3.4.3 Prediction

The results of section 4.3.2 indicate that (1) the price coefficient estimated by MLE-C is inconsistent, but GMM-IV with valid instruments corrects for endogeneity and (2) that there is more variation in the observed coefficient and ASC GMM-IV estimates. We examine the implications of these differences on forecast accuracy by comparing the predictions of the estimated models by the two methods. If relationships that induce bias in the estimated coefficients persist in the prediction data, then models with biased coefficients need not predict any worse than models with unbiased coefficient estimates

---

[15] The slight differences in MLE-C ASC error distributions across the valid and invalid instrument cases are due to the difference in the data generation matrix $\mathbf{\Sigma}_x$ when ξ_z=0 versus ξ_z=0.4. When instruments are invalid there is less overall unexplained variation in the generated estimation data.

[93]. However, if there are fundamental shifts in the structure of the data, then models with biased coefficients are expected to predict worse than models with unbiased coefficients.

We estimate the model in Eq. 15 on the estimation data set using MLE-C and GMM-IV. The models are then used to forecast product market shares 1- and 5-years-forward (including new entrants in the prediction data that are not in the estimation data). The $\widehat{\boldsymbol{\beta}}$ and $\widetilde{\boldsymbol{\beta}}$ parameters are assumed constant across markets, however, the $\hat{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\xi}}$ parameters are market (year) specific and thus a method for determining their values in the prediction years is needed.

Since the ASCs are unobserved, we treat them as a random vector and compare the distribution of share forecasts resulting from different realizations of the ASCs. We are interested in two questions:

- Is there a particular method of generating prediction ASCs that is more robust than the others (in the case that the ASCs represent constant unobserved attributes)?
- How do the models and ASC forecasting methods compare across various estimation and prediction data correlation structures?

Predicting ASCs

We propose two non-parametric distributions for generating ASCs in the prediction data. For "incumbent" products, those in the prediction data that also appear in the estimation data, there are historical estimates of the ASC that may provide information on potential future ASCs if the ASCs do in fact represent unobserved attributes as they do in the synthetic data study. For new "entrant" products introduced in or before the prediction years but after the estimation data years, there are no such historical estimated ASCs. As a result, the ASC forecasting methods we propose differ for incumbents and entrants. Table 5 describes these methods.

Table 5 — Description of ASC forecasting methods

Method name	Method qualitative description	Method mathematical description
Products appearing in the estimation set (incumbents)		
Incumbent	For each product draw uniformly from the product's estimated ASCs	Draw ξ_{j6} or ξ_{j11} uniformly from $\{\xi_{=1},\xi_{j2},\xi_{j3},\xi_{j4}, \xi_{j5}\}$
Products not appearing in estimation set (entrants)		
All (Method 1)	Draw from uniformly from all estimated ASCs	Draw ξ_{j6} or ξ_{j11} uniformly from $\{\xi_{kt} \forall k \in J_t, t = 1,2,3,4,5\}$
Nearest neighbor (Method 2)	For each new product calculate the normalized vector distance of product observed attributes between the new product and each of the estimation data set products. Draw uniformly from the estimated ASCs of the observed product with the smallest vector distance ("nearest neighbor")	Draw ξ_{j6} or ξ_{j11} uniformly from $\{\xi_{k^*1},\xi_{k^*2},\xi_{k^*3},\xi_{k^*4},\xi_{k^*5}\}$ where $k^* = argmin_{k\in J_t, t=1,2,3,4,5}(\|\mathbf{x}^*_{jt} - \mathbf{x}^*_{kt}\|)$ and $\mathbf{x}^*_{jt}$ is the (K×1) vector of normalized observed product attributes

Method 2 would not necessarily be expected to predict ASCs well since we have explicitly generated data in which x_{jt} and ξ_{jt} are uncorrelated so the technology attribute x should not contain any information about ξ. We include this method primarily for comparison to the case study with real data in section 3.5.

Forecasting shares

We evolve each of the 125 estimation data sets once so that we have 125 pairs of estimation and prediction data sets. We make forecasts for 125 prediction data sets so that our prediction results are not sensitive to the random data generation process. For each of the 125 prediction data sets, we generate a Monte Carlo distribution of market share arising from the uncertainty of forecasting ASCs. One hundred vectors of ASCs for each prediction data set are drawn nonparametrically according to either method 1 or method 2 as described in Table 5. ASCs are necessarily drawn independently from one another since there is no available information about their potential correlation structure. For each draw of ASC we simulate a draw of share by integrating over taste heterogeneity as in Eq. 15. Thus for a given time frame (1-year- or 5-year-forward), estimation method (MLE-C or GMM-IV), and ASC forecasting method (method 1 or 2), 12,500 total share forecasts are made (125 prediction data sets × 100 ASC forecasts).

These prediction methods are a departure from the literature that treats the ASCs as point estimates. Resende *et al.* [38] compare profit (as a function of market share) predicted by a mixed logit model and find that the optimal design of a new product is dependent on whether the expected value of profit— the mean of the Monte Carlo distribution of profit— or the point estimate of profit predicted by the expected value of the parameters is used in optimization. We argue that the expected value of share better characterizes future outcomes, and evaluate models on this forecast. However, while our prediction methods account for the uncertainty of *forecasting* future ASCs, they do not represent the statistical uncertainty of the observed attribute coefficients or *estimated* ASCs. Additional investigation would be needed to determine appropriate methods for doing so.

Evaluating forecasts

We compare the accuracy of the predictions using the relative average likelihood (RAL). The RAL is a monotonic transformation of the likelihood of the model share predictions L_{P} divided by the likelihood of an ideal model L_{I} that perfectly predicts the new shares:

$$RAL = \frac{\left(L_{\mathrm{P}}\right)^{1/N}}{\left(L_{\mathrm{I}}\right)^{1/N}} \tag{25}$$

where N is the number of choices observed. When comparing two models on the same data set, the model with a larger RAL is more likely to generate the observed data. Using RAL instead of likelihood is important because markets that have more diffuse choice probabilities will necessarily have lower likelihoods of ideal prediction. RAL normalizes for this effect and can be interpreted as the fraction of the total possible explanatory power a model obtains. RAL is a monotonic transformation of the Kullback-Leibler (KL) divergence if observed shares are assumed to accurately reflect the true choice probabilities [39].

Table 6 and Table 7 contain the mean of the RAL of expected shares across the 125 prediction data sets calculated as:

$$E_d\left[RAL\left(E_a\left[\mathbf{P}(\beta,\xi)\right]\right)\right]=\frac{1}{D}\sum_{d=1}^{D}RAL\left(\frac{1}{A}\sum_{a=1}^{A}\mathbf{P}(\beta_d,\xi_{ad})\right) \tag{26}$$

where d indexes data sets, a indexes ASC draws, D=125, and A=100. Also included are the standard deviation of the RAL over the 125 data sets and a measure "# superior" that indicates the number of paired data sets for which the respective MLE-C or GMM-IV model had a greater RAL of expected share.

In addition to predictions from the models using ASC methods 1 and 2, the tables include predictions for which an ASC is included in estimation (for GMM-IV) but no ASC is included in prediction for either incumbents or entrants (method 0), a "static" model that holds shares of observed products constant from the last year of the estimation data and divides the remaining share equally among the new products, and a "no info" model that assumes all entrant and incumbent products have equal shares in the prediction years. Shaded boxes indicate the greatest RAL, or "best" model, when comparing across models for a given estimation data price correlation and time horizon. For parsimony we refer to models estimated using MLE-C and GMM-IV as "MLE-C models" and "GMM-IV models" respectively, though MLE-C and GMM-IV are estimation techniques not models in and of themselves.

Tested cases

We test four prediction cases. Prediction case 1 ("Base", Table 6): the GMM-IV estimated model uses valid instruments and the price-ASC endogeneity present in the estimation data persists in the prediction data. Prediction case 2 ("Market shift", Table 7): the price correlation ρ_ξ is set to 0.1, 0.4, or 0.7 when generating the *estimation* data, but it is set to 0 when generating the *prediction* data so that price is no longer endogenous in the forecast years and the source of asymptotic coefficient bias disappears. Prediction case 3 ("invalid", Appendix K): the GMM-IV estimated model uses invalid instruments and price-ASC endogeneity persists. Prediction case 4 ("entrants", Appendix K): the shares are forecasted for all products, but shares of the incumbent products are included as a single lump sum in the RAL calculation.

In the following discussion, we remark upon mean RAL comparisons between models and between ASC prediction methods for a given model. Many of the paired RAL differences are statistically significant[16], even some as small as only 3%, but it is subjective as to whether or not this represents a practical difference in forecast accuracy between two sets. Our observations highlight the trends in the data rather than focus on the statistical significance of any individual pairing.

[16] The test for significance is a two-sample t-test for equal means.

Table 6 — BASE case results shown include the mean and std. dev. of the RAL of expected share across 125 data sets for MLE-C and GMM-IV models and the number of data sets for which the MLE-C or GMM-IV model had a greater respective RAL (# superior)

Meth.	1-year-forward None 0	All 1	Neigh. 2	5-year-forward None 0	All 1	Neigh. 2
10% price-ASC correlation (ρ_ξ)						
MLE-C	68%	85%	73%	68%	70%	53%
(std. dev.)	*(7%)*	*(8%)*	*(12%)*	*(8%)*	*(9%)*	*(11%)*
# superior	106	114	110	94	103	113
GMM-IV	65%	74%	59%	65%	65%	40%
(std. dev.)	*(8%)*	*(12%)*	*(16%)*	*(9%)*	*(9%)*	*(14%)*
# superior	19	11	15	31	22	12
Static	68%			42%		
No info	39%			39%		
40% price-ASC correlation (ρ_ξ)						
MLE-C	71%	86%	77%	71%	72%	55%
(std. dev.)	*(7%)*	*(8%)*	*(10%)*	*(7%)*	*(7%)*	*(9%)*
# superior	106	116	116	105	106	115
GMM-IV	66%	72%	60%	66%	66%	41%
(std. dev.)	*(8%)*	*(14%)*	*(16%)*	*(7%)*	*(8%)*	*(13%)*
# superior	19	9	9	20	19	10
Static	71%			48%		
No info	44%			45%		
70% price-ASC correlation (ρ_ξ)						
MLE-C	81%	91%	85%	80%	81%	69%
(std. dev.)	*(5%)*	*(5%)*	*(6%)*	*(4%)*	*(4%)*	*(8%)*
# superior	124	124	119	121	124	121
GMM-IV	72%	74%	68%	71%	71%	52%
(std. dev.)	*(7%)*	*(14%)*	*(17%)*	*(8%)*	*(8%)*	*(15%)*
# superior	1	1	6	4	1	4
Static	77%			56%		
No info	52%			54%		

Table 7 — MARKET SHIFT case results shown include the mean and std. dev. of the RAL of expected share across 100 data sets for MLE-C and GMM-IV models and the number of data sets for which the MLE-C or GMM-IV model had a greater respective RAL (# superior)

Meth.	1-year-forward None 0	All 1	Neigh. 2	5-year-forward None 0	All 1	Neigh. 2
10% price-ASC correlation (ρ_ξ)						
MLE-C	69%	84%	74%	68%	69%	51%
(std. dev.)	*(7%)*	*(8%)*	*(10%)*	*(7%)*	*(8%)*	*(10%)*
# superior	102	112	110	87	87	116
GMM-IV	66%	75%	62%	66%	66%	39%
(std. dev.)	*(8%)*	*(12%)*	*(15%)*	*(9%)*	*(9%)*	*(13%)*
# superior	23	13	15	38	38	9
Static	66%			40%		
No info	38%			37%		
40% price-ASC correlation (ρ_ξ)						
MLE-C	68%	81%	72%	66%	67%	50%
(std. dev.)	*(8%)*	*(9%)*	*(11%)*	*(7%)*	*(8%)*	*(11%)*
# superior	88	95	116	68	76	114
GMM-IV	66%	73%	56%	65%	64%	38%
(std. dev.)	*(9%)*	*(13%)*	*(16%)*	*(8%)*	*(9%)*	*(13%)*
# superior	37	30	9	57	49	11
Static	62%			39%		
No info	40%			37%		
70% price-ASC correlation (ρ_ξ)						
MLE-C	67%	74%	68%	62%	62%	52%
(std. dev.)	*(9%)*	*(11%)*	*(12%)*	*(8%)*	*(8%)*	*(10%)*
# superior	85	56	108	40	34	112
GMM-IV	65%	73%	57%	65%	65%	41%
(std. dev.)	*(9%)*	*(13%)*	*(17%)*	*(9%)*	*(9%)*	*(14%)*
# superior	40	69	17	85	91	13
Static	58%			39%		
No info	42%			37%		

Is there a particular method of generating ASCs for forecasting that generates the best predictions?

In the short term, 1-year-forward forecasts, including an ASC by method 1 (drawing entrant ASCs from all product estimated ASCs) improves predictions over excluding the ASC entirely (method 0). In the long term and for entrant products, including an ASC by method 1 does not improve forecasts. Method 2, or nearest neighbor, improves MLE-C predictions in the short term but results in worse predictions for the long term. Method 2 is always worse than excluding ASCs for the GMM-IV model regardless of time horizon, endogeneity level, or market correlation structure as expected based on the structure of the synthetic data.

How do the models compare across various estimation and prediction data correlation structures?

As discussed previously, we use RAL as a metric of comparison across the models.

Base case: MLE-C is better on average than GMM-IV for any of the ASC forecasting methods, and MLE-C is generally better than GMM-IV at predicting any given data set (comparing "# superior").

Market shift case: MLE-C is better on average than GMM-IV for short term and /or low price-ASC correlation forecasts, but as the time horizon increases or the market shift becomes more dramatic, the GMM-IV model is better on average since the biased MLE-C coefficients are detrimental to prediction.

Invalid case: If invalid instruments are used in the GMM-IV model, we find that MLE-C predicts better on average than the best GMM-IV model. As the endogeneity increases, the expected RAL difference between them decreases since the bias of the invalid GMM-IV coefficient aids in prediction for highly endogenous data.

Entrant case: MLE-C is better on average than GMM-IV.

Regardless of whether the price coefficient is biased, we find that the best attribute based models are at least as good as the naïve (static) model and always better than random guessing (no info), except when predicting the short term market with a GMM-IV model estimated using invalid instruments.

The greatest discrepancy in prediction accuracy between the best MLE-C and GMM-IV models across the four prediction cases occurs when the GMM-IV model is estimated using invalid IVs. The MLE-C model is superior regardless of the level of endogeneity, and the penalty is steep at low endogeneity with an RAL difference of approximately 22%. Conversely, when the GMM-IV model does predict better than the MLE-C model in the high-endogeneity, market-shift case, there is only a 2% lift in RAL from the best MLE-C to the best GMM-IV forecast. This illustrates the greatest drawback to using GMM-IV in a forecasting scenario; it risks making far worse predictions for limited and unlikely upside (because valid IVs are difficult to specify [91]).

3.5 Empirical Case Study

We apply the same approach as described in section 3.4 to a real automotive sales data set in order to investigate the accuracy and uncertainty of vehicle demand. A mixed logit model is estimated by MLE-C and GMM-IV on US consumer new midsize sedan purchase data from 2002 through 2006 and then used to predict market shares for midsize sedans sold in the US during 2007 and 2011. The accuracy and uncertainty of the model forecasts are compared on the RAL measure.

3.5.1 Models and Data

We define an independent random coefficient mixed logit model with a linear utility function including the covariates listed in Table 8 plus an ASC. The attributes capture vehicle price, operating cost, performance, size, and country of origin. This choice of covariates is loosely based on Haaf et al. [93] who compare the market share forecasts of 9,000 possible logit model specifications informed by the vehicle demand literature. They find that predictive accuracy is relatively invariant to the specific form of the covariates (e.g. gallons/mile versus miles/gallon), so long as each covariate is included, and prediction accuracy increases with additional covariates. We omit some covariates that were included in the models tested in Haaf et al. [93] due to the specific nature of the GMM-IV estimator. Dummies for A/C standard and automatic transmission are excluded because dummies increase the number of parameters to be estimated but cannot function as instruments, making it more difficult to meet the GMM-IV estimator requirement that there are more instruments than observed variable coefficients. We proxy dummies for brand (e.g. Honda, Ford, or Volkswagen) by dummies for producer firm geographic location (US, Europe, or Asia), reducing the number of dummies from 23 to 2 (a dummy for US is omitted for identification). This reduces the number of parameters to be estimated as well as results in a statistically significant price coefficient for the estimated model. Additional discussion of parameter selection is included in Section 3.6.2.

Our data set uses vehicle attribute information from Ward's Automotive Index [40] and MSRP and aggregate sales data from Polk [41]. We implicitly assume that all individuals who purchased a vehicle in this class considered only and all of the other midsize sedans available in the same year and made a compensatory decision based on vehicle attributes. Our models consider only new midsize sedan buyers, thus there is no outside good (option to not purchase any midsize sedan). GMM-IV estimation of the model on the full vehicle market would require more years of data (markets) than are available to us[17].

The choice of instruments is non-trivial and a subject of much study in the econometrics literature [37]. By definition they must be correlated with the endogenous variable (in this case price) and uncorrelated with the error (ASC) [37]. We specify instruments similar to those of BLP [58]: for a given vehicle, the IVs are the sum of each of the non-price attributes over all other vehicles in the market offered by the same firm as the given vehicle (excluding the given vehicle itself) and the sum of each of the attributes over all other vehicles in the market sold by the competitor firms. Berry *et al.* [58] argue that these IVs are correlated with the vehicle price but uncorrelated with the ASC because the firms observe the ASCs (it is only unobservable to the researcher) of all vehicles and set prices of their vehicles to be competitive while accounting for the utility derived from the ASCs. The ASC of a given vehicle, however, is not expected to be affected by the non-price attributes of competitors. These instruments are predicated on the interpretation of the ASC as a representation of aggregate unobserved vehicle attributes.

In our formulation of the BLP instruments we exclude attributes that enter as dummies since instruments must exhibit sufficient variation. We also add functions of

[17] We estimated several mixed logit model specifications on all 2004-2006 new vehicle sales (full market), but were unable to obtain a model with any statistically significant coefficients. 2002-2003 full market data was unavailable for the full market.

these instruments to satisfy the requirement that $L \geq 2K$ and to increase the efficiency of our GMM-IV estimator. The instruments are defined as:

$$\begin{aligned}
&\mathbf{z}_{jt}^{(1)} = \mathbf{x}_{jt}^{*},\ \mathbf{z}_{jt}^{(2)} = \sum_{k \backslash j \in F_{Ot}} \mathbf{x}_{kt}^{*},\ \mathbf{z}_{jt}^{(3)} = \sum_{k \backslash j \in F_{Ct}} \mathbf{x}_{kt}^{*} \\
&\mathbf{z}_{jt}^{(4)} = \sum_{k \backslash j \in F_{Ot}} \left(\mathbf{x}_{kt}^{*} \circ \mathbf{x}_{kt}^{*}\right),\ \mathbf{z}_{jt}^{(5)} = \sum_{k \backslash j \in F_{Ct}} \left(\mathbf{x}_{kt}^{*} \circ \mathbf{x}_{kt}^{*}\right) \\
&\mathbf{z}_{jt}^{\sim} = \left[\mathbf{z}_{jt}^{(1)}, \mathbf{z}_{jt}^{(2)}, \mathbf{z}_{jt}^{(3)}, \left(\mathbf{z}_{jt}^{(1)} \circ \mathbf{z}_{jt}^{(1)}\right), \mathbf{z}_{jt}^{(4)}, \mathbf{z}_{jt}^{(5)}, \left(\mathbf{z}_{jt}^{(1)} \circ \mathbf{z}_{jt}^{(2)}\right), \left(\mathbf{z}_{jt}^{(1)} \circ \mathbf{z}_{jt}^{(3)}\right)\right] \\
&\mathbf{z}_{jt} = \mathbf{z}_{jt}^{\sim} \circ \frac{1}{\operatorname{col}\max\left(\mathbf{z}_{jt}^{\sim}\right)},\ \forall j, t
\end{aligned} \tag{27}$$

where $F_{\mathrm{O}t}$ is the set of vehicles made by the same firm in year t, $F_{\mathrm{C}t}$ is the set of vehicles made by competitor firms in year t, and $\mathbf{x}_{jt}^{*}$ is the vector of attributes excluding price and dummies. For numerical purposes we normalize each instrument by dividing by the maximum value of that instrument occurring over all vehicles in all panel years in order to prevent an ill-conditioned weighting matrix $\mathbf{W}=(\mathbf{Z}'\mathbf{Z})$.

There are a total of six covariates: price, three exogenous non-dummy attributes, and two country dummies, yielding ten model parameters to be estimated (non-dummy covariates have mean and variance coefficient components) and 24 total instruments.

3.5.2 Coefficient Estimates

The model coefficients estimated by MLE-C and GMM-IV are shown in Table 8. The mean coefficient estimates are of the same sign (excepting the Asia location dummy) and same order of magnitude, though the only statistically significant difference between the two model mean coefficient estimates is on price[18]. The MLE-C estimated model

[18] The asymptotic distribution of the difference of each mean coefficient is $N(\mu_{MLE} - \mu_{MLE}, s_{MLE}^2 + s_{GMM}^2)$ where s^2 is the square of the standard error of the mean parameter estimate. For each mean coefficient we

coefficients indicate little taste heterogeneity in the population, but the GMM-IV estimated coefficients indicate larger taste heterogeneity. This suggests fundamentally different views of taste preference distribution for this data set. However, only the gallons/mile heterogeneity parameter is statistically significant between the two estimated models.

Table 8 — Coefficients for MLE-C and GMM-IV models estimated on 2002-2006 US midsize sedan new sales data

	Estimated population mean taste (**μ**)		Estimated population taste heterogeneity (**σ**)	
	GMM-IV	MLE-C	GMM-IV	MLE-C
Price ($10,000)	-1.1***	-0.5***	0.0033	0.0039*
(standard error)	(0.3)	(0.0)	(2.1270)	(0.0022)
Gallons/mile (gal./100-mi.)	-0.3	-1.1***	0.9536**	0.0033*
(standard error)	(0.6)	(0.0)	(0.4414)	(0.0018)
Weight/horsepower (10 lbs/hp)	-0.2**	-0.1***	0.0000	0.0003*
(standard error)	(0.1)	(0.0)	(0.2762)	(0.0002)
Length×width (100 ft^2)	2.6	3.8***	1.7229	0.0085*
(standard error)	(1.8)	(0.0)	(1.8169)	(0.0047)
Europe (dummy)	-1.1**	-1.1***		
(standard error)	(0.5)	(0.0)		
Asia (dummy)	-0.1	0.3***		
(standard error)	(0.3)	(0.0)		

Note: US country dummy omitted for identification; zero estimates are zero to the precision shown but are not actually zero
* = significant at the α=0.10 level
**= significant at the α=0.05 level
***= significant at the α=0.01 level

It is somewhat surprising that the MLE-C model indicates little taste heterogeneity in the population, and we take several steps in order to test the validity of our estimation results. First, we use 100 Halton draws of **ν** that are held constant throughout estimation in order to prevent simulation noise. Second, we perform multistart optimization with 100 mean taste parameter starting values drawn randomly from the interval [-10,10] and heterogeneity parameters drawn randomly from the interval [0,10].

---

test the null hypothesis H_0: $\mu_{MLE}=\mu_{GMM}$ against H_A: $\mu_{MLE}\neq\mu_{GMM}$. We reject H_0 at the α=0.1 for price only; for all other mean coefficients we do not reject H_0.

For 87/100 of the starting points the solver converged successfully and to the same objective function value and optimizer. Third, we check that the objective function is not flat in the neighborhood of the optimizer— the log likelihood is -4.52×10^7 at the optimizer versus -4.54×10^7 at more anticipated values of $\boldsymbol{\sigma} = \mathbf{0.1}$. Fourth, the analytical hessian is provided to the solver, and the eigenvalues of the hessian calculated at the optimizer indicate that the convergence is not a specious result of derivatives that vanish near coefficient values of zero. Lastly, when the model is estimated on synthetic data with no price-ASC correlation, the routine successfully recovers the true heterogeneity coefficient values. Two studies from the literature also encounter difficulties with estimating heterogeneity parameters. Train and Winston [47] are unable to obtain statistically significant heterogeneity coefficient estimates when only the vehicle chosen by consumers (and not the specific choice set) is known. Berry *et al.* [49] report algorithm convergence issues when only one market (year) of data is used and suggest that sufficient variation of choice set across markets may enable successful estimation. The limited variation in choice set for our data set may be the cause of the unexpectedly small heterogeneity coefficient estimates.

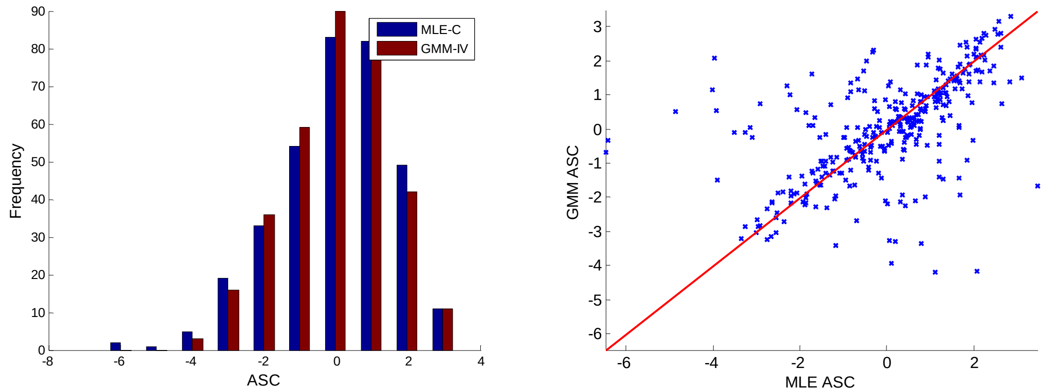
The standard errors for $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are much larger for the GMM-IV estimated coefficients than for the MLE-C estimated coefficients. The GMM-IV is a less efficient estimator than MLE-C so we would expect the standard errors to be somewhat larger. That they are orders of magnitude larger is likely a result of the different estimation strategies for the ASC. There are an additional 339 parameters in the specification of the model estimated by GMM-IV over the model estimated by MLE since MLE-C ASC calibration is performed post the estimation of the observed attribute coefficients. If the

ASCs included in GMM-IV estimation are correlated with price (as we assume they are), then the collinearity may increase the standard errors [97].

If we assume that the GMM-IV estimator successfully corrects for price-ASC endogeneity (which may not be true if the instruments are invalid), then the direction of the bias of the MLE-C price coefficient as compared to the GMM-IV coefficient estimate is reasonable since we would expect the ASC to encompass unobserved features that consumers are willing to pay a premium for, such as options packages. We treat the non-price attributes as exogenous in our model since our focus is on price-ASC endogeneity (and this is common in the literature), but this is not likely to be true. If the non-price attributes are correlated with the ASC, then these coefficients will also be biased (and our instruments will be invalid). In automotive demand models this is virtually certain. Fuel economy, acceleration, weight, dimensions, and manufacturer geographic location are all likely to be strongly correlated with aesthetics and other unobserved or unquantifiable vehicle features, thus there should be nontrivial correlations between observed features and the ASC.

3.5.3 Analyzing Estimated ASCs

Figure 4a contains histograms of ASCs estimated by the two models. The ASCs have a non-normal distribution for both of the models according to a Jarque-Bera test conducted at the α=0.05 significance level. Figure 4b plots the MLE-C versus GMM-IV estimated ASCs. Points falling on the red diagonal line are estimated ASCs that are identical between the two models. There is dispersion of the points from the line, indicating that the models do not agree on the value of the ASCs, but there is no indication that one model consistently under or over estimates them as compared to the other.



(A) **(B)**

Figure 4 — The distribution of estimated ASCs for GMM-IV and MLE-C is non-normal and left skewed (A), but the MLE-C versus GMM-IV estimated ASCs (B) do not indicate that there is a systemic difference in their discrepancy

In order to investigate the (possible) correlation between the ASCs and vehicle attributes we regress MLE-C and GMM-IV estimated ASCs on six sets of dependent variables. The estimated ASC is regressed on: (1) an intercept plus vehicle physical attributes (price, gallons/mile, weight/horsepower, and (length×width)), (2) brand dummies (e.g. Acura, Ford, etc.), and (3) a dummy variable for unique vehicles at the aggregate make-model level (a Toyota Camry and Toyota Camry Solara are both assigned a single ID representing a Toyota Camry). Though we do not include brand dummies in the estimated models of Table 8, we include them in the ASC regressions as covariates since ASCs and brand are likely related. Additional results for an MLE-C estimated model that includes brand dummies are presented in Appendix L.

All three of these regressions yield at least 24% statistically significant coefficients for both MLE-C and GMM-IV estimated ASCs. This supports the use of observed vehicle characteristics in forecasting ASCs, particularly the methods used in this case study. Additionally, that the GMM-IV estimated ASCs are statistically significantly correlated with non-price vehicle characteristics suggests that the BLP instruments were, for our data, invalid.

Additional regressions, results and discussion are included in Appendix M.

3.5.4 Prediction

The coefficients estimated by MLE-C and GMM-IV are used to forecast the 2007 and 2011 sales. In 2007, 33% of the midsize sedans were new (“entrants”, 23 out of 68 did not appear in estimation data), and in 2011, 72% of the luxury sedans were new (34 out of 47 did not appear in estimation data)[19]. We test four methods of predicting new product ASCs. Methods 1 and 2 are the same as described in Table 5 where method 1 draws entrant ASCs from all estimated ASCs and method 2 draws entrant ASCs from estimated ASCs of the “nearest neighbor” vehicle. Additionally we introduce a third method in which entrant vehicle ASCs are drawn from the estimated ASCs of the same brand (“brand”), and a fourth method in which the ASCs are drawn from the estimated ASCs of other trim levels of the same vehicle make-model (“make-model’). If the entrant brand or vehicle make-model is not observed in the estimation data set (e.g. Dodge Caliber), then the entrant vehicle ASCs are drawn from all estimated ASCs as in method 1. Method 3 is similar to that used in Berry *et al.* [49] and Train and Winston [47]. Both method 3 and method 4 are related to the interpretation of the ASC as a vehicle-specific fixed effect that represents aggregate unobserved attribute utility as opposed to a random error term, supported by the regressions of the estimated ASCs on vehicle attributes.

[19] The high turnover is a result of our designation as to what constitutes a “new vehicle”. We consider make-models distinct at the trim level for different body styles (e.g. sedan versus wagon). As a result, several trims appear as “new” in the prediction data when only a subset of body styles is available in each year.

Method 2 assumes that the unobserved attributes represented by the ASC are related to the vehicle's attributes and those of its competitor(s), which is an intuitive means for informing new product ASCs. However, this implies that the ASC is correlated with competitor non-price attributes, violating the assumption required for the BLP specification of instruments that for a given vehicle the sum of competitor attributes serving as instruments are uncorrelated with the ASC. Despite this issue we include results from prediction under this method because it is similar to an approach used in the literature and is tempting to researchers [49].

For each combination of the MLE-C and GMM-IV estimated models and four ASC generation methods, we draw 10,000 sets of ASCs for the predictive year vehicles. For each draw of ASCs we simulate a draw of shares by integrating over taste heterogeneity as in Eq. 15, generating a Monte Carlo simulation of predicted market shares. Forecast uncertainty is represented by the range of the 10,000 sets of share predictions, and our point estimate of share ("expected share") is obtained by averaging over the draws of share. We report the RAL of the expected share predictions in Table 9. The following comparisons of ASC and model estimation methods are based on this measure where greater values of RAL indicate more accurate predictions.

Two additional models are included in Table 9 for comparison purposes. The static model ("static") holds shares of incumbent vehicles constant from the last year of observation in the estimation data and divides the remaining share equally among the entrant vehicles. The no info model ("no info") assumes all vehicles have equal share in the prediction year.

Table 9 — RAL(E[share]) comparison of ASC forecasting methods for MLE-C and GMM-IV models estimated on 2002-2006 midsize sedan data and used to predict 2007 and 2011 midsize sedan market shares

	1-year-forward (2007) forecasts					5-year-forward (2011) forecasts				
	No ASC	All	Near neighbor	Brand	Model	No ASC	All	Near neighbor	Brand	Model
Method:	0	1	2	3	4	0	1	2	3	4
					RAL of expected share					
MLE-C	45%	66%	61%	63%	66%	38%	42%	60%	37%	41%
GMM-IV	34%	68%	64%	70%	70%	28%	33%	45%	34%	33%
Static	*68%*					*39%*				
No info	*32%*					*42%*				
					RAL of expected share— ENTRANTS ONLY					
MLE-C	91%	91%	84%	88%	91%	49%	50%	71%	43%	49%
GMM-IV	86%	85%	81%	87%	88%	43%	43%	57%	43%	42%
Static	*87%*					*54%*				
No info	*87%*					*62%*				

Note: highlighted cells indicate the most accurate model and ASC forecasting method for a given time period and method of calculating RAL

Comparing the ASC generation methods in Table 9, method 4 forecasts best for both the MLE-C and GMM-IV methods for the 1-year-forward ("short term") and method 2 forecasts best for the 5-year-forward ("long term") time horizons. There are two important comparisons in the 5-year-forward forecasts. First, the no info model predicts better than the static model. Second, the best MLE-C model predicts much better than the no info model, but the best GMM-IV model has similar accuracy. The first comparison suggests than the market has changed sufficiently in composition such that competitor products have significantly altered the share of the incumbent products. The second comparison suggests that the underlying relationships in the data, like the correlation of the ASC and price, persist and to some degree are successfully captured by the MLE-C coefficients.

We estimate a model that includes brand dummies by MLE-C and present the results in Appendix L. The MLE-C model predictions with no forecast ASCs are greatly improved by the inclusion of brand dummies as expected (2011 RAL of 51% for no-ASC brand dummy model versus 38% for no-ASC model that does not include brand dummies). The model that includes brand dummies is superior to the model that excludes them and predicts ASCs by the brand method (method 3) but inferior to the model that predicts ASCs by the nearest neighbor method. These comparisons suggest that the brand dummies capture much— but not all— of the explanatory power of the ASCs and that the brand method of forecasting ASCs is unable to recover it. The advantage of explicitly including brand dummies over aggregating the utility contribution with other unobservable characteristics is that brand is observed in future markets so that the predicted brand utility contribution is not susceptible to changes in correlation between brand and the aggregate ASC.

We suspect that the ASC is endogenous with the other vehicle physical attributes in addition to price. For the long term forecasts, method 2 is better than any of the other ASC prediction methods and better than omitting the ASC in prediction for both models. In the synthetic study, when, by design, the ASC was uncorrelated with the non-price attribute *x*, predicting the ASC by method 2 resulted in worse predictions than when no ASC was predicted by method 0. The regression of ASCs on vehicle physical characteristics also provide support for non-price attribute endogeneity with the ASC. This is in direct violation of the assumption for specification of valid instruments that non-dummy attributes and ASCs are uncorrelated, again implying that the instruments in this case study— used frequently in vehicle demand literature [37,39,45,47,52–54]— are invalid.

Table 9 also contains the RAL for entrant products only. The shares are forecasted for all midsize sedans, but the shares of the incumbent products are included as a single lump sum in the RAL calculation. As for the whole market, method 4 is better in the short term, but method 2 is better in the long term. In contrast to the whole market, the best GMM-IV attribute model does not predict better than the no info model in the long term.

3.6 Discussion

3.6.1 Lessons from the Simulation Study

Much of the econometrics literature on vehicle market modeling has presumed that biased coefficients make for bad models. A main method proposed there, GMM-IV, does indeed eliminate coefficient bias in our synthetic data study when valid instruments are used. However, our synthetic data study also shows that correcting coefficient bias does not necessarily produce better forecasts. So long as the underlying source of the bias persists in the forward years, MLE with biased coefficients could result in forecasts at least as accurate as those made by GMM-IV's unbiased coefficients.

We find support for using ASCs in choice modeling when consistent coefficient estimates are required— e.g. when evaluating willingness-to-pay. However, the models that ignore ASCs entirely in prediction and estimation (MLE-C with no forecasted ASC) predict as well or better than the models that estimate ASCs simultaneously with other model parameters (GMM-IV). The MLE-C and GMM-IV models with no ASCs are more accurate than the static and no info model (guessing) for the 5-year-forward-forecasts, meaning attribute-based models yielded more accurate forecasts even if some portion of utility is unaccounted for.

Invalidity of instruments is a significant issue for GMM-IV methods in practice because instrument validity is impossible to verify, and we find that using invalid instruments can potentially result in models with estimated coefficients exhibiting more bias than MLE-C that also make worse forecasts. For long term forecasts, the greatest penalty for using invalid GMM-IV predictions as compared to MLE-C predictions was an RAL difference of 22% (10% price-ASC correlation case). The upside in the only prediction case for which the valid GMM-IV offered an advantage is a mere 2% (market shift case with 70% price-ASC correlation); the greatest penalty for estimating a model with invalid GMM-IV instruments was ~10x greater than the greatest reward for specifying valid GMM-IV instruments in the event a limited market scenario occurs.

3.6.2 Lessons from the Empirical Case Study

We first discuss the empirical case study as if the instruments are valid and then address the finding that our instruments are likely invalid.

Table 9 shows that MLE-C without ASCs (i.e. just MLE) predicts better than GMM-IV without ASCs for both 1-year and 5-year forecasts, at least on the RAL metric. If the instruments are valid, we should expect a GMM-IV model without ASCs to have unbiased coefficients and thus better reflect choice tradeoffs among the observed attributes. That this model predicts worse than an MLE model, which we must assume has more biased coefficients than GMM-IV if the instruments are valid, suggests that there are omitted variables whose utility is correlated with prices. That the MLE-C model predicts better than the GMM-IV model for any ASC forecasting methods suggests that a biased coefficient is more useful for prediction than trying to capture unobserved utility with an ASC alone when this endogeneity persists.

If that is true, do the ASCs improve the predictive performance of the GMM-IV model? The GMM-IV model forecasts are improved by adding ASCs forecasted by any method for both 1-year and 5-year forecasts. Thus adding ASCs to GMM-IV does result in a model that forecasts with unbiased coefficients *and* the influence of omitted attributes. One consequence of this pertains to counterfactual experiments in the explanatory literature. Many papers do not forecast per se, but alter observed market conditions and simulate choice outcomes, market competition, and measure related changes in economic measures such as market power, consumer welfare, fleet fuel economy, etc [47,49,62]. The superiority of near-term forecasts in our study suggest that GMM-IV methods could perform well for counterfactual analysis, if the imposed changes in market structure do not change valuation of omitted attributes.

When only forecast accuracy of entrants are evaluated, adding ASCs to the MLE-C and GMM-IV models does not meaningfully improve predictions in the short term and predicts worse than random guessing for the 5-year-out time horizon for all but the "nearest neighbor" method. These observations lead us to propose that specifying a correct model for the relationship between the observed attributes and ASCs is difficult. Entrants may have an unobserved contribution to utility that is not captured well by historical data. In fact, it may be in automakers' interests to introduce new vehicles with purposefully differentiated (or entirely new) unobserved attributes in order to engage in competitive new vehicle markets. So while the "nearest neighbor" method improves predictions for the data set and long term time horizon tested, this may not be true in future years, and falsely assuming this relationship persists results in worse predictions than excluding the ASC entirely (as seen in the synthetic data study).

The preceding comments are premised on the assumption of instrument validity in GMM-IV. Validity of instruments can only be argued not proven [91], and our interpretation of ASCs as representations of unobserved attributes (consistent with both the explanatory and predictive bodies of literature) as well as the regressions of estimated

ASCs on product characteristics in section 3.5.3 suggests correlations with observations that conflict with IV assumptions. For our data set and choice of instruments, which are popular in the automotive demand literature, the presumably biased MLE-C coefficients were more useful for prediction than the GMM-IV coefficients. If the instruments are valid, then this suggests that bias can *aid* predictions by implicitly capturing persistent unobserved effects. If the instruments are invalid, then predictions are made worse by trying to mitigate the coefficient bias as opposed to ignoring it as under MLE-C. Other choices of instruments may yield different results— predictions from a model estimated using truly valid or better instruments may be superior to a model that ignores endogeneity. There is an inherent risk trade-off: attempting to specify valid instruments risks degrading forecasts if the instruments are invalid.

3.6.3 Computational Issues

The literature on GMM-IV methods has seen a recent focus on the challenges of estimating these types of models. Consistent with the studies discussed in section 3.2 [33,35,94], we encountered several difficulties that we discuss here in order to illuminate technical drawbacks of using GMM-IV. General estimator issues are documented in the marketing and econometrics literature, but we focus on the types of data sets vehicle demand researchers are likely to encounter.

Firstly, we found that the computation time required to estimate a model with GMM-IV was ~2.5x greater than with MLE-C. The GMM-IV optimization problem has nonlinear and computationally expensive share constraints that must be solved at each step, as opposed to only once at the end of the optimization routine as in MLE-C.

Secondly, the GMM-IV estimator is statistically (as opposed to computationally) inefficient relative to the MLE estimator, and it is difficult to specify a sufficient number of instruments for real aggregate data sets of the size and form seen here such that at least some of the coefficients (particularly price) are statistically significant. At least as many instruments as observed coefficients must be used for identification, and we found that at least twice as many instruments as coefficients were needed to ensure a significant estimate of the price coefficient when estimated on our data set. Though any function of the exogenous variables and base instruments can be used, instruments that are too collinear cause numerical difficulties.

GMM-IV estimation was sensitive to the level of aggregation of the vehicle data. We were unable to obtain coefficient estimates when make-model sales were summed over trims.

Table 10 summarizes the model characteristics from the synthetic and empirical case studies. These conclusions are most likely to apply to other products when (1) they are competing in mature markets (relatively constant year over year market shares and limited anticipated structural change), (2) there are a large number of products per market relative to the number of markets available for estimation, and (3) consumer taste preferences are stable.

Table 10 — Comparison of MLE-C and GMM-IV properties and findings

	MLE-C	GMM-IV
Willingness-to-pay estimates	Biased	Can reduce bias given valid IVs
Prediction	More accurate in almost all cases tested	Can be better for markets with high correlation and anticipated dramatic change given valid IVs
Computation	Concave NLP when utility is linear leading to global solution and fast estimation	Nonlinear equality constraints lead to local minima, infeasible regions lacking gradient information toward the feasible domain, and potentially multiple local minima; slow and requires multi start
ASC	Calibration constant does not necessarily represent missing attributes	Estimation parameter represents missing attributes as long as IVs are valid

3.7 Limitations of the Study

Our investigation examines only a portion of the factors affecting vehicle demand prediction uncertainty, and our case study models have error resulting from misspecification and missing information (as do all models). We lack individual-level choice data with consumer covariates, such as demographics or usage variables [42], and are unable to quantify some key purchase drivers, such as aesthetics. We assumed that coefficient attributes are constant over time, i.e. that there are no changes in consumer valuation of attributes. The ASC is intended to capture commonly-held evaluations of unobserved attributes (in the GMM-IV estimated model) and/or other sources of error (in the MLE-C estimated model). However if we suspect that they are correlated with observed attributes other than price (and we do), our coefficient estimates for the observed attributes will be inconsistent.

We restrict our study of demand forecasting to random utility DCMs that treat consumers as observant rational utility maximizers with consistent preferences that fully consider every option in the relevant market. Several studies offer critiques or alternative treatments such as preferences that evolve over time [36], incorporate cultural factors

[43,44], or are adapted to a specific choice situation [45,46]. Especially relevant to the data sources of this study, the Lucas critique warns against use of aggregated historical data to predict outcomes in counterfactual scenarios [47].

We did not consider alternative estimation methods beyond MLE-C and GMM-IV. For example, we do not consider Bayesian methods which estimate coefficients of the same model forms and are asymptotically equivalent to MLE [23], and thus if our MLE estimates are reasonably good we should not expect to see significantly different results with Bayesian methods. Nor did we investigate alternative heterogeneity specifications, e.g. latent class models, mixed logit model with joint parameter distributions, mixture models, and generalized logit models that account for scale and coefficient heterogeneity [18].

We assume in generating prediction ASCs that they are uncorrelated with one another, but this is a restrictive assumption. For GMM-IV estimation we specify a weighting matrix, **W** (Eq. 19) that is less efficient than other candidates. An alternative choice of **W** (e.g. the inverse of the covariance matrix of the moments $\mathbf{W}=(\mathbf{Z}'\xi\xi'\mathbf{Z})^{-1}$) may lead to more accurate estimates for our finite data sample at increased computational cost, though Nevo [67] suggests that this is not a primary concern.

3.8 Conclusion

Including ASCs in product utility functions of DCMs improves model fit but not necessarily forecast accuracy. Our simulation study compares the estimation of a mixed logit model using MLE-C and GMM-IV methods to investigate the effects of mitigating the endogeneity of the price coefficient with the ASC on both the estimated model parameters as well as the accuracy of 1-year- and 5-year-forward forecasts. Several methods for forecasting ASCs are tested to determine the sensitivity of share predictions.

We propose these methods as this issue has not been addressed by the literature. We also examine an empirical case study, estimating mixed logit models on midsize sedan sales from 2002 through 2006 and predicting vehicle market shares in 2007 and 2011.

Given the results of the synthetic data study and the case study, our recommendation is that researchers and practitioners primarily interested in vehicle demand forecasting exclude ASCs entirely and estimate the model using MLE. For purely predictive purposes, the drawbacks of GMM-IV— challenging model estimation and difficult instrument specification— outweigh the expected benefits of potentially mitigating price endogeneity.

4. A Comparison of Vehicle Market Share Forecasts from Bayesian and Frequentist Mixed Logit Models

This third study covers two of the major research themes: comparing model specifications and forecasting methods in terms of prediction accuracy and comparing design implications of different model specifications and forecasting methods. We ask specific research questions: How should estimates of past alternative-specific constants (ASCs) be used in future share forecasts? How do the share forecasts compare when ASCs are treated as model parameters with full distributions of uncertainty rather than as calibration constants? Of the assumptions and models in Table 1 this study assumes the ASC is independent of price and compares the predictions made by mixed logit models with both independent and correlated coefficients[20].

We compare the implications on forecast accuracy and uncertainty when models are estimated on revealed preference, aggregate data sets by frequentist maximum likelihood estimation (MLE) methods to Bayesian-estimated models that permit greater structural flexibility. Models that include an ASC in the utility function are estimated on 2002-2006 U.S. new midsize sedan sales and are used to predict market shares for the 2007 and 2011 midsize sedan markets.

We find that the increased structural flexibility of the mixed logit model with a full covariance matrix estimated by Bayesian methods does not result in meaningfully more accurate predictions. Treatment of the ASCs as model parameters rather than calibration constants improves short term predictions, but the best long term predictions are made by the MLE model. There is comparable uncertainty in forecasts from models

[20] This chapter is included in a working paper. In this chapter the use of first person plural includes coauthors Elea Feit and Jeremy Michalek.

estimated by either method when ASCs are included in predictions, despite differences in their estimation treatment.

4.1 Introduction

Traditionally, discrete choice models (DCMs) have been estimated by frequentist methods, especially MLE and generalized method of moments (GMM). When an ASC is included, instrumental variables (IVs) are often introduced into GMM estimation in order to mitigate price-ASC endogeneity [33,35–37,49,52–54,58]. As Bayesian estimation has become more popular in marketing and statistics, it has also been adopted by automotive demand researchers [17,61,99–105]. Musalem *et al.* [106] outline four main advantages of Bayesian over classical techniques. Firstly, complex models with many parameters, e.g. mixed logit models with full coefficient covariance matrices and parametrically distributed ASCs, are conceptually simple to estimate. Secondly, obtaining distributions of functions of the estimated parameters, e.g. market shares, is straightforward. Thirdly, Bayes estimators allow researchers to obtain finite sample inferences about parameters directly rather than relying on asymptotic results. And lastly, auxiliary information, like expert knowledge, is easy to incorporate through the use of informative priors.

We are interested in the accuracy and uncertainty of DCM automotive demand forecasts. We compare how the increased model flexibility resulting from the use of a Bayes estimator compares to more structurally restricted models estimated by frequentist MLE methods (MLE). Specifically we address:

(Q4.1) How do the accuracy and uncertainty of share predictions from a full covariance mixed logit model compare to an independent mixed logit model?

(Q4.2) How do the share forecasts compare under the Bayesian versus frequentist treatment of ASCs?

The study proceeds as follows: section 4.2 contains a review of Bayesian-estimated models in the automotive demand literature; section 4.3 describes the MLE and Bayesian estimation and prediction methods used in this study; section 4.4 presents the results of a case study that estimates mixed logit models using Bayesian and MLE techniques on 2002-2006 U.S. aggregate midsize sedan sales and predicts 2007 and 2011 market shares; and section 4.6 concludes.

4.2 Literature Review

The application of Bayesian estimation methods to DCMs in the transportation domain is severely limited as compared to other fields [100], and Yonetani *et al.* [107] is the only study that forecasts future vehicle market shares. They predict 1-year-forward vehicle market shares and find that the model predicts well for 36/50 vehicle make-models but ascribe the over/under prediction to predicted ASC inaccuracy and lack of data generally. To the authors' knowledge this study is the only one that compares the accuracy and uncertainty of forecasts when using Bayesian estimation techniques on aggregate data sets to classically estimated model forecasts.

A number of automotive demand studies have applied Bayesian estimation methods to DCMs (Table 11), though they are not concerned with forecasting. A mix of structural specifications appear in the Bayesian automotive demand literature with no dominant favorite: multinomial logit ("logit") [99,101], independent mixed logit [17,107], correlated mixed logit[21] [61,99,105], independent probit [102], and correlated probit [100]. Some studies include an ASC in the utility function [100,101,103,105,107] and others do not [17,61,99,102].

[21] "Correlated mixed logit" is used throughout to refer to a mixed logit model with correlated utility part-worths such that the coefficient covariance matrix has non-zero off diagonal elements

Bayesian-estimated automotive demand DCMs are primarily estimated on disaggregate, stated preference data sets [17,61,99–105]. As exceptions Yonetani *et al.* [107] estimates a model on aggregate and revealed preference data, Daziano and Chiew [102] estimate a model on stated preference data using priors informed by revealed preference data, and Feit *et al.* [61] propose a method to combine disaggregate stated and revealed preference data. This study specifies a correlated mixed logit model and includes a parametrically distributed ASC estimated on aggregate, revealed preference data. (See Table 11 for a comparison of utility and structural specifications in the literature.)

Yang *et al.* [108] propose the first estimation technique in a Bayesian framework for a mixed logit model that includes an ASC, which is the structural specification introduced by Berry *et al.* [58] ("BLP"). Yang *et al.* [108] illustrate the method on a disaggregate data set of household beer purchases. They warn that that classical estimators rely on asymptotic results, yet the data typically used in the model is a small sample for which the properties of the estimators are not well known. Though the total number of observations in a data set may be large— the number of annual new car purchases in the US over twenty years in the case of the Berry *et al.* [58]— it may be considered a small sample if there are many units (households) but only a few observations per unit [108].

Several studies have developed methods for Bayesian DCM estimation on aggregate data. Musalem *et al.* [106] and Jiang *et al.* [109] propose techniques for estimation of a demand model decoupled from a supply-side mode. Musalem *et al.* [106] augment the data with unobserved individual choices, treating the latent individual choices as a distribution with parameters to be estimated. Romeo [110] employs data augmentation but estimates a joint model of supply and demand. Jiang *et al.* [109] propose an estimation method that avoids augmentation with contraction mapping as in BLP.

Table 11 — Bayesian-estimated DCM literature comparison

Author	Year	Model structure[1]	Market structure	A priori ASC distribution	Product	Data source	Case study focus
Jiang *et al.* [109]	2009	Correlated MIXL	Demand-side	IID normal	Canned tuna	Aggregate	Parameter and elasticity distributions
Musalem *et al.* [106]	2009	Correlated MIXL	Demand-side	Corr. multivariate normal	Facial tissue	Aggregate	Parameter and elasticity distributions; out-of-sample elasticity prediction
Romeo [110]	2007	Correlated MIXL	Demand-side, joint supply-demand	IID inverted gamma[2]	Bath tissue	Aggregate	Parameter and elasticity distributions
Daziano [99]	2013	MNL, corr. MIXL, semi-para. MIXL	Demand-side	No ASC	Vehicles	Disaggregate	Electric range willingness-to-pay and elasticity distributions
Daziano and Achtnicht [100]	2013	Correlated mixed probit	Demand-side	Non parametric	Vehicles	Disaggregate	Parameter and market share distributions; counterfactual market shares
Daziano and Bolduc [101]	2013	MNL	Demand-side [3]	Point estimate	Vehicles	Disaggregate	Counterfactual market shares
Daziano and Chiew [102]	2013	Independent mixed probit	Demand-side	No ASC	Vehicles	Disaggregate	Electric range willingness-to-pay and elasticity distributions
Zhang *et al.* [17]	2011	Independent MIXL	Demand-side [3]	No ASC	Vehicles	Disaggregate	Optimize alternative fuel vehicle designs
Feit *et al.* [61]	2010	Correlated MIXL	Demand-side	No ASC	Vehicles	Disaggregate	Parameter distributions and out-of-sample market share predictions
Ahn *et al.* [103]	2008	MIXL[4]	Demand-side	Multivariate normal	Vehicles	Disaggregate	Counterfactual market shares
Sonnier *et al.* [104]	2007	MIXL[4]	Demand-side	Multivariate normal	Vehicles	Disaggregate	Willingness-to-pay distributions; optimal vehicle price
Train and Sonnier [105]	2005	Correlated MIXL	Demand-side	No ASC	Vehicles	Disaggregate	Parameter distributions

Note: Papers may include more than one model or data set; characteristics listed are for case study or model most relevant to this paper

[1] MIXL=mixed logit, MNL=multinomial logit

[2] Includes mathematical specifications for other correlation structures, but case study only includes results of independent ASCs

[3] DCM demand model is incorporated into larger model other than joint supply and demand

[4] Correlation structure not specified

4.3 Methods

To address the questions posed in section 4.1 we compare the forecasts from Bayesian and MLE estimated models for independent and correlated mixed logit structural specifications that include an ASC in the utility function. We estimate the models on combined vehicle attribute information from Ward's Automotive Index [73] and MSRP and aggregate sales data from Polk [74] for 2002-2006 US midsize sedans. We then compare the predictions of 2007 and 2011 US midsize sedan market shares to actual market shares in terms of accuracy and uncertainty.

The utility function is given by:

$$u_{ijt} = \mathbf{x}'_{jt}\boldsymbol{\beta}_i + \xi_{jt} + \varepsilon_{ijt} \tag{28}$$

where $\mathbf{x}_{jt}$ is a (K×1) vector of attributes specific to product j in market (year) t, $\boldsymbol{\beta}_i$ is a (K×1) vector of taste parameters for consumer i, ξ_{jt} is a random variable ASC representing the aggregate utility contribution of unobserved attributes of product j in market t, and ε_{ijt} is an idiosyncratic error term. Note that we use the terms market and year interchangeably, though they have different implications in some contexts. We assume the attribute vector $\mathbf{x}_{jt}$ for a given product to be similar but not necessarily constant in time or across markets. For example, the price or weight of a Ford Focus may vary slightly from year to year though it is still considered the same product[22].

We assume that the distribution of preferences is constant in time or across markets such that $\boldsymbol{\beta}$ is indexed only by i and not t. (This is a standard assumption in the vehicle demand literature. See Axsen *et al.* [41] for an exception.) The coefficients are assumed to follow a multivariate normal distribution:

[22] In the case study data the greatest year over year non-price attribute change is less than 15%.

$$\boldsymbol{\beta} \sim N\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\beta}\right) \tag{29}$$

where $\boldsymbol{\mu}$ is a (K×1) mean vector and $\boldsymbol{\Sigma}_{\beta}$ is a (K×K) covariance matrix. A general independent mixed logit specification $\boldsymbol{\Sigma}_{\beta}$ has off-diagonal elements equal to zero and a (K×1) diagonal element vector $\boldsymbol{\sigma}^2$. In the case study, two dummy variables are assumed to be constant (homogenous) across the population and their corresponding diagonal elements in $\boldsymbol{\Sigma}_{\beta}$ are 0. This is because MLE fails to converge to a reasonable solution when the dummy coefficients are permitted to vary over the population[23]. This is likely due to insufficient variation in the offered products and the resulting market shares across years so that this source heterogeneity is not well identified. In the correlated mixed logit case $\boldsymbol{\Sigma}_{\beta}$ is a full covariance matrix.

If ε_{ijt} is assumed to follow an independent and identically distributed (iid) extreme value type I distribution, the probability of individual i selecting a product j in market t is then given by the logit probability:

$$P_{ijt} = \frac{\exp\left(\mathbf{x}'_{jt}\boldsymbol{\beta}_i + \xi_{jt}\right)}{\sum_{k \in J_t} \exp\left(\mathbf{x}'_{kt}\boldsymbol{\beta}_i + \xi_{kt}\right)} \tag{30}$$

where J is the set of distinct products observed across all markets and J_t is a subset of J containing the products that appear in market t. We model product choice conditional on a consumer entering the market, thus we do not include an outside good in the model (option to not purchase any product).

[23] The Knitro solver for Matlab converges with a valid exit flag of zero, but with brand dummy $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}_{\beta}$ elements that are on the order of 100x larger than other coefficient estimates (all covariates are on the order of one) when they are not bounded in estimation. For bounded estimation they converge to the bounds. Though the estimation reaches a valid convergence, predictions from the resulting models are meaningless.

For aggregate data models, the share P_{jt} of product j in year t is typically obtained by numerical integration over population taste heterogeneity:

$$P_{jt} = \int_{\mathbf{y}} \frac{\exp\left(\mathbf{x}'_{jt}\mathbf{y} + \xi_{jt}\right)}{\sum_{k \in J_t} \exp\left(\mathbf{x}'_{kt}\mathbf{y} + \xi_{kt}\right)} f_{\beta}\left(\mathbf{y}\right)\left(d\mathbf{y}\right) \tag{31}$$

where $f(\circ)$ is the multivariate normal probability density function.

4.3.1 Maximum Likelihood Estimation

The model in Eq. 31 excluding the ASC can be estimated by MLE[24] and the ASCs can be calculated post-hoc [23]. The likelihood of the estimated parameters L is defined as the probability of generating the observed data given the estimated parameter values:

$$L\left(\mu, \sigma^2 | \mathbf{X}\right) = \prod_{t=1}^{T} \left(\prod_{k \in J_t} \left(P_{kt}\right)^{n_{kt}} \right) \tag{32}$$

where $\mathbf{X}$ is a $(V \times K)$ matrix of stacked transposed attribute vectors $\mathbf{x}_{jt}$ for all products in all markets, V is the total number of vehicles across all years in the data set, n_{kt} is the observed sales of vehicle k in time t, T is the total number of markets, and P_{jt} in Eq. 31 is modified to exclude ξ_{jt}. The MLE estimator of the parameters for the independent mixed logit model is the value of the $(2K \times 1)$ vector that maximizes L. The monotonic transformation $\ln(L)$ is typically used as the objective function for computational benefit. The ASCs are calibrated post-hoc by solving the system of equations:

[24] When the numerical integration is incorporated the more precise term is Simulated Maximum Likelihood Estimation (SMLE), though MLE is frequently used interchangeably in the literature.

$$
\begin{aligned}
&\ln\left(P_{jt}\left(\boldsymbol{\xi}_t \middle| \mathbf{X}_t, \boldsymbol{\mu}, \boldsymbol{\sigma}^2\right)\right) = \ln\left(s_{jt}\right), \forall j \in J_t^-, t \\
&\sum_{k \in J_t} \xi_{kt} = 0, \forall t
\end{aligned}
\tag{33}
$$

where $\boldsymbol{\xi}_t$ is the stacked vector of all ξ_{jt} in market t, $\mathbf{X}_t$ is the matrix of product attributes for all products in market t, s_{jt} is the observed share of product j in market t, P_{jt} includes the ASC as written in Eq. 31, and J_t^- is the set of products in market t excluding one. We constrain the ASCs so that they sum to zero for a given panel year since the utility of all vehicles can be shifted by an arbitrary constant if no outside good is present and the predicted shares will remain constant. For more detail on mixed logit models see Train [23].

4.3.2 Bayesian Estimation

We follow the procedure for Bayesian estimation of a mixed logit model on aggregate data proposed by Jiang *et al.* 2009 [109]. Some details are omitted here for conciseness, and interested readers are referred to the source. A key difference between the Bayesian framework and the frequentist literature on choice models is the treatment of the ASC. Under the Bayesian framework, the ASCs are treated similarly to the observed attribute coefficients and assumed to follow an iid parametric distribution:

$$
\xi_{jt} \sim N\left(0, \tau^2\right) \tag{34}
$$

That is, we assume that the ASCs are normally distributed across the population of vehicles. This additional assumption allows us to treat the ASCs as parameters in estimation.

The likelihood is rewritten to include this assumption:

$$L\left(\boldsymbol{\mu},\boldsymbol{\Sigma}_{\beta},\tau^{2}\right)=\prod_{t=1}^{T}\left(\pi\left(\mathbf{P}_{t}\left|\mathbf{X}_{t},\boldsymbol{\mu},\boldsymbol{\Sigma}_{\beta},\tau^{2},\mathbf{s}_{t}\right.\right)\right) \tag{35}$$

where $\mathbf{P}_t$ is the stacked vector of predicted market shares P_{jt} for market t, $\mathbf{s}_t$ is the stacked vector of observed market shares, and π is the joint conditional probability density.

Under the Bayesian paradigm we specify priors on $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}_{\beta}$, and τ^2. These priors are typically specified as diffuse parametric distributions so that the likelihood will dominate the posterior, producing parameter estimates that are similar to MLE estimates. Following Jiang *et al.* [109] we specify independent priors as follows:

$$\begin{aligned}&\boldsymbol{\mu}\sim MVN\left(\boldsymbol{\mu}_{0},\mathbf{V}_{\mu}\right)\\&\tau^{2}\sim \nu_{0}s_{0}^{2}/\chi_{\nu_{0}}^{2}\end{aligned} \tag{36}$$

where $\boldsymbol{\mu}_0$ and $\mathbf{V}_{\mu}$ are the respective mean and variance of the prior distribution on population mean taste parameter $\boldsymbol{\mu}$, and v_0 and s_0^2 are the respective degrees of freedom and scaling parameter of as scaled inverse chi-squared distribution.

The covariance matrix $\boldsymbol{\Sigma}_{\beta}$ is reparameterized as $\boldsymbol{\Sigma}_{\beta}$=$\mathbf{U}'\mathbf{U}$ where:

$$\mathbf{U}=\begin{bmatrix}e^{r_{11}} & r_{12} & \cdots & r_{1K}\\ 0 & e^{r_{22}} & \ddots & \vdots\\ \vdots & \ddots & \ddots & e^{r_{K-1,K}}\\ 0 & \dots & 0 & e^{r_{KK}}\end{bmatrix} \tag{37}$$

with associated priors:

$$\begin{aligned}&r_{mm}\sim N\left(0,\sigma_{r_{mm}}^{2}\right),\ m=1,\ldots,K\\&\mathrm{r}_{mn}\sim N\left(0,\sigma_{roff}^{2}\right),\ \ m,n=1,\ldots,K\ \ \forall m<n\end{aligned} \tag{38}$$

The vector $\mathbf{r}$ is the $[K(K+1)/2 \times 1]$ stacked vector of all r. We specify weakly informative priors so that the utility part-worth contributions are constrained to a reasonable interval to prevent the algorithm from exploring infeasible regions. See Appendix O for further discussion of the priors.

We want to express the posterior distribution of shares as a function of model parameters $\boldsymbol{\mu}$, $\mathbf{r}$, and τ^2. To do this we use the Change-of-Variables Theorem and define two additional quantities. Define $\mathbf{J}$ as the Jacobian of the predicted market shares at time t excluding one for identification (because an outside good is not included) with respect to vector $\boldsymbol{\xi}_t$. Define function h:

$$P_{jt} = h\left(\boldsymbol{\xi}_t \left| \mathbf{X}_t, \boldsymbol{\mu}, \boldsymbol{\Sigma}_\beta, \mathbf{s}_t \right.\right) \tag{39}$$

so that predicted shares are a function of only the variable $\boldsymbol{\xi}_t$ conditional on values of $\boldsymbol{\mu}$, and $\boldsymbol{\Sigma}_\beta$ and the data $\mathbf{X}_t$ and $\mathbf{s}_t$. The function h can be inverted by the nested fixed point algorithm outlined in Berry *et al.* [58], and each resulting $\boldsymbol{\xi}_t$ can be shifted by a constant so that the vector sums to zero. The joint posterior distribution of the parameters is then:

$$\begin{aligned}
&\pi\left(\boldsymbol{\mu},\mathbf{r},\tau^2 \left| \left\{\mathbf{s}_t,\mathbf{X}_t\right\}_{t=1}^{T}\right.\right) \propto L\left(\boldsymbol{\mu},\mathbf{r},\tau^2 \left| \left\{\mathbf{s}_t,\mathbf{X}_t\right\}_{t=1}^{T}\right.\right) \times \pi\left(\boldsymbol{\mu},\mathbf{r},\tau^2\right) \\
&= \prod_{t=1}^{T}\left(\mathbf{J}^{-1}\left(\mathbf{P}_t,\mathbf{r},\mathbf{X}_t,\right)\prod_{j\in J_t}\phi\left(\frac{h^{-1}\left(\mathbf{P}_t \left| \mathbf{X}_t,\boldsymbol{\mu},\mathbf{r},\mathbf{s}_t\right.\right)}{\tau}\right)\right) \times \left|\mathbf{V}_\mu\right|^{-1/2} \exp\left\{-\tfrac{1}{2}\left(\boldsymbol{\mu}-\boldsymbol{\mu}_0\right)' \mathbf{V}_\mu^{-1}\left(\boldsymbol{\mu}-\boldsymbol{\mu}_0\right)\right\} \quad (40) \\
&\times \prod_{m=1}^{M}\exp\left\{-\frac{r_{mm}^2}{2\sigma_{r_{mm}}^2}\right\} \times \prod_{m=1}^{K-1}\prod_{n=m+1}^{K}\exp\left\{-\frac{r_{mn}^2}{2\sigma_{r_{off}}^2}\right\} \times \left(\tau^2\right)^{-\left(\frac{\nu_0}{2}+1\right)} \exp\left\{-\tfrac{\nu_0 s_0^2}{2\tau^2}\right\}
\end{aligned}$$

The hybrid Markov-Chain Monte Carlo (MCMC) sampler proposed by Jiang *et al.* combines a Gibbs sampler to obtain draws of $\boldsymbol{\mu}$ and τ^2 from a univariate Bayes regression followed by a Random Walk (RW) Metropolis step to obtain $\mathbf{r}$. See Appendix O for more detail on specification and the MCMC algorithm.

4.3.3 Prediction

We estimate three models: (1) a Bayesian-estimated mixed logit model with a full covariance matrix ("Bayesian correlated model"), (2) a Bayesian-estimated mixed logit model with independent coefficients ("Bayesian independent model"), and (3) a maximum likelihood estimated mixed logit model with independent coefficients ("MLE model"). For each of these models we predict distributions of shares of the 1-year-forward and 5-year-forward vehicle markets. We include cases in which an ASC is predicted for forward markets and in which all future ASCs are equal to zero.

Bayesian methods account for uncertainty in parameters and forecasts with a posterior distribution for each parameter. For the Bayesian models, we use each of the estimation draws (excluding the burn-in) of $\boldsymbol{\mu}$, $\mathbf{r}$ and $\boldsymbol{\xi}$ to predict an implied posterior distribution of shares. In order to forecast shares we need draws of $\boldsymbol{\beta}$ and future $\boldsymbol{\xi}$. Obtaining draws of $\boldsymbol{\beta}$ are straightforward: posterior draws of $\boldsymbol{\Sigma}_{\beta}$ are calculated from draws of $\mathbf{r}$, and then 100 posterior draws of $\boldsymbol{\beta}$ are taken from $N(\boldsymbol{\mu},\boldsymbol{\Sigma}_{\beta})$. One hundred Halton draws from the standard normal distribution used to draw $\boldsymbol{\beta}$ are held constant throughout to eliminate sampling variation.

Drawing $\boldsymbol{\xi}$ for the prediction data set is less straightforward. For incumbent vehicles, we have estimates of how consumers may value the utility contribution of the ASC if it is assumed to represent aggregate unobserved attributes, and we forecast incumbent ASCs based on past estimates. For products that are newly introduced in the prediction data sets ("entrants") we propose four methods of forecasting ASCs described in Table 12. From the estimated $\boldsymbol{\xi}$ corresponding to each posterior draw of $\boldsymbol{\mu}$ and $\mathbf{r}$, we draw an ASC for each vehicle in the prediction data. In total there are 5 sets of predictions made (one for each of the entrant ASC prediction methods plus one in which all prediction ASCs are set to zero), and each set contains 49,600 implied draws of share (from 50,000 model parameter posterior draws less 400 burn-in draws). Our forecast uncertainty is represented by the range of share predictions we obtain from these 49,600 draws, and our point estimate for share is then obtained by averaging over the draws.

For the MLE model, we forecast ASCs by taking 50,000 draws from the nonparametric distributions described in Table 12. However, unlike the Bayesian models, each of the 50,000 ASC draws are drawn from the same estimated $\boldsymbol{\xi}$**.** For each of the 50,000 forecast ASCs, we simulate a draw of share by integrating over taste heterogeneity (Eq. 31) using 100 Halton draws of $\boldsymbol{\beta}_i$ from $N(\boldsymbol{\mu},\boldsymbol{\sigma}^2)$ where $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}^2$ are the likelihood maximizing parameters from Eq. 35. Our forecast uncertainty is represented by the range of share predictions from the 50,000 draws of the ASC, and our point estimate for share is again obtained by averaging over the draws.

Table 12 — Description of ASC forecasting methods

Method name	Method qualitative description	Method mathematical description
Products appearing in the estimation set (incumbents)		
Incumbent	For each product draw uniformly from the product's estimated ASCs	Draw ξ_{j6} or ξ_{j11} uniformly from $\{\xi_{=1},\xi_{j2},\xi_{j3},\xi_{j4}, \xi_{j5}\}$
Products not appearing in estimation set (entrants)		
All (Method 1)	Draw from uniformly from all estimated ASCs	Draw ξ_{j6} or ξ_{j11} uniformly from $\{\xi_{kt} \forall k \in J_t, t = 1,2,3,4,5\}$
Nearest neighbor (Method 2)	For each new product calculate the normalized vector distance of product observed attributes between the new product and each of the estimation data set products. Draw uniformly from the estimated ASCs of the observed product with the smallest vector distance ("nearest neighbor")	Draw ξ_{j6} or ξ_{j11} uniformly from $\{\xi_{k*1},\xi_{k*2},\xi_{k*3},\xi_{k*4},\xi_{k*5}\}$ where $k^* = argmin_{k \in J_t, t=1,2,3,4,5}(\lVert \mathbf{x}_{jt}^* - \mathbf{x}_{kt}^* \rVert)$ and $\mathbf{x}_{jt}^*$ is the (K×1) vector of normalized observed product attributes
Brand (Method 3)	Draw from all estimated ASCs of vehicles of the same brand	Draw ξ_{j6}^b or ξ_{j11}^b uniformly from $\{\boldsymbol{\xi}_1^b, \boldsymbol{\xi}_2^b, \boldsymbol{\xi}_3^b, \boldsymbol{\xi}_4^b, \boldsymbol{\xi}_5^b\}$ where the superscript b indexes brand (Chevy, Ford, Toyota, etc.)
Vehicle model (Method 4)	Draw from all estimated ASCs of the same vehicle trim (e.g. a Toyota Camry and a Toyota Camry Solara are assigned the same aggregate vehicle model for *prediction* but ASCs are *estimated* as if they are distinct)	Draw ξ_{j6}^m or ξ_{j11}^m uniformly from $\{\boldsymbol{\xi}_1^m, \boldsymbol{\xi}_2^m, \boldsymbol{\xi}_3^m, \boldsymbol{\xi}_4^m, \boldsymbol{\xi}_5^m\}$ where the superscript m indexes vehicle make-model

The accuracy of the predictions made by each model is compared using the relative average likelihood (RAL). The RAL is a monotonic transformation of the aggregate likelihood of the model share predictions L_{P} divided by the likelihood of an ideal model L_{I} that perfectly predicts the new shares:

$$RAL = \frac{\left(L_{\mathrm{P}}\right)^{1/N}}{\left(L_{\mathrm{I}}\right)^{1/N}} \tag{41}$$

where N is the number of choices observed. When comparing two models on the same data set, the model with a larger RAL is more likely to generate the data observed. Using RAL instead of likelihood is important because markets that have more diffuse choice probabilities will necessarily have lower likelihoods of ideal prediction. RAL normalizes for this effect, and can be interpreted as the fraction of the total possible explanatory power a model obtains. The RAL is calculated for each draw of share predictions to obtain a distribution of RAL (posterior distribution for the Bayesian models and a Monte Carlo distribution for the MLE model).

4.4 Results

4.4.1 Estimation Results

We estimate the random coefficient mixed logit model with a linear utility function including the covariates listed in Table 13 plus an ASC. The attributes capture price, fuel economy, performance, size, and country of origin. This choice of covariates is loosely based on Haaf *et al.* [93] who compare the market share forecasts of 9,000 possible logit model specifications informed by the vehicle demand literature and find predictive accuracy is insensitive to the form of the covariates (e.g. miles per gallon versus gallons per mile) so long as they are included in some form. We proxy dummies for brand (e.g. Honda, Ford, or Volkswagen) by dummies for producer firm geographic location (Europe or Asia, US is excluded for identification) because the dummies for brand are not separably identifiable from the ASC in our data set. The posterior means and standard deviations of the mean taste parameter vector $\boldsymbol{\mu}$ and taste heterogeneity parameter $\boldsymbol{\sigma}$ are shown in Table 13 for the Bayesian models, and the point estimates and standard errors of $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ for the MLE model are shown. The full estimated taste heterogeneity matrix $\boldsymbol{\Sigma}_{\beta}$ for the Bayes correlated model is presented in Appendix P. Appendix Q contains select trace plots for the Bayesian estimation routine.

The signs on the price, gallons/mile, and weight/horsepower coefficients for all three estimated models have intuitive signs, meaning consumers prefer vehicles with lower prices, greater fuel economy, and greater performance (HP/weight is a proxy), though the true values of the taste parameters for the population are unknown and could actually be positive. The three estimated models generally agree on coefficient sign and order of magnitude.

Bayes and MLE methods are asymptotically equivalent, meaning that for a large amount of data they are expected to result in the same coefficient estimates for a given model specification. Despite similar structural specifications, our estimates of population mean taste parameter $\boldsymbol{\mu}$ noticeably differ across models. These discrepancies occur for two reasons. Firstly, the ASCs are simultaneously estimated with the observed attribute coefficients for Bayesian estimation, but for MLE they are calibrated post-hoc; there are an additional 339 parameters in the specification of the model estimated by Bayesian methods over the model estimated by MLE. Secondly, our specification of the priors for the Bayesian models is not truly flat. Our data is not informative enough to entirely overcome the effect of the priors so the Bayesian estimates are influenced by this extra information that is not included in MLE estimation.

There is more uncertainty around the mean taste coefficients $\boldsymbol{\mu}$ (and heterogeneity coefficients) under Bayesian estimation than under MLE estimation; the standard deviations of the posterior distribution of the coefficients in the Bayesian models are greater than the standard errors of the MLE model coefficients. The ASCs included in Bayesian estimation are likely correlated with price so that we are including an additional 339 variables, and the collinearity may increase the posterior uncertainty. See a related frequentist argument in Greene [97]. Alternatively, if the sample size is too small then asymptotic standard errors for the MLE estimates may be in accurate. If price and ASC are correlated in the data, then both the Bayesian and MLE estimated price coefficients would be expected to exhibit bias due to endogeneity since no steps are taken to correct for it (e.g. incorporating instrumental variables (IVs)).

The MLE-estimated heterogeneity parameters $\boldsymbol{\sigma}$ are substantially smaller than those estimated by Bayesian methods. The Bayesian estimates are likely influenced by the priors because the heterogeneity parameters are not well identified by the data (as discussed in Chapter 3) and our Bayesian priors are especially informative for the heterogeneity parameters (see Appendix O).

Table 13 — Bayes and MLE model coefficient estimates

	Bayes full		Bayes independent		MLE	
	Mean	**Std. dev.**	**Mean**	**Std. dev.**	**Mean**	**Std. err.**
Mean taste parameters μ						
Price	**-1.0**	**0.2**	**-1.0**	**0.2**	-0.5***	0.001
Gallons/mile	-0.5	0.3	-0.6	0.3	-1.1***	0.001
Weight/HP	**-0.1**	**0.0**	**-0.1**	**0.0**	-0.1***	0.000
Length × width	**3.2**	**1.0**	**3.0**	**1.0**	3.9***	0.003
Europe	**-1.9**	**0.9**	**-1.0**	**0.3**	-1.1***	0.002
Asia	0.1	0.3	-0.1	0.2	0.3***	0.001
Heterogeneity taste parameters σ						
Price	0.310	0.120	0.290	0.080	0.003*	0.002
Gallons/mile	0.730	0.280	0.580	0.220	0.003*	0.000
Weight/HP	0.100	0.030	0.100	0.020	0.000*	0.080
Length × width	2.090	0.540	1.560	0.730	0.009**	0.000
Europe	1.660	0.550				
Asia	3.110	1.060				
Est. time (min.):	495		65		<1	

Note: US dummy omitted for identification; bolded coefficients indicate that 95% credible interval does not straddle zero; coefficients shown are zero to precision shown but not actually zero

* Coefficient is significant at α=0.10 level

** Coefficient is significant at α=0.05 level

*** Coefficient is significant at α=0.01 level

4.4.2 Prediction Results

Table 14 contains the RAL of the expected share for each of the models and ASC forecasting methods. Also included are the RALs of a ”static” model that holds all shares constant from the last observed value in the estimation data set and divides the remaining market share equally across new products and a “no info” model that assigns all products an equal share (equivalent to random guessing). Highlighted cells indicate the attribute-based model that predicts best for a given time period and ASC forecasting method.

Table 14 — RAL(E[share]) comparison of ASC forecasting methods for Bayesian full and independent and MLE models estimated on 2002-2006 midsize sedan data and used to predict 2007 and 2011 midsize sedan market shares

	1-year-forward (2007) forecasts					**5-year-forward (2011) forecasts**				
Method:	**No ASC 0**	**All 1**	**Near neighbor 2**	**Brand 3**	**Model 4**	**No ASC 0**	**All 1**	**Near neighbor 2**	**Brand 3**	**Model 4**
RAL of expected share — RAL(E[**P**(**β**,ξ)])										
Bayes full	41%	72%	68%	73%	74%	36%	42%	42%	41%	41%
Bayes ind.	38%	71%	69%	71%	72%	35%	42%	41%	40%	41%
MLE	45%	66%	61%	63%	66%	38%	42%	60%	37%	41%
Static	*68%*					*39%*				
No info	*32%*					*42%*				
RAL of expected share ENTRANTS ONLY — RAL(E[**P**(**β**,ξ)])										
Bayes full	90%	89%	83%	90%	91%	51%	54%	52%	53%	52%
Bayes ind.	88%	88%	85%	88%	90%	50%	52%	51%	50%	52%
MLE	91%	91%	84%	88%	91%	49%	50%	70%	43%	49%
Static	*87%*					*54%*				
No info	*87%*					*62%*				

Note: highlighted cells indicate the most accurate model and ASC forecasting method for a given time period and means of calculating RAL

(Q4.1) How do the accuracy and uncertainty of share predictions from a full covariance mixed logit model compare to an independent mixed logit model?

The Bayesian correlated and independent models predict similarly accurately for both 2007 and 2011 for all methods of ASC forecasting (RAL differences of only 1-3%). Both Bayesian models predict better than the MLE model in the short term, but the MLE model using the nearest neighbor ASC forecast method (method 2) forecasts best in the long term. When no ASC is included in prediction years, the MLE model predicts best, followed by the Bayesian correlated model, and then the Bayesian independent model.

(Q4.2) How do the share forecasts compare under the Bayesian versus MLE treatment of ASCs?

The MLE model forecasts reflect uncertainty from *forecasting* ASCs only, whereas the Bayesian model forecasts additionally reflect the joint uncertainty in observed attribute coefficients and *estimated* ASCs. (The straightforward ability to incorporate the joint parameter estimation uncertainty is an advantage of Bayesian methods.) Despite the additional sources of uncertainty in the Bayes model predictions, the overall share forecast uncertainty is comparable between the Bayes and MLE models since the majority of it results from forecasting ASCs. (See actual versus predicted share plots in Appendix R). Forecast accuracy of all models is improved by including an ASC in prediction years, but only the MLE model is improved enough by the forecast ASC to be superior to random guessing (no info model) in the long term.

The MLE model is sensitive to choice of ASC forecasting method, but the Bayes models are not. This is a result of the parameter estimation uncertainty that is propagated to the Bayes model forecasts. The nearest neighbor ASC forecasting method introduces the least amount of uncertainty of all ASC forecasting methods, and the reduction in uncertainty has greater impact on share forecast uncertainty when it is not confounded by parameter estimation uncertainty as in the Bayes model forecasts.

Table 14 also contains the RALs calculated on entrant vehicles only. For this calculation, all incumbent vehicle market shares are summed and represented as single "non-entrant" market share. The conclusions are similar to those of the full market, but Bayes model forecasts are relatively less superior than the MLE model forecasts in the short term.

4.5 Discussion

A correlated mixed logit did not meaningfully improve predictions relative to the independent mixed logit. Treating ASCs as model parameters in Bayesian estimation results in better short term forecasts than calibrating them post-hoc as in MLE estimation, but the best long term forecasts are made by the MLE model.

If MLE and nearest neighbor forecasted ASCs yield the best long term expected forecasts, then should these be the estimation and ASC methods of choice for modelers?

Not necessarily. For this particular data set, vehicles' observed characteristics are good indictors of what the aggregate utility value of their unobserved attributes is to consumers, but this need not be true. For example, an auto manufacturer may choose to differentiate a midsize sedan from the nearest competitor sedans through the unobserved attribute(s) for competitive purposes. In this case, the nearest neighbor method would be misleading in terms of determining entrant ASCs. For this data, the Bayesian model offers an advantage in that it is more robust to selection of ASC generating method, and less reliant on persistent market structures. If the brand method is used to forecast ASCs then the MLE model predicts worse than excluding an ASC entirely, though the difference is slight.

The main disadvantage of the Bayesian estimator is the difficulty of implementation relative to the familiar and straightforward MLE estimator. There are only a handful of studies in the literature that address Bayesian estimation of a DCM on aggregate data sets, and it is a relatively new— and therefore untested— area of exploration with fewer resources for modelers who are not already experts in Bayesian statistics. Estimation takes much longer for data sets of the structure tested here (few markets with many products per market) by orders of magnitude over the MLE model (see Table 13). Our Bayesian estimator converged relatively quickly, so we could have made fewer draws and reduced the estimation run time, but this is not known *a priori* and it would still not approach the ~30 seconds the MLE estimator required.

Table 15 summarizes the comparison of the estimators.

Table 15 — Comparison of MLE and Bayes properties and findings

	MLE	Bayes
Uncertainty	No share uncertainty from observed coefficients; uncertainty equal to that of Bayes when predictive ASCs are included	Share uncertainty from predictive ASCs greater than uncertainty from observed coefficients
Prediction	Most accurate long term expected share predictions given correct ASC forecasting method	Most accurate for short term predictions
Computation	Fast estimation	Longer estimation times; more difficult to implement than MLE methods
ASC	Performance relative to excluding ASC is sensitive to ASC forecasting method	Robust to selection of ASC forecasting method; draws of predicted ASCs are made jointly with observed coefficient draws

4.6 Conclusion

Bayesian estimation enables modelers to estimate more flexible DCM structural specifications and to obtain distributions of functions of model parameters directly rather than relying on Monte Carlo sampling from asymptotic distributions. We use Bayesian estimation to fit both correlated mixed logit and independent mixed logit models and MLE to estimate an independent mixed logit model. The models are fit to 2002-2006 US new midsize sedan aggregate sales data and used to predict shares of the 2007 and 2011 markets. Forecasts from the three estimated models are compared to investigate the effects on prediction accuracy and uncertainty of a more flexible structural specification (correlated mixed logit versus independent mixed logit) and richer description of the ASC (treatment as an estimation parameter with full distribution of uncertainty versus a post-hoc calibration constant).

The additional structural specification flexibility permitted by Bayesian estimation was not advantageous for prediction, but additional uncertainty from parameter estimation is more easily reflected in forecast share uncertainty. Aggregate data sets with limited variation in market shares— like the type used here— are not well suited to Bayesian estimation and we recommend modelers carefully evaluate the suitability of their data set.

5. Summary and Conclusions

This thesis explores the use of discrete choice models (DCMs) for forecasting automotive demand with a focus on engineering and policy contexts. I address several questions regarding the accuracy and uncertainty of market share predictions resulting from choice of utility function and structural specification, estimation method, and data structure assumptions (i.e. endogeneity of price and the alternative-specific constant (ASC) and treatment of the ASC as a representation of unobserved vehicle attributes).

The three studies address three major research themes: defining how to measure forecast accuracy, comparing model specification and forecasting methods in terms of prediction accuracy, and comparing design implications of different model specifications and forecasting methods. With regards to defining how to measure forecast accuracy, I find in study 1 that for the automotive case study examined, determination of the best models did not depend strongly on potentially arbitrary selection of the measure used to evaluate predictive accuracy. I pose the relative average likelihood (RAL) as an intuitive likelihood measure and use it primarily in further studies. I also find that better model fit correlates well with better predictive accuracy. The match between fit and predictive accuracy, suggesting no major overfitting issues, is particularly encouraging, since the modeler has access to choice data for estimation but not choice data in the counterfactual predictive context.

I address the second research theme— comparing models and forecasting methods on their forecast accuracy— by examining utility function and structural specifications. I find in study 1 that including more covariates generally improves predictive accuracy. While including an appropriate set of product attributes as model covariates is important, the form those covariates take (e.g. miles/gallon versus gallons/mile) in the utility function is less important in this application.

Including ASCs in product utility functions of DCMs improves model fit but not necessarily forecast accuracy. In study 2, all proposed ASC prediction methods improve forecasts of the midsize sedan market over models that exclude ASCs entirely. However, in evaluating entrant vehicles, only the "nearest neighbor" ASC prediction method improves forecasts. If the relationships between observed attributes and ASCs change as product turnover from estimation to prediction data sets increases, forecasting ASCs by "nearest neighbor" is risky since falsely assuming this relationship may result in worse forecasts than excluding the ASC, as seen in the synthetic data study. Furthermore, my methods for forecasting ASCs rely on their estimated values, which is valid only if there is persistence in the utility of unobserved attributes.

If the ASC is treated as a model parameter to be estimated complete with a full distribution of uncertainty, like under Bayesian estimation in study 3, then forecasting future ASCs generally improves long term expected share predictions relative to excluding the ASC from predictions. This treatment is also more robust to selection of ASC generating method and less reliant on persistent market structures. However, the best long term expected share predictions over all estimation and ASC forecasting methods are from the frequentist MLE estimated model that treats ASCs as post-model-estimation calibration constants.

I address the third research theme— comparing models and forecasting methods on their implications for design— in studies 1 and 2. In study 1 the limited predictive power of standard models on real data in a canonical product category suggests designers should apply DCMs cautiously, though predictions may be substantially better in domains with fewer unobserved attributes or with conjoint data (where all attributes are observed).

In study 2 GMM-IV methods are used in an effort to mitigate the potential endogeneity bias of the price coefficient when ASCs are included in the utility function. Inconsistent coefficients are of particular interest for designers since they affect estimates of consumer willingness-to-pay for vehicle attributes. I observe that the GMM-IV models based on invalid instruments did indeed predict worse than the MLE-C models in the synthetic data study, and there is evidence that commonly used instruments in real automotive demand models (e.g. instruments grounded in those of BLP) are invalid. Invalidity of instruments is a significant issue for GMM-IV methods in practice because instrument validity is impossible to verify.

For purely predictive purposes, the drawbacks of GMM-IV— challenging model estimation and difficult instrument specification— outweigh the expected benefits of potentially mitigating price endogeneity. The risk of specifying invalid instruments is high, and my results suggest that unbiased coefficients are not necessary for maximizing the *predictive* power of the model. Bias can even *aid* predictions by implicitly capturing persistent unobserved effects in some circumstances.

5.1 Contributions

I assess and test classes of existing methods for forecasting automotive demand using DCMs in order to understand which methods are more theoretically grounded and which perform better in practice.

Contributions from study 1 (Chapter 2) include an exhaustive evaluation of the utility function covariates found in the automotive demand literature and a comparison of measures of model fit and prediction accuracy, including one proposed here, the cumulative distribution function of error tolerance (CDFET).

Contributions from study 2 (Chapter 3) are two-fold. Firstly, the application of a canonical DCM estimation technique frequently used in automotive demand contexts (GMM-IV as proposed by Berry *et al.* [58]) is evaluated on its suitability for DCM models used to forecast. Secondly, I investigate whether ASCs should be interpreted as pure error terms or representations of unobserved vehicle attributes and propose several means for predicting future ASCs of vehicles not observed in the estimation data.

The contribution from study 3 (Chapter 4) is primarily a comparison of the predictive accuracy when the ASC is treated as a parameter to be estimated complete with an uncertainty distribution as opposed to a point estimate in model estimation and prediction. I also compare the relative uncertainty in forecasts resulting from the use of frequentist (MLE) versus Bayesian estimators.

5.2 Recommendations for Future Work

For study 1, future work could include evaluating the sensitivity of predictions to inclusion of individual-level choice data with consumer covariates, such as demographics or usage variables, in the utility function and alternative structural specifications like latent class and generalized logit models. It would be interesting to extend the case study to include the design of an optimal vehicle under various utility function and structural specifications to better understand the implications of share prediction uncertainty on vehicle design decisions.

Future work for study 2 includes estimating a model by MLE and incorporating IVs via a control function as in [47]. Ideally the case study would include not only prediction of *future* market shares but also prediction of market shares under counterfactual scenarios. The unbiased coefficients of the GMM-IV estimator, if valid, may prove better for counterfactual forecasts than the MLE-C coefficients.

Future work for study 3 could include the addition of IVs to evaluate their behavior in a Bayesian framework, but similar results to study 2 are anticipated.

In addition to the study-specific future work, areas of exploration for DCM forecasting include alternative assumptions to treating consumers as observant rational utility maximizers with consistent preferences, consumer choice set definition, and longer forecast time horizons. Additional types and sources of data like disaggregate household purchase information and stated preference surveys in combination with revealed preference data may yield more accurate forecasts. The models tested here represent the demand-side of the market only, and incorporating a joint supply-side model may make the results of thesis more applicable to marketing researchers.

This work forecasts shares under a multitude of DCM structures, but stops short of contextualizing those forecasts in terms of decision outcomes. Comparing the profit-optimal design of a new vehicle predicted by models with different utility function or structural specifications as suggested in the future work of study 1 is one example. Recent federal vehicle emission legislation includes a regulatory impact analysis to evaluate manufacturer compliance implementation costs, likely vehicle pricing responses, and the resulting consumer welfare [111], and DCMs could be used to predict the downstream implications of market share distributions on these estimated costs and benefits.

6. References

[1] Environmental Protection Agency, 2010, Light-Duty Vehicle Greenhouse Gas Emissions Standards and Corporate Average Fuel Economy Standards; Final Rule, Title 40 Code of Federal Regulations, Pts. 85, 86, 600. Title 49 Code of Federal Regulations Pts. 531, 533, 536 et al. 2010 ed.

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

[3] 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," J. Mech. Des., **134**(10), p. 100901.

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

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

[6] Patel P., and Clarke L., 2012, "Global Change Assessment Model Documentation," Univ. Maryl. wiki [Online]. Available: http://wiki.umd.edu/gcam/index.php?title=Transportation&oldid=350. [Accessed: 29-Oct-2014].

[7] TA Engineering Inc., 2012, LVChoice: Light Vehicle Market Penetration Model Documentation, Baltimore, MD.

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

[9] Bass F. M., 1969, "A New Product Growth Model for Consumer Durables," Manage. Sci., **15**(5), pp. 215–227.

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

[11] National Research Council, 2010, Transitions to Alternative Transportation Technologies- Plug-in Hybrid Electric Vehicles, Washington, D.C.

[12] Argonne National Laboratory, 2007, "The VISION Model."

[13] U.S. Energy Information Administration, 2009, Assumptions to the Annual Energy Outlook 2008 with Projections for 2030, Washington, D.C.

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

[15] 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.

[16] U.S. Department of Energy Office of Energy Efficiency and Renewable Energy, 2007, Plug-In Hybrid Electrical Vehicle R&D Plan, Washington, D.C.

[17] Zhang T., Gensler S., and Garcia R., 2011, "A Study of the Diffusion of Alternative Fuel Vehicles: An Agent-Based Modeling Approach*," J. Prod. Innov. Manag., **28**(2), pp. 152–168.

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

[19] Payne J. W., 1976, "Task Complexity and Contingent Processing in Decision Making: An Information Search," Organ. Behav. Hum. Perform., **16**, pp. 366–387.

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

[21] Environmental Protection Agency, 2012, Preamble to Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy MY2017-2025.

[22] Hazelrigg G. A., 1998, "A framework for decision-based engineering design," J. Mech. Des., **120**(4), pp. 653–658.

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

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

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

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

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

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

[29] He L., Chen W., Hoyle C., and Yannou B., 2012, "Choice Modeling for Usage Context-Based Design," J. Mech. Des., **134**(3), p. 031007.

[30] Morrow W. R., Long M., and MacDonald E. F., 2014, "Market-System Design Optimization with Consider-Then-Choose Models," J. Mech. Des., **136**(3).

[31] 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**(1), pp. 402–411.

[32] MacDonald E. F., Gonzalez R., and Papalambros P. Y., 2009, "Preference Inconsistency in Multidisciplinary Design Decision Making," J. Mech. Des., **131**(3), p. 031009.

[33] Knittel C. R., and Metaxoglou K., 2012, "Estimation of Random Coefficient Demand Models: Two Empiricists' Perspective," Rev. Econ. Stat.

[34] Allcott H., and Wozny N., 2012, "Gasoline prices, fuel economy, and the energy paradox," Rev. Econ. Stat., (Forthcoming).

[35] Dubé J.-P., Fox J. T., and Su C.-L., 2012, "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," Econometrica, **80**(5), pp. 2231–2267.

[36] Beresteanu A., and Li S., 2011, "Gasoline Prices, Government Support, and the Demand for Hybrid Vehicles in the United States," Int. Econ. Rev. (Philadelphia)., **52**(1), pp. 161–182.

[37] Copeland A., Dunn W., and Hall G., 2011, "Inventories and the automobile market," RAND J. Econ., **42**(1), pp. 121–149.

[38] Hoyle C., Chen W., Wang N., and Gomez-Levi, 2011, "Understanding and modelling heterogeneity of human preferences for engineering design," J. Eng. Des., **22**(8), pp. 583–601.

[39] Li S., Liu Y., and Zhang J., 2011, "Lose Some, Save Some: Obesity, Automobile Demand, and Gasoline Consumption in the United States," J. Environ. Econ. Manage., **61**(1), pp. 52–66.

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

[41] 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," Resour. Energy Econ., **31**(3), pp. 221–238.

[42] Dagsvik J. K., and Liu G., 2009, "A Framework fo Analyzing Rank-ordered Data with Application to Automobile Demand," Transp. Res. Part A Gen., **43**(1), pp. 1–12.

[43] Hoyle C., Chen W., Ankenman B., and Wang N., 2009, "Optimal Experimental Design of Human Appraisals for Modeling Consumer Preferences in Engineering Design," J. Mech. Des., **131**(7), p. 071008.

[44] Shiau C.-S. N., and Michalek J. J., 2009, "Should Designers Worry About Market Systems?," J. Mech. Des., **131**(1), p. 011011.

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

[46] Dasgupta S., Siddarth S., and Silva-Risso J., 2007, "To Lease or to Buy? A Structural Model of a Consumer's Vehicle and Contract Choice Decisions," J. Mark. Res., **44**(3), pp. 490–502.

[47] Train K., and Winston C., 2007, "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers," Int. Econ. Rev. (Philadelphia)., **48**(4), pp. 1469–1497.

[48] Kumar D., Kim H. M., and Chen W., 2006, "Multi-level Optimization for Enterprise-Driven Decision-Based Product Design," Decision Making in Engineering Design, K. Lewis, W. Chen, and L.C. Schmidt, eds., ASME Press, New York, NY, pp. 203–216.

[49] Berry S., Levinsohn J., and Pakes A., 2004, "Differentiated Products Demand Systems from a Combination of Micro and Macro Data : The New Car Market," J. Polit. Econ., **112**(1), pp. 68–105.

[50] 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.

[51] 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," J. Mech. Des., **126**(6), p. 1062.

[52] Petrin A., 2002, "Quantifying the Benefits of New Products: The Case of the Minivan," J. Polit. Econ., **110**(4), pp. 705–729.

[53] Sudhir K., 2001, "Competitive pricing behavior in the auto market: A structural analysis," Mark. Sci., **20**(3), pp. 42–60.

[54] Berry S., Levinsohn J., and Pakes A., 1999, "Voluntary export restraints on automobiles: Evaluating a trade policy," Am. Econ. Rev., **89**(3), pp. 400–430.

[55] Brownstone D., and Train K., 1999, "Forecasting new product penetration with flexible substitution patterns," J. Econom., **89**, pp. 109–129.

[56] Dagsvik J. K., Wennemo T., Wetterwald D. G., and Aaberge R., 2002, "Potential demand for alternative fuel vehicles," Transp. Res. Part B Methodol., **36**(4), pp. 361–384.

[57] McCarthy P. S., 1996, "Market price and income elasticities of new vehicle demands," Rev. Econ. Stat., **78**(3), pp. 543–547.

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

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

[60] Lave C., and Train K., 1979, "A Disaggregate Model of Auto-type Choice," Transp. Res. Part A Gen., **13**(1), pp. 1–9.

[61] Feit E. M., Beltramo M. a., and Feinberg F. M., 2010, "Reality Check: Combining Choice Experiments with Market Data to Estimate the Importance of Product Attributes," Manage. Sci., **56**(5), pp. 785–800.

[62] Bunch D., Greene D., Lipman T., and Shaheen S., 2011, Potential Design, Implementation, and Benefits of a Feebate Program for New Passenger Vehicles in California, Davis, CA.

[63] Gigerenzer G., and Brighton H., 2009, "Homo Heuristicus: Why Biased Minds Make Better Inferences," Top. Cogn. Sci., **1**(1), pp. 107–143.

[64] Train K., 1979, "Consumers' responses to fuel-efficient vehicles: a critical review of econometric studies," Transportation (Amst)., **8**(3), pp. 237–258.

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

[66] 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," Transp. Res. Part A Policy Pract., **43**(9-10), pp. 814–828.

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

[68] Skrainka B., and Judd K., 2011, High Performance Quadrature Rules: How Numerical Integration Affects a Popular Model of Product Differentiation.

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

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

[71] McFadden D., and Train K., 2000, "Mixed MNL Models for Discrete Response," J. Appl. Econ., **15**, pp. 447–470.

[72] Selby B., and Kockelman K. M., 2012, "Microsimulating Automobile Markets: Evolution of Vehicle Holdings and Vehicle-Pricing Dynamics," Transp. Res. Forum, **51**(2), pp. 83–96.

[73] Ward's Automotive Group, 2012, "Ward's Yearly Vehicle Specifications," Ward's Yrly. Veh. Specif. [Online]. Available: http://wardsauto.com/data-center.

[74] Polk, 2012, "Automotive Sales Data," IHS Polk Insight Yrly. Vol. Data [Online]. Available: http://blog.polk.com/blog/new-vehicle-sales.

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

[76] Kullback S., and Leibler R. A., 1951, "On information and sufficiency," Ann. Math. Stat., **22**(1), pp. 79–86.

[77] Akaike H., 1973, "Information Theory and an Extension of the Maximum Likelihood Principle," Proceeding of the Second International Symposium on Information Theory, B.N. Petrov, and F. Csaki, eds., Akademiai Kiado, Budapest, Hungary, pp. 267–281.

[78] Schwarz G., 1978, "Estimating the dimension of a model," Ann. Stat., **6**(2), pp. 461–464.

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

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

[81] Heffner R. R., Kurani K. S., and Turrentine T. S., 2007, "Symbolism in California's early market for hybrid electric vehicles," Transp. Res. Part D Transp. Environ., **12**(6), pp. 396–413.

[82] 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," Transp. Res. Part D Transp. Environ., **16**(2), pp. 150–159.

[83] Hauser J., and Wernerfelt B., 1990, "An Evaluation Cost Model of Consideration Sets," J. Consum. Res., **16**(4), pp. 393–408.

[84] Gilbride T. J., and Allenby G. M., 2004, "A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules," Mark. Sci., **23**(3), pp. 391–406.

[85] Lucas, Robert E., 1976, "Econometric policy evaluation: A critique," Carnegie-Rochester Conference Series on Public Policy, K. Brunner, and A. Meltzer, eds., Elsevier Science Publishers, North-Holland, MI, pp. 19–46.

[86] 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.

[87] Swait J., Louviere J., and Williams M., 1994, "A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice," Transportation (Amst)., **21**, pp. 135–152.

[88] Berry S., 1994, "Estimating discrete-choice models of product differentiation," RAND J. Econ., **25**(2), pp. 242–262.

[89] Villas-Boas J. M., and Winer R. S., 1999, "Endogeneity in Brand Choice Models," Manage. Sci., **45**(10), pp. 1324–1338.

[90] Chintagunta P. K., 2001, "Heterogeneity Endogeneity and Demand Model: Probit Estimation Using Aggregate Data," **20**(4), pp. 442–456.

[91] Rossi P., 2013, "Even the Rich Can Make Themselves Poor: The Dangers of IV and Related Methods in Marketing Applications," Available SSRN, pp. 1–53.

[92] Wooldridge J. M., 2010, Econometric Analysis of Cross Section and Panel Data, MIT Press, Cambridge, Massachusetts.

[93] Haaf C. G., Michalek J. J., Morrow W. R., and Liu Y., 2014, "Sensitivity of Vehicle Market Share Predictions to Discrete Choice Model Specification," J. Mech. Des., **126**(12).

[94] Su C.-L., and Judd K. L., 2012, "Constrained Optimization Approaches to Estimation of Structural Models," Econometrica, **80**(5), pp. 2213–2230.

[95] Jenn A., Azevedo I., and Ferreira P., 2013, "The impact of federal incentives on the adoption of hybrid electric vehicles in the United States," Energy Econ., **40**, pp. 936–942.

[96] 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," Transp. Res. Part A Policy Pract., **38**(3), pp. 201–222.

[97] Greene W. H., 2003, Econometric Analysis, Prentice Hall, Upper Saddle River, NJ.

[98] Min J., Azevedo I. L., Michalek J., and Bruine de Bruin W., 2014, "Labeling energy cost on light bulbs lowers implicit discount rates," Ecol. Econ., **97**, pp. 42–50.

[99] Daziano R. A., 2013, "Conditional-logit Bayes estimators for consumer valuation of electric vehicle driving range," Resour. Energy Econ., **35**(3), pp. 429–450.

[100] Daziano R., and Achtnicht M., 2013, "Forecasting adoption of ultra-low-emission vehicles using Bayes estimates of a multinomial probit model and the GHK simulator," Transp. Sci., (January).

[101] Daziano R. A., and Bolduc D., 2013, "Incorporating pro-environmental preferences towards green automobile technologies through a Bayesian hybrid choice model," Transp. A Transp. Sci., **9**(1), pp. 74–106.

[102] Daziano R. A., and Chiew E., 2013, "On the effect of the prior of Bayes estimators of the willingness to pay for electric-vehicle driving range," Transp. Res. Part D Transp. Environ., **21**, pp. 7–13.

[103] Ahn J., Jeong G., and Kim Y., 2008, "A forecast of household ownership and use of alternative fuel vehicles: A multiple discrete-continuous choice approach," Energy Econ., **30**(5), pp. 2091–2104.

[104] Sonnier G., Ainslie A., and Otter T., 2007, "Heterogeneity distributions of willingness-to-pay in choice models," Quant. Mark. Econ., **5**(3), pp. 313–331.

[105] Train K., and Sonnier G., 2005, “Mixed Logit with Bounded Distributions of Correlated Parthworths,” Appl. Simul. Methods Environ. Resour. Econ., pp. 117–134.

[106] Musalem A., Bradlow E. T., and Raju J. S., 2009, “Bayesian estimation of random-coefficients choice models using aggregate data,” J. Appl. Econom., **24**, pp. 490–516.

[107] Yonetani Y., Kanazawa Y., Myojo S., and Turnbull S. J., 2010, A Bayesian simultaneous demand and supply model for aggregate data in a differentiated product market.

[108] Yang S., Chen Y., and Allenby G. M., 2003, “Bayesian Analysis of Simultaneous Demand and Supply,” Quant. Mark. Econ., **1**(3), pp. 251–275.

[109] Jiang R., Manchanda P., and Rossi P. E., 2009, “Bayesian Analysis of Random Coefficient Logit Models Using Aggregate Data,” J. Econom., **149**(2), pp. 136–148.

[110] Romeo C. J., 2007, “A Gibbs Sampler for Mixed Logit Analysis of Differentiated Product Markets Using Aggregate Data,” Comput. Econ., **29**(1), pp. 33–68.

[111] Environmental Protection Agency, 2012, Regulatory Impact Analysis: Final Rulemaking for 2017-2025 Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards.

[112] U.S. Energy Information Administration, 2010, Annual Energy Outlook 2010 With Projections to 2035, Washington, D.C.

[113] U.S. Energy Information Administration, 2010, Transportation Sector Module of the National Energy Modeling System : Model Documentation 2010.

[114] Park S., and Gupta S., 2011, “Comparison of SML and GMM estimators for the random coefficient logit model using aggregate data,” Empir. Econ., **43**(3), pp. 1353–1372.

[115] U.S. Energy Information Administration, 2013, “U.S. All Grades All Formulations Retail Gasoline Prices” [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].

[116] “Annual Average Indexex,” Bur. Labor Statitstics Data [Online]. Available: http://www.bls.gov/cpi/cpi_dr.htm#2007. [Accessed: 01-Aug-2013].

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

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

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

[120] 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.

7. Appendices

7.1 Appendix A

Comparison of predictive and explanatory vehicle demand literature

Table 16 — Explanatory and predictive literature

		Author type				**Goal**		**Study**		
Author (Individual or Institute)	Year	Academic	Government	Other research (note 1)	Consulting firm	Explanatory	Predictive	Reproducible from documentation	Peer-reviewed journal publication	Uses discrete choice model
Lave and Train [60]	1979	x				x		x	x	x
Boyd and Mellman [59]	1980				x	x		x	x	x
Berry *et al.* [58]	1995	x				x		x	x	x
McCarthy [57]	1996	x				x		x	x	x
Brownstone and Train [55]	1999	x				x		x	x	x
Electric Power Research Institute [117]	2001			x		x				See note 2
Dagsvik *et al.* [56]	2002		x			x		x		x
Choo and Mokhtarian [96]	2004	x				x		x	x	x
Oak Ridge National Laboratory [15]	2004		x				x	x		x
Greene *et al.* [50]	2005		x			x		x	x	x
Electric Power Research Institute [4]	2007			x			x			See note 2
Train and Winston [47]	2007	x				x		x	x	x
Nat.'l Research Council, Nat.'l Academies [10]	2008			x			x			
Pacific Northwest National Laboratory [14]	2008		x				x			See note 3
Center for Entrepreneurship and Technology [8]	2009	x					x			
Dagsvik and Liu [42]	2009		x			x		x	x	x
Lin and Greene [86]	2009		x				x	x	See note 4	x
Vance and Mehlin [45]	2009	x				x		x		x
Frischknecht *et al.* [20]	2010	x				x		x	x	x
Argonne National Laboratory [118]	2011		x			x		x		x
Electric Power Research Institute [119]	2011			x			x			
Energy Information Administration [5]	2011		x				x	x		x
Musti and Kockelman [40]	2011	x				x		x	x	x
Zhang *et al.* [17]	2011	x				x		x	x	x
Whitefoot and Skerlos [31]	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 [15] report, so it is at least partially based on choice modeling but extension methodology is not explicitly described; 4.) Published in conference proceedings

7.2 Appendix B

Comparison of predictive and explanatory vehicle demand literature

Table 17 — Literature demand covariates

Demand covariates	Lave 1979 [60]	Boyd 1980 [59]	Berry 1995 [58]	Mc-Carthy 1996 [57]	Brown-stone 1999 [55]	Dags-vik 2002 [56]	ORNL 2004 [15]	ANL 2005 [120]	Greene 2005 [50]	Train 2007 [47]	Dags-vik 2009 [42]	Vance 2009 [45]	Frischk-necht 2010 [20]	EIA AEO 2011 [5]	Musti 2011 [40]	Zhang 2011 [17]	White-foot 2012 [31]
Price																	
price						x	x	x	x	x				x	x		x
price + fuel cost/50000mi																	
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											
1/mpg										x	x		x				x
levels of mpg																x	
levels of miles/charge																x	
NPV of fuel savings									x								

Demand covariates	Lave 1979	Boyd 1980	Berry 1995	Dags-vik 1996	Mc-Carthy 1996	Brown-stone 1999	ORNL 2004	ANL 2005	Greene 2005	Train 2007	Dags-vik 2009	Vance 2009	Frischk-necht 2010	EIA AEO 2011	Musti 2011	Zhang 2011	White-foot 2012
Maintenance cost																	
repair rating																	
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																	
f(hp/wt)														x			
known seconds						x											
Other performance																	
handling rating																	
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										

Demand covariates	Lave 1979	Boyd 1980	Berry 1995	Dags-vik 1996	Mc-Carthy 1996	Brown-stone 1999	ORNL 2004	ANL 2005	Greene 2005	Train 2007	Dags-vik 2009	Vance 2009	Frischk-necht 2010	EIA AEO 2011	Musti 2011	Zhang 2011	White-foot 2012
Intangibles																	
style: (length+width)/height																	
luxury: noise rating																	
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											
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										

Demand covariates	Lave 1979	Boyd 1980	Berry 1995	Dags-vik 1996	Mc-Carthy 1996	Brown-stone 1999	ORNL 2004	ANL 2005	Greene 2005	Train 2007	Dags-vik 2009	Vance 2009	Frischk-necht 2010	EIA AEO 2011	Musti 2011	Zhang 2011	White-foot 2012
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		

Demand covariates	Lave 1979	Boyd 1980	Berry 1995	Dags-vik 1996	Mc-Carthy 1996	Brown-stone 1999	ORNL 2004	ANL 2005	Greene 2005	Train 2007	Dags-vik 2009	Vance 2009	Frischk-necht 2010	EIA AEO 2011	Musti 2011	Zhang 2011	White-foot 2012
Transaction																	
search process					x					x							
financing										x							
Data source	**SP**	**P**	**RP**	**SP**	**SP**	**SP**	**RP**	**RP/ SP**	**RP**	**SP**	**SP**	**RP**	**RP**	**N/A**	**SP**	**SP**	**N/A**

7.3 Appendix C

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

Proof that KL is a monotonic transformation of likelihood:

$$
\begin{aligned}
&KL\left(s_j \| P_j\right) = \Sigma_{j=1}^{J} \ln\left(\frac{s_j}{P_j}\right) s_j \\
&= \Sigma_{j=1}^{J}\left(s_j \ln\left(s_j\right) - s_j \ln\left(P_j\right)\right) \\
&= \Sigma_{j=1}^{J} s_j \ln\left(s_j\right) - \Sigma_{j=1}^{J} \frac{n_j}{J} \ln\left(P_j\right) \\
&= \Sigma_{j=1}^{J} s_j \ln\left(s_j\right) - \frac{1}{J} \ln\left(\Pi_{j=1}^{J} P_j^{n_j}\right) \\
&= k - \frac{\ln(L)}{J}
\end{aligned}
\tag{7}
$$

7.4 Appendix D

Estimated coefficients and evaluation measures for selected models with discussion

For each of the covariates listed in Table D.1 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- One 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 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 measures, but the simple hp/wt measures 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.).

<u>Acceleration</u>- All of the signs are as expected for all of the acceleration forms.

<u>Size</u>- A size covariate is included in all but one of the best models. The positive sign indicates that consumers prefer larger cars when vehicle class is controlled for.

<u>Style</u>- 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.

<u>Luxury (A/C dummy) and transmission (standard auto dummy)</u>- 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 measures' 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 best models, the coefficients are always positive, except for sports cars which presents with mixed signs across models.

Table 18 — Coefficient estimates for selected logit models

		Literature informed models				Best estimation set fit	Best prediction set fit			
	Covariate units	**BM-A like**	**BM-B/C like**	**BLP-like**	**Whitefoot-like**	**AIC/BIC/KL/EAL**	**AIC/BIC/KL/EAL**	**CDFET 25%**	**CDFET 50%**	**CDFET 75%**
Cost to consumer										
PRICE										
price	10k $	-0.39			-0.44			-0.42	-0.38	
price+$/50000mi	10k $ +10k $/50k mi		-0.28			-0.46	-0.40			-0.19
ln(price)	ln(10k $)			-1.56						
Operating										
$/mi	$/10 mi									
mi/$	10 mi/$			5.74		-7.02				
mpg	10 mi/gal	0.02								
gal/mi	gal/10 mi				-41.22		50.67	-14.97	-19.41	
Performance										
Acceleration										
hp/wt	hp/10 lbs			1.36	1.37			0.99		0.81
wt/hp	10 lbs/hp	-0.46	-0.40						-0.32	
f(hp/wt)	exp(hp/wt)									
hp	hp					0.00				
Size										
Physical										
length	ft									
width	ft						0.72			
length-width	ft									
length*width	100 sq-ft			4.85	5.40	7.09		4.79	4.71	
Intangibles										
Style										
(length*width)/height	100-ft	21.69	21.97			-17.18	9.52			17.07
Luxury										
a/c std	1			0.09		0.00				-0.06
Transmission										
automatic std.	1					-0.03		-0.10	-0.07	

	Covariate units	Literature informed models: BM-A like	BM-B/C like	BLP-like	Whitefoot-like	Best estimation set fit: AIC/BIC/KL/EAL	Best prediction set fit: AIC/BIC/KL/EAL	CDFET 25%	CDFET 50%	CDFET 75%
Intangibles										
Manufacturer										
geographical (US, Asia, Europe)										
Europe	1									
Asia	1									
brand dummies included	1					x	x	x	x	x
Class dummies										
compact	1	-0.22	-0.53	1.01	0.99	2.48	1.32	1.72	1.66	0.15
fullsize	1	-0.87	-1.20	0.73	0.40	1.81	0.53	0.92	0.84	-0.58
luxury sedan	1	-1.09	-1.55	0.59	0.36	2.42	1.20	1.55	1.47	-0.18
luxury SUV	1	0.49	0.31	1.06	1.14	1.43	0.83	1.22	1.15	0.52
midsize	1	-0.15	-0.48	1.40	1.17	2.64	1.37	1.80	1.71	0.23
minivan	1	0.00	-0.28	0.70	0.47	1.09	0.49	0.84	0.77	0.06
pickup	1	0.59	0.45	1.18	1.10	1.61	1.25	1.26	1.24	0.72
sports	1	-1.69	-2.08	0.18	0.03	1.58	-0.07	0.45	0.44	-1.57
SUV	1	0.30	0.03	1.22	1.12	1.83	1.13	1.56	1.48	0.51
Total number of covariates		**13**	**12**	**14**	**13**	**52**	**49**	**50**	**50**	**49**

Table 19 — Measures for selected logit models

Measure	Min. value over all models	Max. value over all models	Literature informed models: BM-A like	BM-B/C like	BLP-like	Whitefoot-like	Best estimation set fit: AIC/BIC/KL/EAL	Best prediction set fit: AIC/BIC/KL/EAL	CDFET 25%	CDFET 50%	CDFET 75%
ESTIMATION SET											
AIC (10e7)	-5.2511	-4.9934	-5.1930	-5.1983	-5.1733	-5.1559	$\boxed{-4.9934}$	-5.0301	-5.0065	-5.0046	-5.0449
BIC (10e7)	-5.2511	-4.9934	-5.1930	-5.1983	-5.1733	-5.1559	$\boxed{-4.9934}$	-5.0302	-5.0065	-5.0046	-5.0449
KL	0.1978	0.4587	0.4000	0.4053	0.3800	0.3624	$\boxed{0.1978}$	0.2350	0.2111	0.2092	0.2499
EAL	0.0049	0.0064	0.0052	0.0052	0.0053	0.0054	$\boxed{0.0064}$	0.0061	0.0063	0.0063	0.0060
ASE	0.0021	0.0031	0.0029	0.0029	0.0028	0.0028	0.0021	0.0023	0.0021	0.0021	0.0023
PREDICTION SET											
KL	0.2416	0.4622	0.4204	0.4174	0.4144	0.4026	0.2661	$\boxed{0.2416}$	0.2603	0.2602	0.2655
EAL	0.0048	0.0060	0.0050	0.0050	0.0050	0.0051	0.0058	$\boxed{0.0060}$	0.0059	0.0059	0.0058
ASE	0.0024	0.0032	0.0030	0.0030	0.0031	0.0030	0.0025	0.0024	0.0025	0.0025	0.0025
CDF cutoff											
Within 25%	0.0004	0.0010	0.0007	0.0007	0.0008	0.0007	0.0004	0.0004	$\boxed{0.0004}$	0.0004	0.0004
Within 50%	0.0011	0.0023	0.0020	0.0020	0.0019	0.0017	0.0013	0.0013	0.0012	$\boxed{0.0011}$	0.0012
Within 75%	0.0027	0.0042	0.0038	0.0038	0.0037	0.0037	0.0032	0.0030	0.0030	0.0029	$\boxed{0.0027}$
Within 100%	0.0247	0.0377	0.0363	0.0362	0.0346	0.0345	0.0313	0.0285	0.0311	0.0309	0.0317

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 measures 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=0.01 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 measures indicate the measure value when it was the selection criterion for the model

7.5 Appendix E

Selected results from mixed and nested logit model estimation

The evaluation measures 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 measures 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 measures calculated from the prediction data. The "estimation set" measures are the measures evaluated for each model on the 2004-2006 estimation data and the "prediction set" measures are the measures evaluated for each model on the 2007 prediction data.

Table 20 — Logit, mixed logit, and nested logit measure comparison

	Logit		Mixed logit		Nested logit	
Measure	**Est.**	**Pred.**	**Est.**	**Pred.**	**Est.**	**Pred.**
ESTIMATION SET						
AIC (10e7)	-4.9934	-5.0301	-4.9601	-4.9677	-4.9925	-5.0360
BIC (10e7)	-4.9934	-5.0302	-4.9601	-4.9677	-4.9925	-5.0361
KL	0.1978	0.2350	0.1641	0.1718	0.1969	0.2410
EAL	0.0064	0.0061	0.0066	0.0065	0.0064	0.0061
ASE	0.0021	0.0023	0.0018	0.0018	0.0021	0.0023
PREDICTION SET						
KL	0.2661	0.2416	0.2426	0.2219	0.3057	0.2644
EAL	0.0058	0.0060	0.0060	0.0061	0.0056	0.0058
ASE	0.0025	0.0024	0.0023	0.0022	0.0027	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. 42 as opposed to representing them as class dummies in the utility function:

$$P_j = \frac{\exp\left(v_j / \lambda_k\right)\left(\sum_{i \in N_k} \exp\left(v_i / \lambda_k\right)\right)^{\lambda_k - 1}}{\sum_{l=1}^{K}\left(\sum_{m \in N_l} \exp\left(v_m / \lambda_l\right)\right)^{l}} \tag{42}$$

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 21 — 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	-0.61	0.18	-0.59
gpm	50.67	90.83	0.01*	87.68
width	0.72	0.95	0.19	0.99
(len.*wid.)/hei.	9.52	9.62	0.00*	11.37
Class dummies				
compact	1.32	1.58	0.21	1.55
fullsize	0.53	0.86	0.00*	1.30
luxury sedan	1.20	-110.45	70.51	1.56
luxury SUV	0.83	-2.68	3.46	1.39
midsize	1.37	-1.87	4.83	1.63
minivan	0.49	-67.25	52.15	1.22
pickup	1.25	-20.75	22.46	1.42
sports	-0.07	-3.87	5.17	1.13
SUV	1.13	1.00	1.16	1.39
van				0.78

	Logit	Mixed logit		Nested logit
	Mean	Mean	St. dev.	Nominal mean
Brand dummies				
Audi	-1.30	-1.49	0.16	-1.96
BMW	0.41	0.61	0.00*	0.47
Buick	0.15	-0.15	0.92	0.03
Cadillac	-0.06	-0.10	0.00*	-0.21
Chevrolet	0.84	-2.65	6.25	1.04
Chrysler	0.48	-0.86	2.06	0.51
Dodge	0.71	-1.48	2.47	0.89
Ford	1.23	1.26	0.04	1.57
GMC	0.25	-0.06	0.61	0.29
Honda	1.07	-11.91	10.84	1.40
Hummer	-0.31	-0.22	0.00*	-0.56
Hyundai	0.00**	0.26	0.00*	-0.21
Infiniti	-0.41	-20.90	9.89	-0.75
Isuzu	-1.97	-2.04	0.00*	-2.88
Jaguar	-1.70	-8.45	4.18	-2.49
Jeep	0.76	-10.28	8.78	0.81
Kia	-0.44	-0.24	0.00	-0.84
Land Rover	-0.75	-0.72	0.13	-1.21
Lexus	0.07	0.15	0.00*	0.09
Lincoln	-0.55	-0.67	0.00†	-0.84
Mazda	-0.29	-0.10	0.49	-0.58
Mercedes	0.60	0.59	0.18	0.75
Mercury	-0.46	-0.69	0.77	-0.82
Mitsubishi	-0.85	-0.85	0.45	-1.43
Nissan	0.38	-28.29	26.79	0.37
Oldsmobile	-1.00	-0.67	0.00*	-1.69
Pontiac	0.06	-0.37	1.82	-0.10
Porsche	0.30	0.42	0.14	0.53
Saab	-1.21	-1.18	0.00*	-1.91
Saturn	0.10	-20.10	14.82	0.01***
Scion	-0.29	-1.52	2.33	-0.69
Subaru	-0.17	-5.87	5.11	-0.42
Suzuki	-1.43	-1.17	0.00*	-2.40
Toyota	0.94	-123.46	113.43	1.20
Volkswagen	-0.17	-3.77	3.03	-0.53
Volvo	-0.84	-2.46	2.72	-1.27

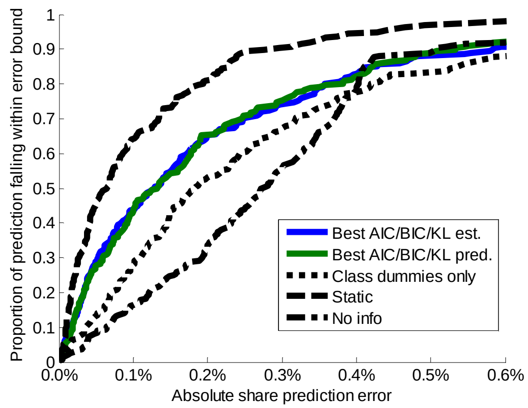
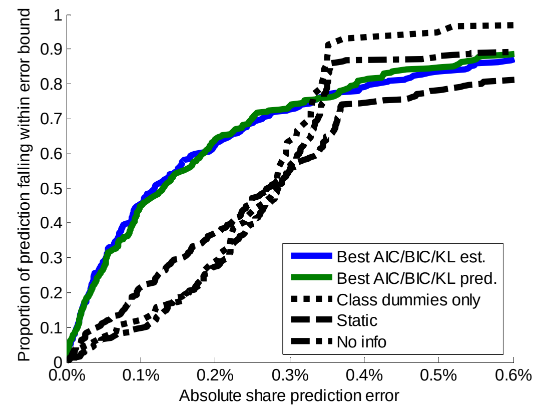
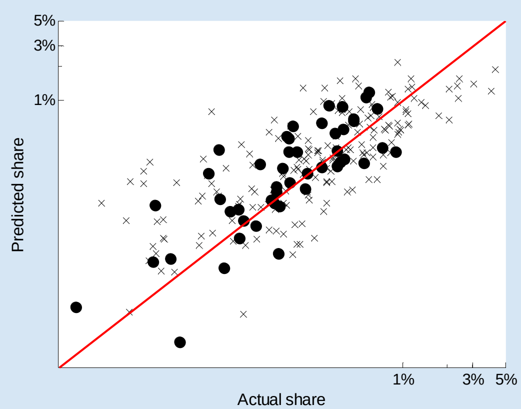
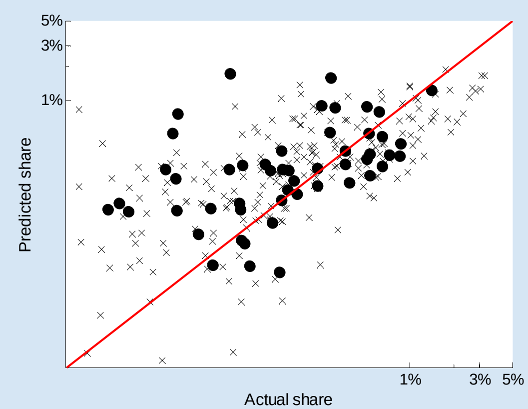
Note: All coefficients are significant at the α=0.01 level unless otherwise indicated
* Not significant at any level
** Significant at the α=0.10 level
*** Significant at the α=0.05 level

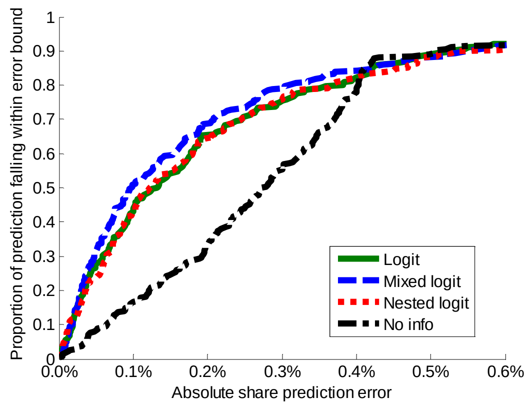
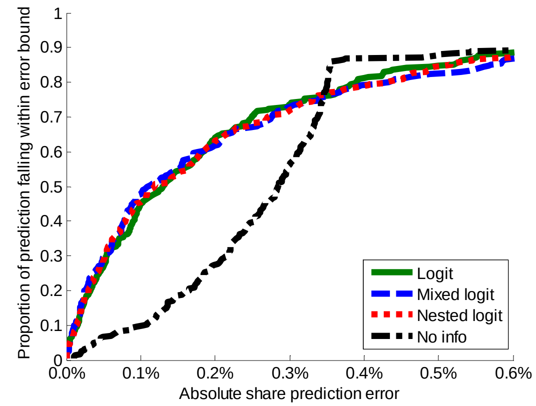
7.6 Appendix F

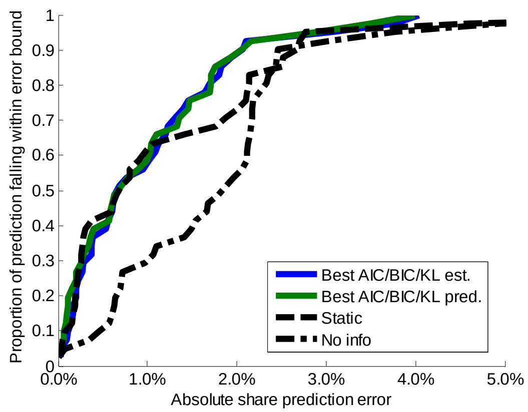
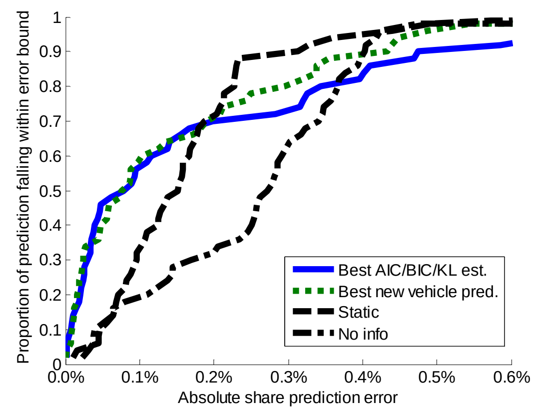
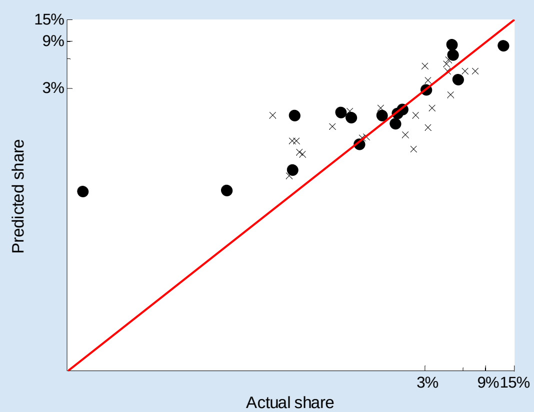
CDFET and actual versus predicted share plots for selected scenarios 1, 4-6 from Table 3.

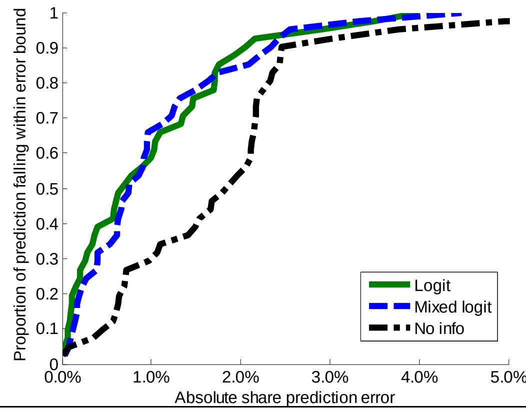
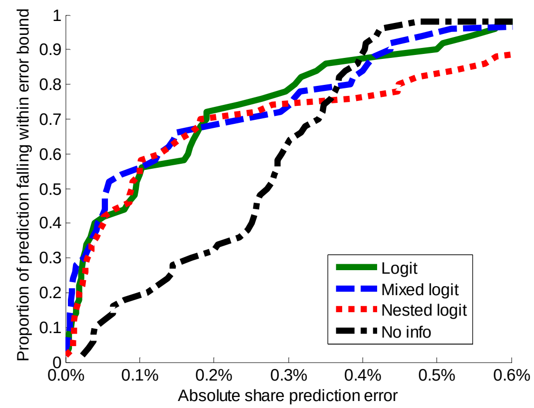
In the actual versus predicted scatter plots the solid line represents the space where the predicted share would be identical to the actual share, and the vehicles that appear in the prediction set but not the estimation set are represented by circles rather than crosses. The scatter plots for scenarios 1 and 6 are identical. Scenarios 2 and 3 are omitted because of their similarity to scenario 1. All scatter plots are on a log-log scale.

Table 22 — CDFET by scenario

Scenario	1	4
Fit data	2004-2006	2004-2006
Prediction data	2007	2010
Market	Full market	Full market
CDFET of best KL/AIC/BIC 2004-2006 estimative and predictive logit		
Scatter plot of best KL/AIC/BIC 2004-2006 predictive logit		

<table>
<tr><th>Scenario</th><th>1</th><th>4</th></tr>
<tr><td>Fit data</td><td>2004-2006</td><td>2004-2006</td></tr>
<tr><td>Prediction data</td><td>2007</td><td>2010</td></tr>
<tr><td>Market</td><td>Full market</td><td>Full market</td></tr>
<tr><td>CDFET of mixed and nested logits fit to best KL/AIC/BIC estimative logit</td><td>
</td><td>
</td></tr>
</table>

<table>
<tr><th>Scenario</th><th>5</th><th>6</th></tr>
<tr><td>Fit data
Prediction data
Market</td><td>2004-2006
2007
Luxury sedan</td><td>2004-2006
2007
New designs</td></tr>
<tr><td>CDFET of best fit and best prediction logit specification using L/AIC/BIC</td><td>
</td><td>
</td></tr>
<tr><td>Scatter plot of best fit and best prediction logit specification using L/AIC/BIC</td><td>
</td><td>See scenario 1</td></tr>
</table>

<table>
<tr><th>Scenario</th><th>5</th><th>6</th></tr>
<tr><td>Fit data</td><td>2004-2006</td><td>2004-2006</td></tr>
<tr><td>Prediction data</td><td>2007</td><td>2007</td></tr>
<tr><td>Market</td><td>Luxury sedan</td><td>New designs</td></tr>
<tr><td>CDFET of mixed logit and nested logit using best fit logit covariates</td><td>
</td><td>
</td></tr>
</table>

7.7 Appendix G

Summary of best vehicle class model specifications and measure comparison

In Table 23 an “x” indicates that the covariate was included in the best estimative model fit to 2004-2006 class level data only. Table 23 compares the average likelihood (AL) for models the best estimative models fit to the column heading data and used to predict the row heading data. It shows that models fit to specific classes in 2004-2006 sales data predict 2007 sales in that class better than models fit to the full market in 2004-2006 (exception: midsize and sport). Further, models fit to specific classes in 2004-2006 data are frequently better at predicting 2007 sales in that class than models fit directly to the full market in 2007 (exceptions for luxury SUV, midsize, and sports). This implies that the improvement in fit from modeling a subset of the market can be a larger factor influencing prediction accuracy than the differences in sales patterns between 2004-2006 and 2007.

Table 23 — Class level model specification for the class's best estimative model

	com-pact	full-size	luxury sedan	luxury SUV	mid-size	mini-van	pick-up	sports	SUV	van
Cost to consumer										
Price										
price		x					x			
price+$/50000mi						x			x	
ln(price)	x		x	x	x			x		x
Operating										
$/mi		x			x					
mi/$	x			x				x		
mpg							x		x	x
gal/mi			x			x				
Performance										
Acceleration										
hp/wt										
wt/hp	x	x	x	x	x					
f(hp/wt)										
hp						x	x	x	x	x
Size										
Physical										
length			x			x			x	
width		x		x	x		x			
length-width	x									x
length*width								x		
Intangibles										
Style										
(length*width)/height	x	x	x	x	x	x	x	x	x	
Luxury										
a/c std	x	x	x	x	x	x	x	x	x	x
Transmission										
automatic std.	x		x		x		x	x	x	
Manufacturer										
geographical (US, Asia, Europe)										
brand dummies included	x	x	x	x	x	x	x	x	x	x

Table 24 — Average Likelihood (AL) vehicle class model comparison

	Estimation data											
			2004-2006 data, class only									
2007 Prediction data:	2004-2006 full market	2007 full market	com-pact	full-size	luxury sedan	luxury SUV	mid-size	mini-van	pick-up	sports	SUV	van
full market	0.0058	0.0062										
compact	0.0366	0.0373	0.0378									
fullsize	0.0546	0.0607		0.0639								
luxury sedan	0.0317	0.0323			0.0336							
luxury SUV	0.0688	0.0721				0.0690						
midsize	0.0945	0.1001					0.0913					
minivan	0.0998	0.1106						0.1244				
pickup	0.0713	0.0766							0.0807			
sports	0.0657	0.0692								0.0519		
SUV	0.0181	0.0185									0.0185	
van	0.5935	0.5990										0.6204

7.8 Appendix H

RAL for model scenarios tested on alternate data years

To verify that our results are not specific to the 2004-2006 timeframe, we estimate 360 logit models with alternate utility function specifications on 1971-1973 and 1981-1983 data. We predict 1 year and 4 year forward market shares for each of the estimation data sets. We use data from Berry *et al.* [58] that is posted online as a companion to Knittel and Metaxoglou [33]. Due to limited covariate availability we are not able to estimate all 9,000 utility functions that were applied to the 2004-2006 data. The conclusions from the analysis of 2004-2006 data generally hold for these two historical data sets. Specifically:

- Our accuracy measures are concordant: model specifications that perform well on one measure tend to also perform well on other measures for both fit and prediction
- Even the best DCMs exhibit substantial prediction error, stemming largely from limited model fit due to unobserved covariates (more so than the 2004-2006 data since the 1971-1987 data does not include class or brand dummies)
- The static model (share in the forecast year = share in the last available year) outperforms all 360 attribute-based models when predicting the full market 1-year-forward
- Attribute-based models can predict better for 4-year-forward forecasts or new vehicle designs
- Share predictions are sensitive to the presence of utility covariates but less sensitive to covariate form (e.g. miles per gallons versus gallons per mile)
- Mixed logit specifications do not produce more accurate forecasts

- The 1971-1973 models with best predictions do not necessarily have expected coefficient signs, however, the best predictive models in 1981-1983 do have the expected coefficient signs

Table 25 compares the RAL for tested scenarios. Note that the available data set does not contain sales data, rather only share is provided. Consequently, sales were imputed from the shares and an assumed market size of 100,000,000. This is an approximation of the number of US households for each of the years, which is the market defined by Berry *et al.* [58]

Table 25 — Relative Average Likelihood (RAL) calculated on the prediction data set for select model specifications and data sets

Scenario	**1**	**2**	**3**	**4**	**5**	**6**
Estimation data	1971-1973	1971-1973	1971-1973	1981-1983	1981-1983	1981-1983
Prediction data	1974	1977	1974	1984	1987	1984
Market	Full market	Full market	New designs[1]	Full market	Full market	New designs[1]
AL of ideal model (predicted shares=actual shares)	*0.0243*	*0.0179*	*0.7353*	*0.0133*	*0.0119*	*0.1110*
RAL of no info model	57.2%	58.9%	89.2%	66.6%	58.8%	82.5%
RAL of static model	**83.1%**	62.9%	93.0%	**77.4%**	62.9%	82.4%
RAL of best logit model for likelihood of prediction data	70.0%	70.0%	98.6%	76.9%	**65.1%**	**89.0%**
RAL of best fit logit model for L/AIC/BIC of estimation data	69.0%	**74.3%**	98.6%	76.4%	64.3%	88.9%
RAL of mixed logit with best logit estimation fit covariates	68.8%	73.0%	**98.9%**	76.5%	62.5%	**89.0%**

[1] Full market used for estimation, measures assessed for prediction of new vehicles only

7.9 Appendix I

Explicit formulation of the GMM-IV optimization problem

The optimization problem for GMM-IV estimation as stated in Eq. 19 is:

$$\begin{aligned}
&\underset{\beta,\xi}{\text{minimize}}: \left(\mathbf{Z}'\xi\right)' \mathbf{W}\left(\mathbf{Z}'\xi\right)/T \\
&\text{subject to: } \ln\left(P_{jt}\left(\beta,\xi_t\right)\right) = \ln\left(s_{jt}\right), \quad \forall j \in J_t^-, t \\
&\qquad\qquad \sum_{k\in J_t} \xi_{kt} = 0, \forall t
\end{aligned}$$

and we specify $\mathbf{W}=(\mathbf{Z}'\mathbf{Z})^{-1}$ as the weighting matrix. For execution, we transform the optimization problem so that the weighting matrix can be incorporated by singular value decomposition, rather than directly using Matlab's "inverse" function since it is less numerically stable, and we rewrite the objective as a simple inner product plus a linear constraint which is more computationally efficient for the KNITRO solver. We obtain matrices **U**, **D**, and **V** from the singular value decomposition of **z** such that:

$$\mathbf{UDV}' = \mathbf{Z} \tag{43}$$

U is a ($V\times V$) orthogonal matrix, **V** is a ($K\times K$) orthogonal matrix, and **D** is ($V\times K$) matrix composed of a stacked diagonal ($K\times K$) matrix with positive entries and a ((V-K) $\times K$) matrix of zeros. We call the diagonal ($K\times K$) upper matrix **D***. The product $\mathbf{Z}'\mathbf{Z}$ can be written:

$$\mathbf{Z}'\mathbf{Z} = \mathbf{VD}'\mathbf{U}'\mathbf{UDV}' = \mathbf{VD}'\mathbf{DV}' = \mathbf{V}\left(\mathbf{D}^*\right)^2\mathbf{V}' \tag{44}$$

so that $\mathbf{W}=(\mathbf{Z}'\mathbf{Z})^{-1}$ can be expressed:

$$\mathbf{W} = \left(\mathbf{Z}'\mathbf{Z}\right)^{-1} = \left(\mathbf{V}\left(\mathbf{D}^*\right)^2\mathbf{V}'\right)^{-1} = \mathbf{V}\left(\mathbf{D}^2\right)^{-2}\mathbf{V}' \tag{45}$$

We can now rewrite the objective function as a simple quadratic equation with the addition of a linear constraint:

$$
\begin{aligned}
&\underset{\boldsymbol{\beta},\boldsymbol{\xi}}{\text{minimize}} : (\mathbf{h}'\mathbf{h}) / T \\
&\text{subject to: } \ln\left(P_{jt}(\boldsymbol{\beta},\boldsymbol{\xi}_t)\right) = \ln\left(s_{jt}\right), \quad \forall j \in J_t^-, t \\
&\qquad\qquad \sum_{k \in J_t} \xi_{kt} = 0, \forall t \\
&\qquad\qquad \mathbf{h} = \left(\mathbf{D}^*\right)^{-1} \mathbf{V}'\mathbf{Z}'\boldsymbol{\xi}
\end{aligned} \tag{46}
$$

where **h** is (L×1). This objective function of Eq. 46 is equivalent to that of Eq. 19 as can be seen by substituting in the expression for **h** in the constraints and using the singular value decomposition.

7.10 Appendix J

Synthetic data estimation results

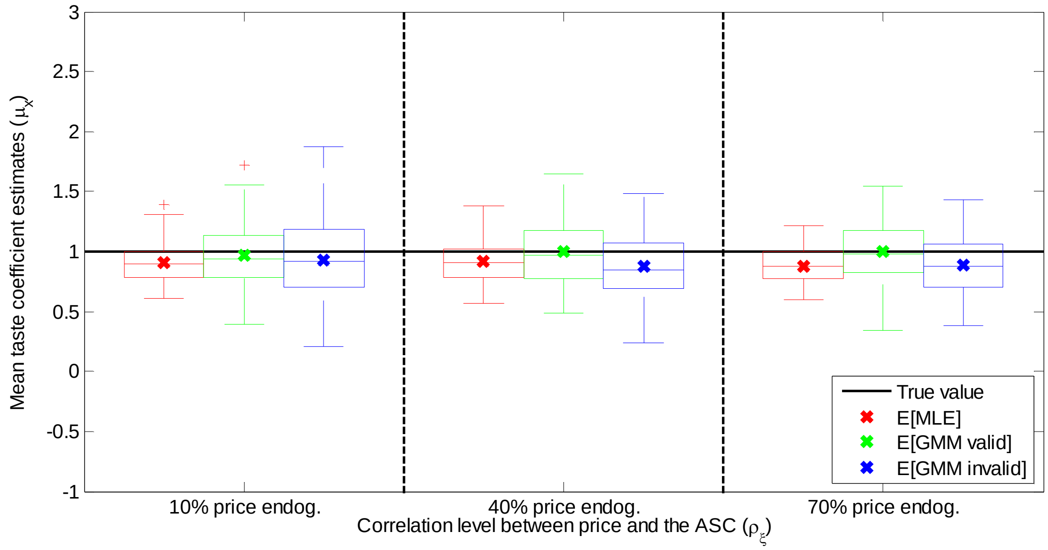


Figure 5 — Estimates of population mean taste parameter μ_x

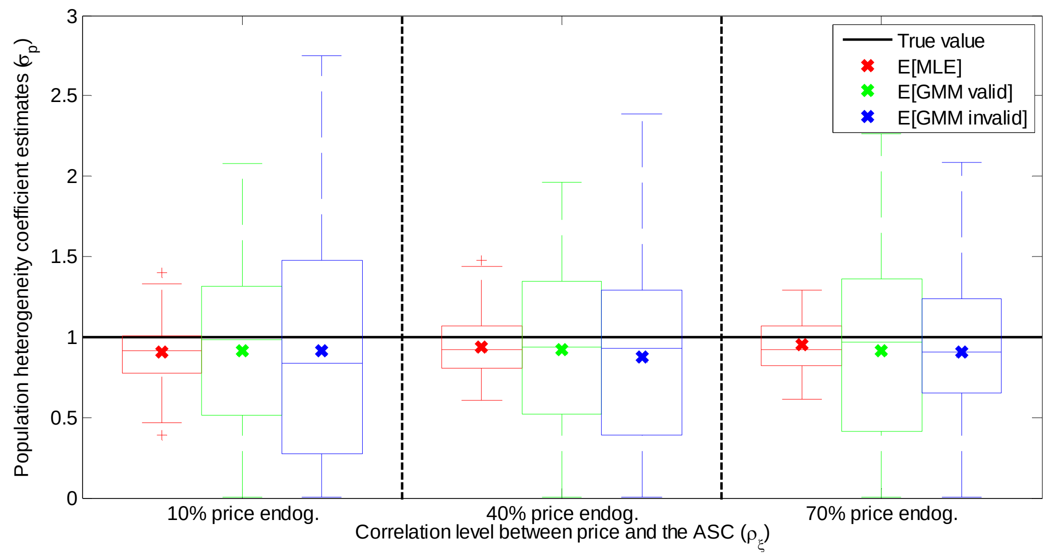


Figure 6 — Estimates of population heterogeneous portion of taste parameter price σ_p

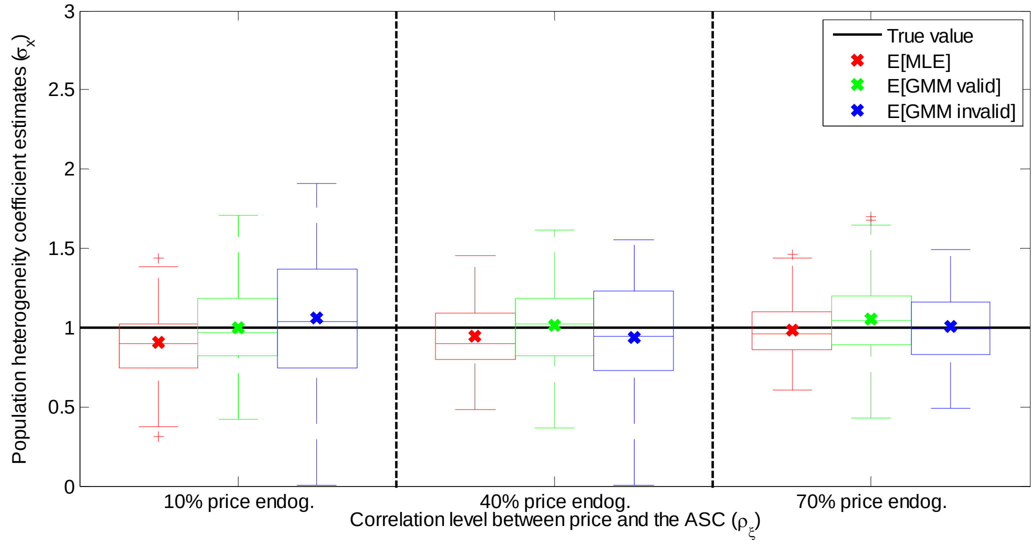


Figure 7 — Estimates of population heterogeneous portion of taste parameter technology σ_t

7.11 Appendix K

Synthetic data prediction results

Table 26 — INVALID case results shown include the mean and std. dev. of the RAL of expected share across 125 data sets for MLE-C and GMM-IV models and the number of data sets for which the MLE-C or GMM-IV model had a greater respective RAL (# superior)

	1-year-forward			5-year-forward		
	None (0)	All (1)	Neigh. (2)	None (0)	All (1)	Neigh. (2)
10% price-ASC correlation (ρ_ξ)						
MLE-C	72%	85%	67%	85%	74%	68%
(std. dev.)	*(5%)*	*(6%)*	*(8%)*	*(8%)*	*(12%)*	*(8%)*
# superior	100	100	119	123	121	116
GMM-IV	61%	53%	54%	63%	52%	55%
(std. dev.)	*(5%)*	*(4%)*	*(13%)*	*(18%)*	*(20%)*	*(13%)*
# superior	0	0	6	2	4	9
Static	72%		68%			42%
No info	42%		38%			39%
40% price-ASC correlation (ρ_ξ)						
MLE-C	72%	86%	78%	71%	72%	56%
(std. dev.)	*(6%)*	*(6%)*	*(8%)*	*(6%)*	*(7%)*	*(8%)*
# superior	120	121	119	114	117	121
GMM-IV	63%	67%	56%	62%	61%	38%
(std. dev.)	*(10%)*	*(17%)*	*(19%)*	*(9%)*	*(9%)*	*(15%)*
# superior	5	4	6	11	8	4
Static	72%			47%		
No info	44%			45%		
70% price-ASC correlation (ρ_ξ)						
MLE-C	81%	91%	85%	80%	82%	69%
(std. dev.)	*(5%)*	*(4%)*	*(6%)*	*(4%)*	*(4%)*	*(7%)*
# superior	116	124	113	113	117	122
GMM-IV	76%	81%	71%	76%	76%	53%
(std. dev.)	*(7%)*	*(9%)*	*(14%)*	*(6%)*	*(7%)*	*(13%)*
# superior	9	1	12	12	8	3
Static	76%			56%		
No info	52%			53%		

Table 27 — ENTRANT ONLY case results shown include the mean and std. dev. of the RAL of expected share for MLE-C and GMM-IV models and the number of data sets for which the MLE-C or GMM-IV model had a greater respective RAL (# superior)

	1-year-forward			5-year-forward		
	None (0)	All (1)	Neigh. (2)	None (0)	All (1)	Neigh. (2)
10% price-ASC correlation (ρ_ξ)						
85%	85%	74%	69%	70%	53%	85%
0.54%	*0.55%*	*1.49%*	*0.73%*	*0.78%*	*1.22%*	*0.54%*
87	103	103	92	101	113	87
83%	80%	64%	66%	65%	40%	83%
0.67%	*0.77%*	*2.32%*	*0.78%*	*0.77%*	*1.84%*	*0.67%*
38	22	22	33	24	12	38
68%			42%			68%
69%			41%			69%
40% price-ASC correlation (ρ_ξ)						
MLE-C	86%	86%	77%	72%	72%	55%
(std. dev.)	0.52%	0.55%	0.91%	0.48%	0.50%	0.83%
# superior	81	107	114	102	104	115
GMM-IV	84%	81%	67%	68%	67%	41%
(std. dev.)	0.67%	0.75%	1.87%	0.54%	0.55%	1.56%
# superior	44	18	11	23	21	10
Static	72%			48%		
No info	73%			47%		
70% price-ASC correlation (ρ_ξ)						
MLE-C	91%	91%	85%	81%	81%	69%
(std. dev.)	0.20%	0.21%	0.36%	0.20%	0.20%	0.62%
# superior	112	122	118	120	121	120
GMM-IV	87%	82%	74%	72%	72%	52%
(std. dev.)	0.37%	0.58%	1.68%	0.60%	0.56%	2.26%
# superior	13	3	7	5	4	5
Static	78%			56%		
No info	77%			56%		

7.12 Appendix L

MLE-C estimated model with brand dummies

Estimation results

Table 28 — MLE-C estimated parameters for a mixed logit model that includes brand dummies

Coefficient	Mean taste parameter μ	Mean taste parameter (std. err.)	Heterogeneity taste parameter σ	Heterogeneity taste parameter (std. err.)
Price ($10,000)	-0.9***	(0.001)	0.4141***	(0.001)
Gallons/mile (gal./100-mi.)	0.1***	(0.002)	0.0027	(0.002)
Weight/HP (10 lbs/hp)	0.1***	(0.000)	0.0001	(0.000)
Len. × wid. (100 ft^2)	1.4***	(0.005)	1.1885***	(0.006)
Acura	-0.8***	(0.003)		
Cadillac	0.4***	(0.002)		
Chevrolet	-0.4***	(0.002)		
Chrysler	-1.6***	(0.002)		
Dodge	-1.2***	(0.003)		
Ford	0.0***	(0.002)		
Honda	0.5***	(0.002)		
Hyundai	-1.9***	(0.002)		
Infiniti	0.2***	(0.003)		
Kia	-2.6***	(0.003)		
Lincoln	-0.4***	(0.003)		
Mazda	-1.5***	(0.002)		
Mercury	-1.8***	(0.002)		
Mitsubishi	-1.8***	(0.003)		
Nissan	0.1***	(0.002)		
Oldsmobile	-1.2***	(0.011)		
Pontiac	-0.4***	(0.002)		
Saab	-1.9***	(0.005)		
Saturn	-2.9***	(0.004)		
Suzuki	-3.9***	(0.009)		
Toyota	0.0	(0.002)		
Volkswagen	-1.9***	(0.002)		
Volvo	-2.1***	(0.004)		

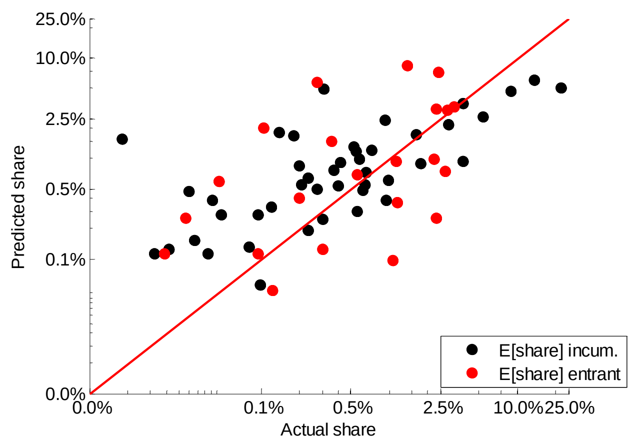
*** Coefficient is significant at the α=0.01 level

Prediction results

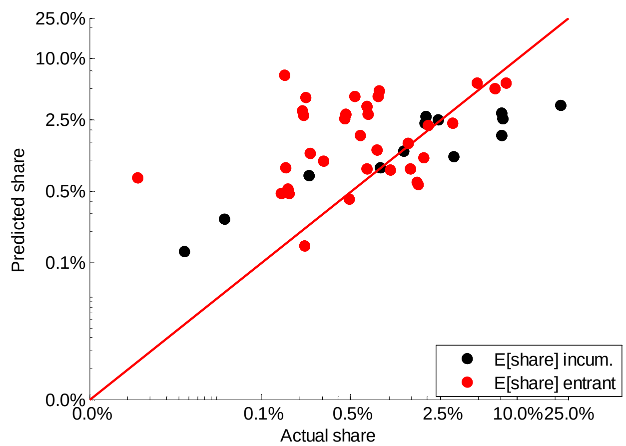
Table 29 — RAL(E[share]) comparison of ASC forecasting methods for MLE-C models including brand dummies estimated on 2002-2006 midsize sedan sales and used to predict 2007 and 2011 midsize sedan market shares

	1-year-forward (2007) forecasts					5-year-forward (2011) forecasts				
Method:	No ASC 0	All 1	Near neighbor 2	Brand 3	Model 4	No ASC 0	All 1	Near neighbor 2	Brand 3	Model 4
MLE-C	54%	66%	65%	68%	67%	51%	45%	44%	48%	51%
Static	*68%*					*39%*				
No info	*32%*					*42%*				
MLE-C	79%	78%	77%	80%	79%	60%	57%	55%	60%	64%
Static	*87%*					*54%*				
No info	*87%*					*62%*				

Note: highlighted cells indicate the most accurate model and ASC forecasting method for a given time period and means of calculating RAL



(A)



(B)

Figure 8 — Actual versus predicted shares of midsize sedans predicted by a MLE-estimated mixed logit model that includes brand dummies and excludes ASCs in the prediction year for the 2007 (A) and 2011 (B) markets

7.13 Appendix M

Estimated ASC regressions on observed vehicle attributes

We regress MLE-C and GMM-IV estimated ASCs on six sets of dependent variables in order to investigate the (possible) correlation between the ASCs and vehicle attributes. Table 30 contains the number of coefficients that are statistically significant at the α=0.05 for six linear regression models. The estimated ASC is regressed on: (1) an intercept plus vehicle physical attributes (price, gallons/mile, weight/horsepower, and (length x width)), (2) geographic dummies for the US, Europe, and Asia, (3) the covariates of models (1) and (2) excluding the US dummy for identification, (4) brand dummies (e.g. Acura, Ford, etc.), (5) a dummy variable indicating each unique vehicle, and (6) a dummy variable for unique vehicles at the aggregate model level (a Toyota Camry and Toyota Camry Solara are both assigned a single ID representing a Toyota Camry). There are 339 total observations across five estimation data set years but only 153 unique vehicles since vehicles appear in multiple years. Though we do not include brand dummies in the estimated models of Table 8, we include them in the ASC regressions as covariates since ASCs and brand are likely related.

The number of statistically significant coefficients for a given regression is nearly identical between the MLE-C and GMM-IV estimated models. For both models, regressions 4 and 5 yielded statistically significant coefficients for ~1/3 of the covariates, suggesting that the ASCs may be better described as normally distributed about brand or vehicle-specific means as opposed to randomly drawn from a mean-zero normal distribution. This supports the assumption that the ASC represents unobserved vehicle attributes as opposed to functioning only as a mathematical tool (i.e. regression residuals). However, GMM-IV estimated ASCs are statistically significantly correlated with non-price vehicle characteristics suggesting that the BLP instruments were, for our data, invalid.

Table 30 — Regression of MLE-C and GMM-IV estimated ASCs on select dependent variables

Dependent variables (regression #)	Physical attributes (1)	Geographic dummies (2)	Physical attributes + geographic dummies (3)	Brand dummies (4)	Fixed effects (5)	Aggregate fixed effects (6)
Total covariates in regression	5	3	7	24	153	66
Number of statistically significant regression coefficients at the α = 0.05 level						
MLE-C estimated ASCs	0	0	3	9	52	14
Significant coefficients	*price, gal./mi, len. x wid.*		*gal./mi, len. x wid.*	*not listed for brevity*	*not listed for brevity*	*not listed for brevity*
GMM-IV estimated ASCs	2	0	2	7	52	15
Significant coefficients	*price, gal./mi.*		*price, gal./mi.*	*not listed for brevity*	*not listed for brevity*	*not listed for brevity*

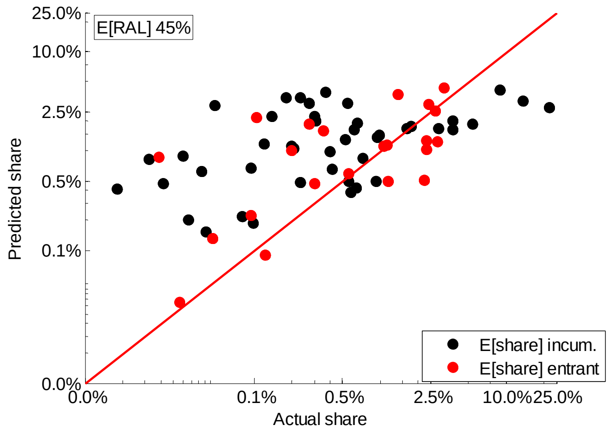
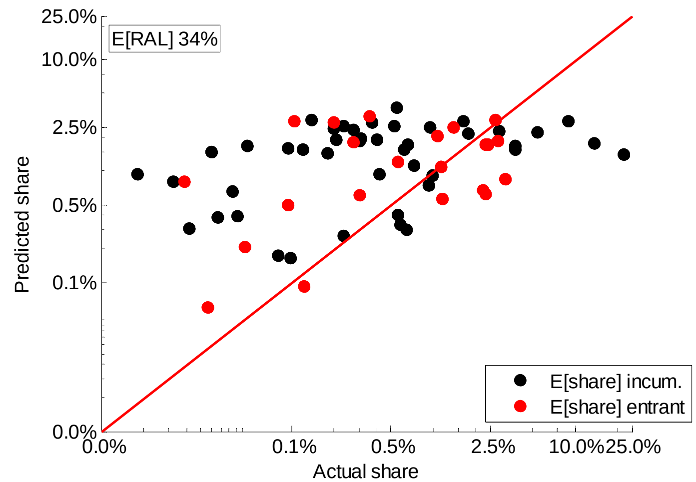
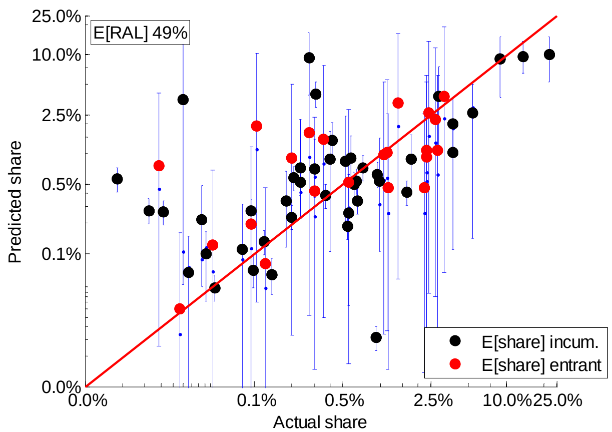
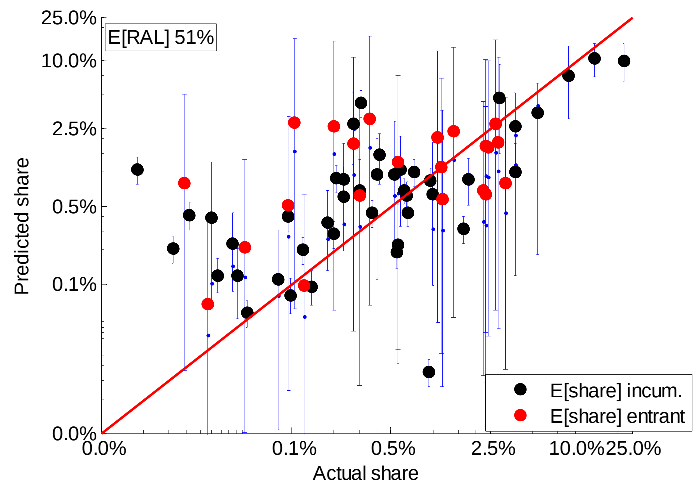
Note: A constant is excluded in regressions 2, 4, 5, and 6 for identification

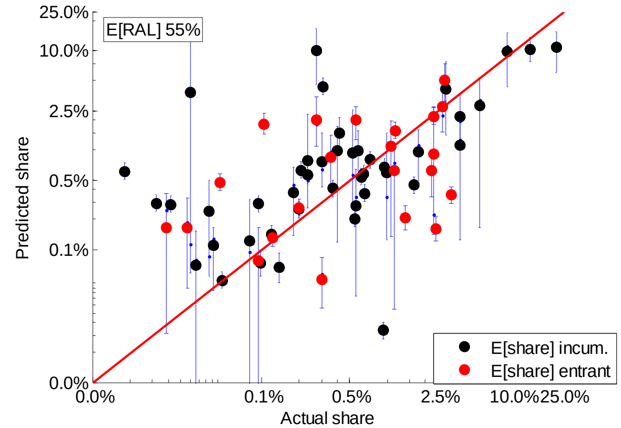
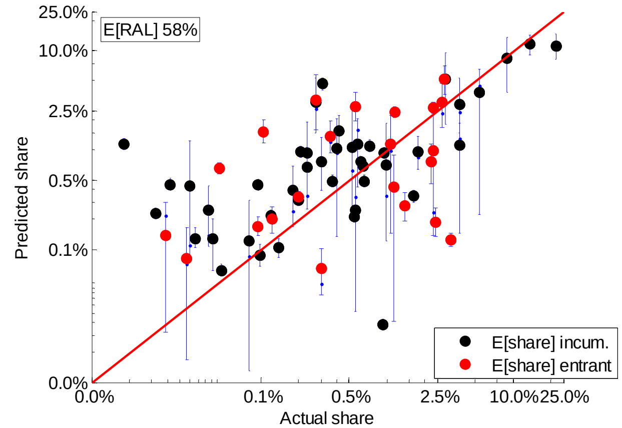
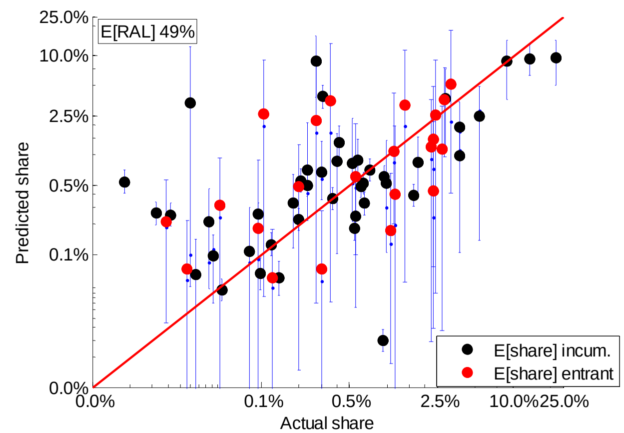
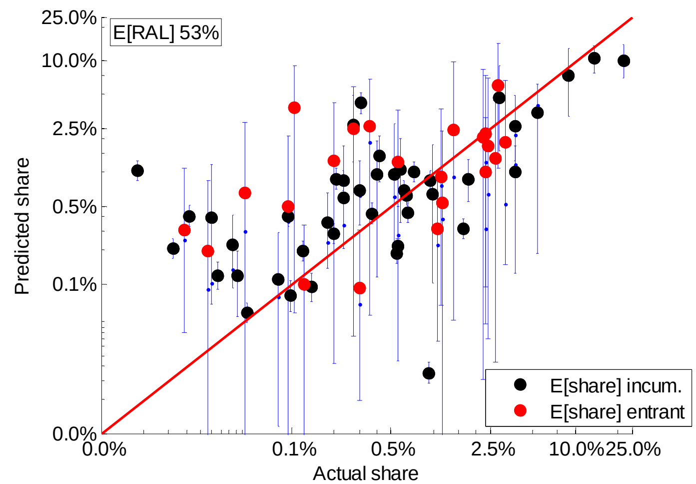
7.14 Appendix N

Case study prediction results

Table 31 and Table 32 contain plots of the actual versus predicted shares for the 1-year-forward and 5-year-forward forecasts of the MLE-C and GMM-IV models using all four ASC generation methods. The range for each of the forecasts covers the 2.5%-97.5% percentile of the simulated shares. Note that the ranges shown are independent of one another, meaning that the 2.5% percentile share shown for vehicle 1 may have occurred in a different draw of shares than the 2.5% percentile share shown for vehicle 2. Since shares of a given vehicle are related to the shares of all the other vehicles, the distributions would likely be tighter if the correlation were accounted for.

Table 31 — Actual versus predicted 2.5-97.5% interval of shares of 1-year-forward predictions for the MLE-C and GMM-IV models using each of the ASC generation methods

	MLE-C	GMM-IV
Actual v. predicted Method 0 (None)		
Actual versus predicted Method 1 (All)		

Actual versus predicted Method 2 (Nearest neighbor)		
Actual versus predicted Method 3 (Brand)		

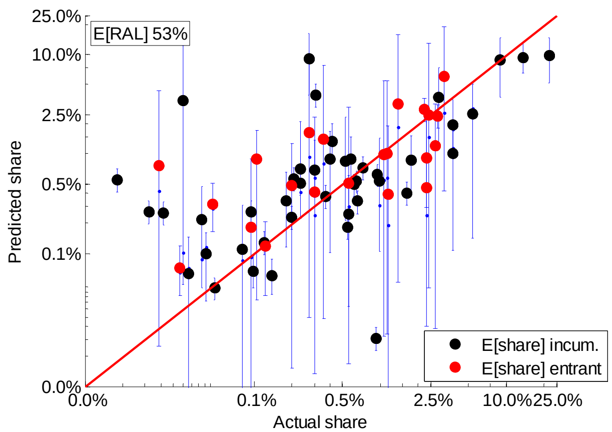
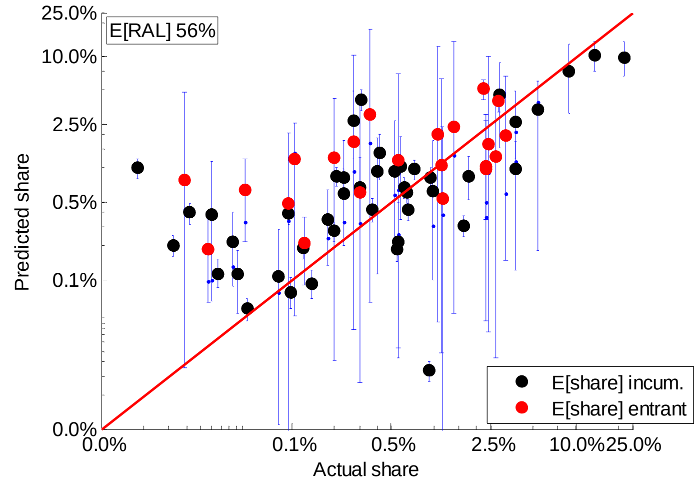
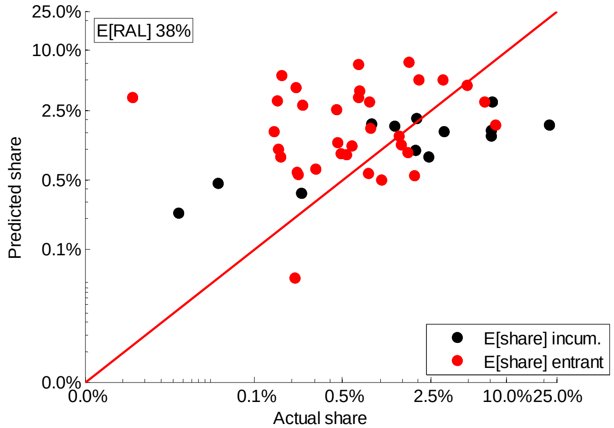
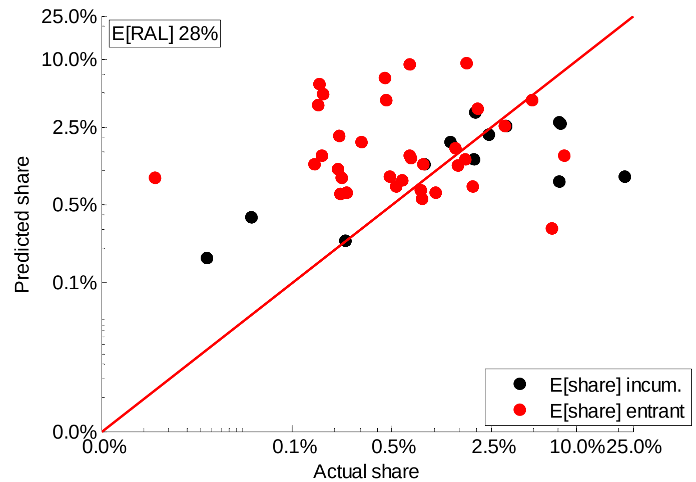
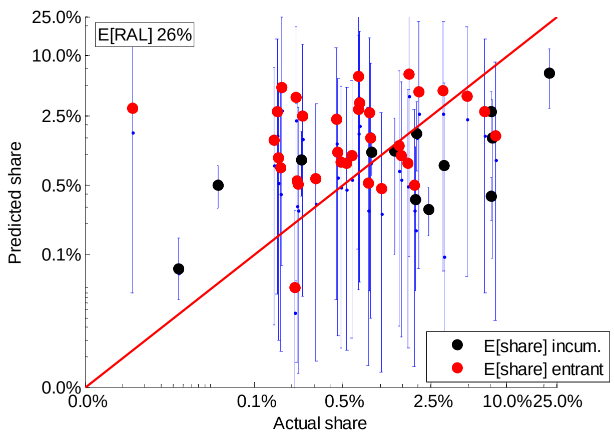
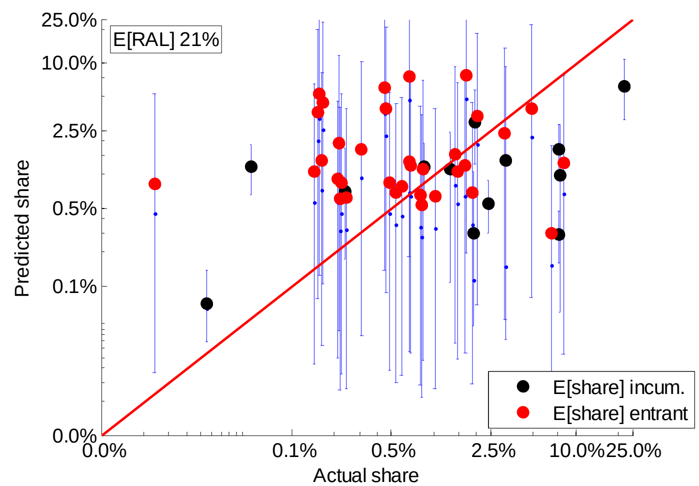
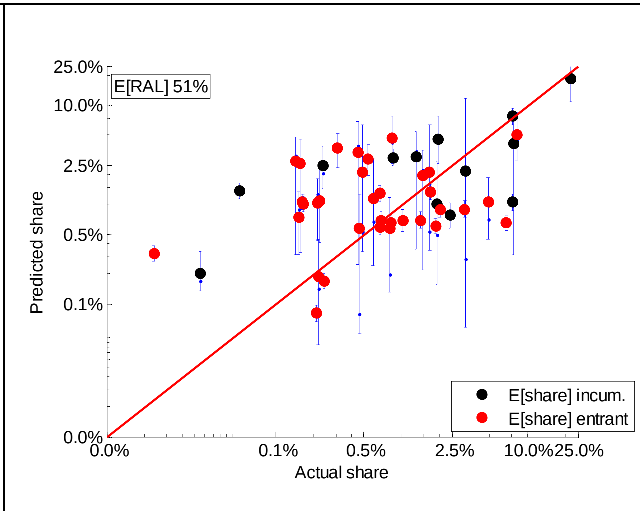
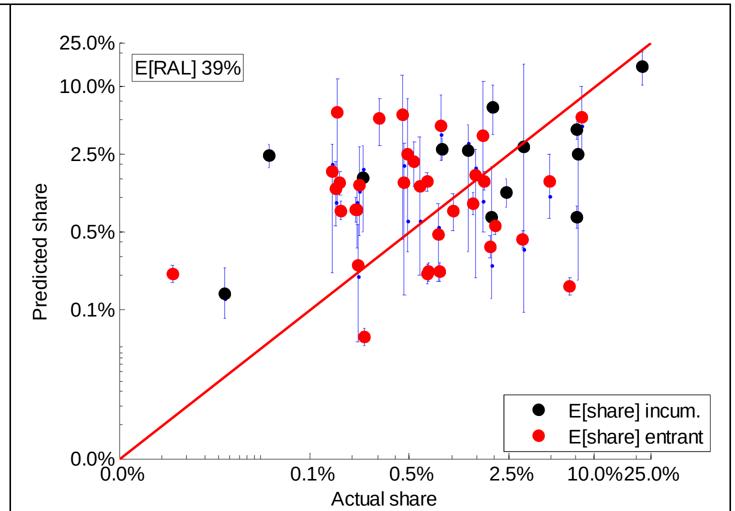
Actual versus predicted Method 4 (Make-Model)		

Table 32 — Actual versus predicted 2.5-97.5% interval of shares of 5-year-forward predictions for the MLE-C and GMM-IV models using each of the ASC generation methods

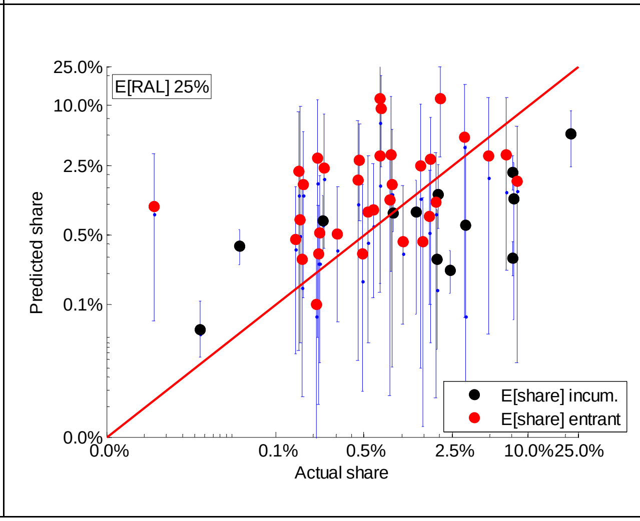
<table>
<tr><th></th><th>MLE-C</th><th>GMM-IV</th></tr>
<tr><td>Actual versus predicted Method 0 (None)</td><td>
</td><td>
</td></tr>
<tr><td>Actual versus predicted Method 1 (All)</td><td>
</td><td>
</td></tr>
</table>

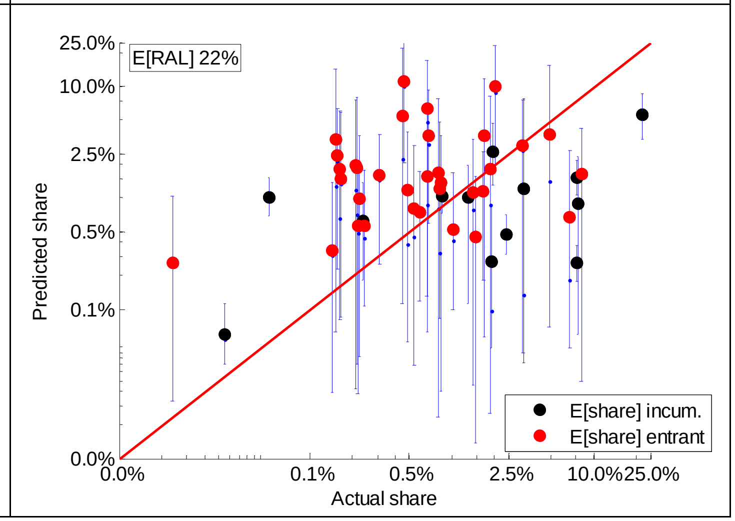
Actual versus predicted Method 2 (Nearest neighbor)

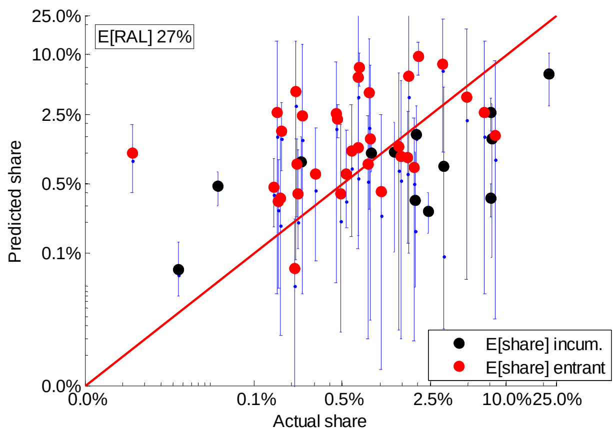
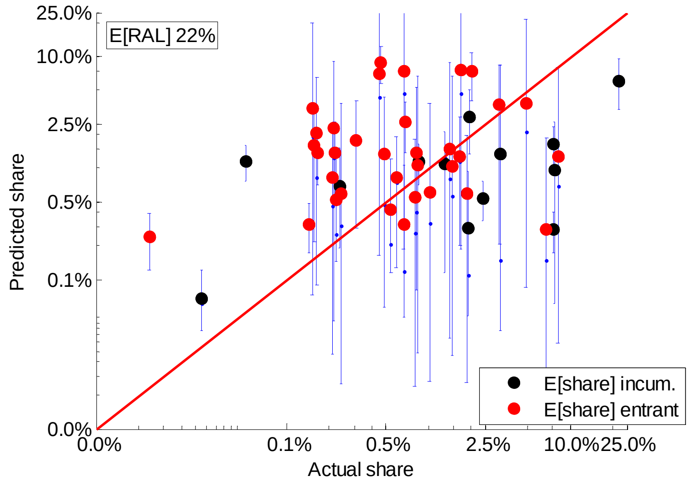




Actual versus predicted Method 3 (Brand)





Actual versus predicted Method 4 (Make-Model)		

7.15 Appendix O

Bayesian estimation specification and algorithm details

We generally follow the procedure outlined in Jiang *et al.* [109] for estimation of a Bayesian mixed logit model on aggregate data. Our steps are detailed below, and interested readers are referred to the source for more information.

Express utility as:

$$u_{ijt} = \delta_{jt} + \mathbf{x}'_{jt}\mathbf{v}_i + \varepsilon_{ijt} \tag{47}$$

where mean utility $\delta_{jt} = \mathbf{x}'_{jt}\boldsymbol{\mu} + \xi_{jt}$ and $\mathbf{v} \sim N(0, \boldsymbol{\Sigma}_\beta)$, and rewrite the share equation in terms of mean utility as:

$$P_{jt} = \int_{\mathbf{y}} \frac{\exp\left(\delta_{jt} + \mathbf{x}'_{jt}\mathbf{y}\right)}{\sum_{k \in J_t} \exp\left(\delta_{kt} + \mathbf{x}'_{kt}\mathbf{y}\right)} f_{\mathbf{v}}\left(\mathbf{y}\right) d\mathbf{y} \tag{48}$$

The expected value of predicted share P_{jt} is obtained by numerical integration. Given a value of $\boldsymbol{\Sigma}_\beta$, δ_{jt} can be obtained by nested fixed point iteration as in Berry *et al.* [58]:

$$\boldsymbol{\delta}_t^{(k+1)} = \boldsymbol{\delta}_t^{(k)} + \ln\left(\mathbf{s}_t\right) - \ln\left(\mathbf{P}_t\right), \forall t \tag{49}$$

where $\boldsymbol{\delta}_t$, $\mathbf{s}_t$, and P_t are the stacked vectors of mean utility, observed market share, and expected predicted market share, respectively, in market t. The nested fixed point iteration is terminated when $\max\left(\left|\boldsymbol{\delta}_t^{(k+1)} - \boldsymbol{\delta}_t^{(k)}\right|\right) < C$, and C is a convergence tolerance here set equal to 1e-6. Parameterize the covariance matrix $\boldsymbol{\Sigma}_\beta$:

$$\mathbf{\Sigma}_\beta = \mathbf{U}'\mathbf{U}$$
$$\mathbf{U} = \begin{bmatrix} e^{r_{11}} & r_{12} & \cdots & r_{1K} \\ 0 & e^{r_{22}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & e^{r_{K-1,K}} \\ 0 & \dots & 0 & e^{r_{KK}} \end{bmatrix} \tag{50}$$

Define priors:

$$\begin{aligned} &\boldsymbol{\mu} \sim \mathrm{MVN}\left(\boldsymbol{\mu}_0, \mathbf{V}_\mu\right) \\ &\tau^2 \sim \nu_0 s_0^2 / \chi^2_{\nu_0} \\ &r_{mm} \sim N\left(0, \sigma^2_{r_{mm}}\right) \forall m = 1,\dots,K \\ &r_{mn} \sim N\left(0, \sigma^2_{r_{off}}\right) \forall m,n = 1,\dots K, m < n \end{aligned} \tag{51}$$

7.15.1 Initialization

Assess hyperparameters for the full covariance matrix structural specification:

$$\begin{aligned} &\boldsymbol{\mu}_0 = \mathbf{0} \\ &\mathbf{V}_\mu = 10\mathbf{I} \\ &\nu_0 = K + 1 \\ &s_0^2 = 1 \\ &\sigma^2_{r_{off}} = .1 \\ &c = 50 \\ &\sigma^2_{r_{mm}} = \frac{1}{4}\log\left(\frac{1 + \sqrt{1 - 4\left(2\left(j-1\right)\sigma^4_{r_{off}} - c\right)}}{2}\right), \forall m = 1,\dots,K \end{aligned} \tag{52}$$

where **I** is the identity matrix. For the independent mixed logit specification there is no σ^2_{off} or c and:

$$\sigma^2_{r_{mm}} = 1, \forall m = 1,\dots,K \tag{53}$$

All other hyperparameter values are identical across the independent and full covariance matrix structural specifications. Set initial values **μ**=**0**, τ^2=0, and **r**=**0**.

7.15.2 Algorithm

Draw from first set of conditionals:

$$\boldsymbol{\mu},\tau^2\left|\mathbf{r},\left\{\mathbf{s}_t,\mathbf{X}_t\right\}_{t=1}^{T},\boldsymbol{\mu}_0,\mathbf{V}_{\mu},\nu_0,s_0^2\right. \tag{54}$$

by obtaining $\boldsymbol{\delta}$ for a given value of $\mathbf{r}$ and carrying out Bayesian linear regression:

$$\boldsymbol{\delta}=\mathbf{x}'\boldsymbol{\mu}+\xi,\quad \xi\sim N\left(0,\tau^2\right) \tag{55}$$

Draw $\boldsymbol{\mu}\sim N((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\delta},\tau^2(\mathbf{X}'\mathbf{X})^{-1})$ followed by

$\tau^2\sim IG\left(a+\frac{J}{2},b+\frac{1}{2}(\boldsymbol{\delta}-\mathbf{X}'\boldsymbol{\mu})'(\boldsymbol{\delta}-\mathbf{X}'\boldsymbol{\mu})\right)$ where $a=\frac{\nu_0}{2}$, and $=\frac{\nu_0 s_0^2}{2}$.

Draw from second set of conditionals:

$$\mathbf{r}\left|\boldsymbol{\mu},\tau^2,\left\{\mathbf{s}_t,\mathbf{X}_t\right\}_{t=1}^{T},\sigma_{r_{off}}^2,\sigma_{r_{mm}}^2\right. \tag{56}$$

using a Random-Walk (RW) Metropolis chain:

$$\mathbf{r}^{new}=\mathbf{r}^{old}+MVN\left(\mathbf{0},\sigma_r^2\mathbf{D}_r\right) \tag{57}$$

where σ_r^2 is a scaling constant and $\mathbf{D}_r$ is the candidate covariance matrix. We set $\mathbf{D}_r=\mathbf{I}$ and σ_r^2=0.1.

Accept draw of $\mathbf{r}^{\text{new}}$ with probability α by calculating the posterior probability:

$$\pi\left(\boldsymbol{\mu},\mathbf{r},\tau^2\left|\left\{\mathbf{s}_t,\mathbf{X}_t\right\}_{t=1}^{T}\right.\right)\propto L\left(\boldsymbol{\mu},\mathbf{r},\tau^2\left|\left\{\mathbf{s}_t,\mathbf{X}_t\right\}_{t=1}^{T}\right.\right)\times\pi\left(\boldsymbol{\mu},\mathbf{r},\tau^2\right)$$
$$=\prod_{t=1}^{T}\left(\mathbf{J}^{-1}\left(\mathbf{P}_t,\mathbf{r},\mathbf{X}_t,\right)\prod_{j\in J_t}\phi\left(\frac{h^{-1}\left(\mathbf{P}_t\left|\mathbf{X}_t,\boldsymbol{\mu},\mathbf{r},\mathbf{s}_t\right.\right)}{\tau}\right)\right)\times\left|\mathbf{V}_{\mu}\right|^{-1/2}\exp\left\{-\tfrac{1}{2}\left(\boldsymbol{\mu}-\boldsymbol{\mu}_0\right)'\mathbf{V}_{\mu}^{-1}\left(\boldsymbol{\mu}-\boldsymbol{\mu}_0\right)\right\} \tag{58}$$
$$\times\prod_{m=1}^{M}\exp\left\{-\frac{r_{mm}^2}{2\sigma_{r_{mm}}^2}\right\}\times\prod_{m=1}^{K-1}\prod_{n=m+1}^{K}\exp\left\{-\frac{r_{mn}^2}{2\sigma_{r_{off}}^2}\right\}\times\left(\tau^2\right)^{-\left(\frac{\nu_0}{2}+1\right)}\exp\left\{-\tfrac{\nu_0 s_0^2}{2\tau^2}\right\}$$

under $\mathbf{r}^{\text{new}}$ and $\mathbf{r}^{\text{old}}$ and calculating the ratio:

$$\alpha = \frac{\pi\left(\boldsymbol{\mu}, \mathbf{r}^{new}, \tau^2 \middle| \left\{\mathbf{s}_t, \mathbf{X}_t\right\}_{t=1}^{T}\right)}{\pi\left(\boldsymbol{\mu}, \mathbf{r}^{old}, \tau^2 \middle| \left\{\mathbf{s}_t, \mathbf{X}_t\right\}_{t=1}^{T}\right)} \tag{59}$$

If a randomly drawn number from the uniform interval [0,1] is less than α, then update $\mathbf{r}=\mathbf{r}^{\text{new}}$, else $\mathbf{r}=\mathbf{r}^{\text{old}}$. Note:

$$J^{-1}\left(\mathbf{P}_t, \mathbf{x}_t, \mathbf{r}\right) = \left\| \begin{bmatrix} \partial P_{1t} / \partial \xi_{1t} & \partial P_{1t} / \partial \xi_{2t} & \cdots & \partial P_{1T} / \partial \xi_{JT} \\ \vdots & \ddots & & \\ \partial P_{(J-T)T} / \partial \xi_{1T} & & \cdots & \partial P_{(J-T)T} / \partial \xi_{JT} \end{bmatrix}_{(V-T)xV} \right\| \tag{60}$$

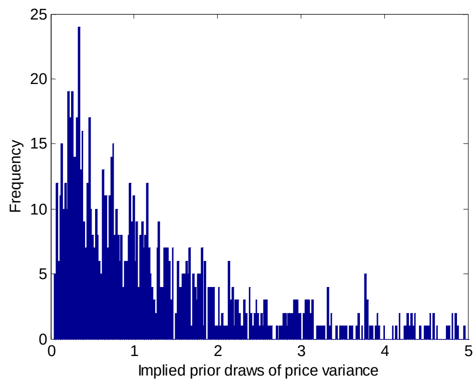
or the determinant of the $(V\text{-}T)\times V$ matrix with elements:

$$\partial P_{jt} / \partial \xi_{jt} = \begin{cases} \int_{\boldsymbol{\beta}} \left(-P_{ijt}\right)\left(P_{ikt}\right) f\left(\boldsymbol{\beta}\right) d\boldsymbol{\beta} \text{ if } k \neq j \\ \int_{\boldsymbol{\beta}} -P_{ijt}\left(1 - P_{ijt}\right) f\left(\boldsymbol{\beta}\right) d\boldsymbol{\beta} \text{ if } k = j \end{cases} \tag{61}$$

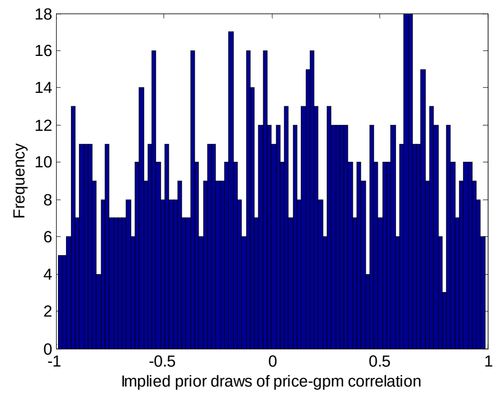
Also note $\boldsymbol{\xi}_t = h^{-1}(\mathbf{P}_t | \mathbf{X}_t, \boldsymbol{\mu}, \mathbf{r}, \mathbf{s}_t)$ is the implicit share inversion function of equation Eq. 54. We run a chain of 100,000 draws and discard 500 burn-in draws to obtain the final model parameter posterior distributions.

7.15.3 Implied Priors

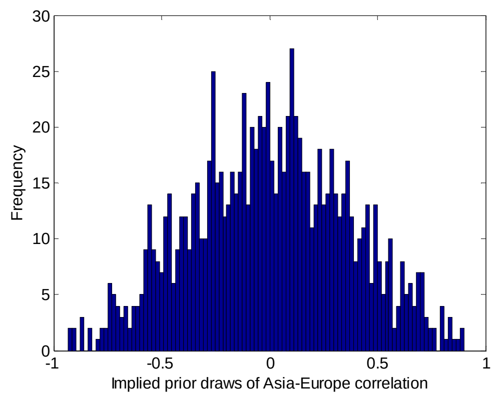
Monte Carlo simulations of the implied priors on select population taste heterogeneity parameters are shown in Figure 9.



(A)



(B)



(C)

Figure 9 – Implied prior distributions on select population taste heterogeneity parameters: price variance (A), price-gallons/mile correlation (B), and Asia-Europe dummy correlation (C)

7.16 Appendix P

Estimated covariance matrices

Table 33 contains the correlation matrix of taste parameters for the Bayes model with a full covariance matrix. The estimates shown are the posterior estimates implied by the posterior distribution of **r**.

Table 33 — Mean and (standard deviation) of CORRELATION matrix elements for the full covariance Bayesian-estimated model

	Price	Gallons/mile	Weight/HP	Length*width	Europe	Asia
Price	1.0 (0.0)	0.2 (0.5)	0.2 (0.4)	0.1 (0.4)	-0.3 (0.4)	-0.1 (0.3)
Gallons/mile		1.0 (0.0)	0.2 (0.4)	-0.1 (0.3)	-0.1 (0.4)	0.3 (0.3)
Weight/HP			1.0 (0.0)	-0.1 (0.4)	-0.1 (0.4)	0.1 (0.3)
Length*width				1.0 (0.0)	0.1 (0.4)	0.0 (0.3)
Europe					1.0 (0.0)	0.1 (0.3)
Asia						1.0 (0.0)

Note: matrix is symmetric, only upper triangular portion shown; zero values are only zero to precision shown but are non-zero

7.17 Appendix Q

Trace plots of parameter estimates for the Bayesian-estimated correlated mixed logit model

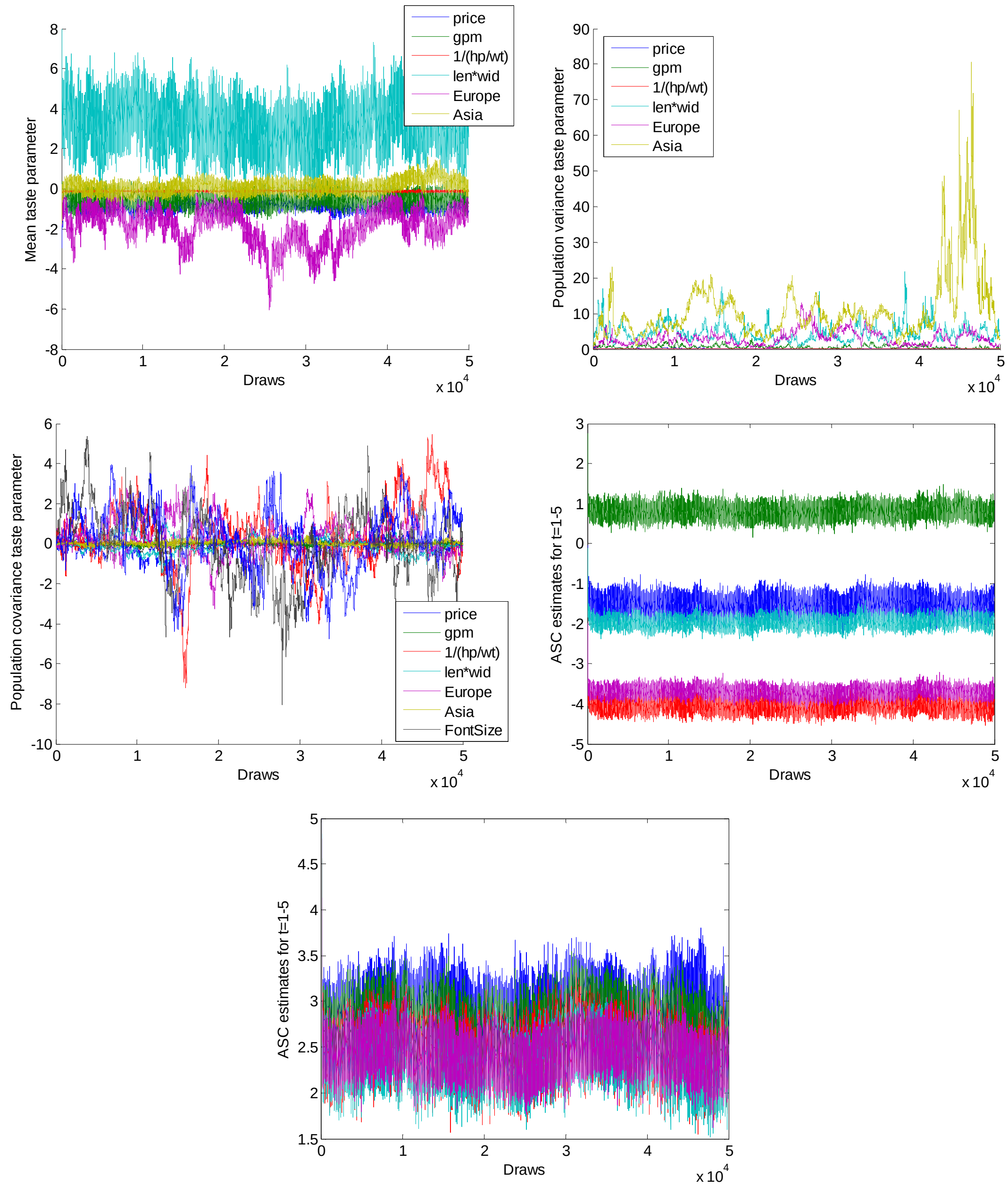
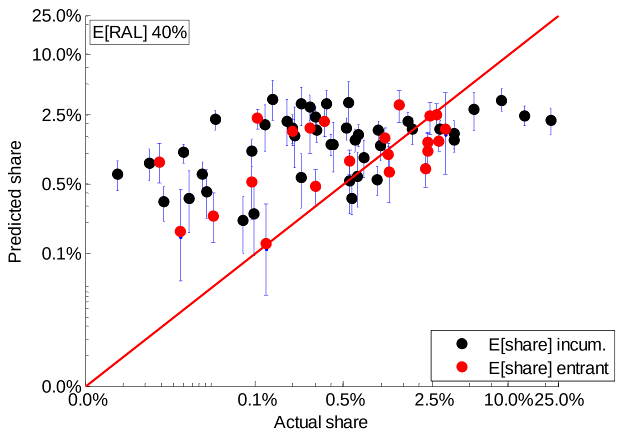
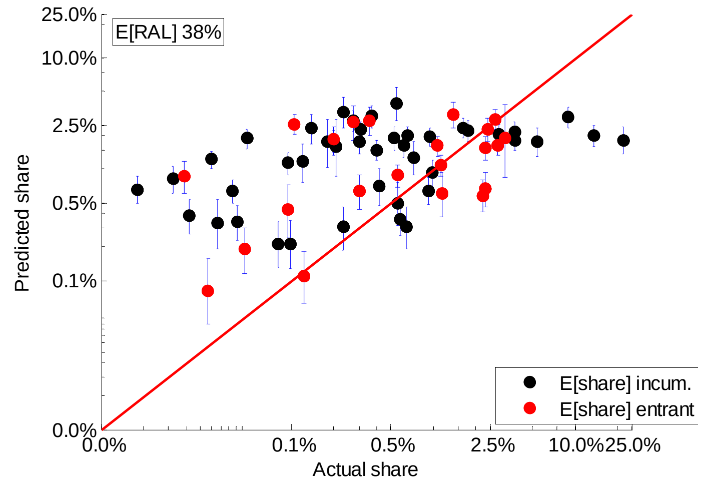


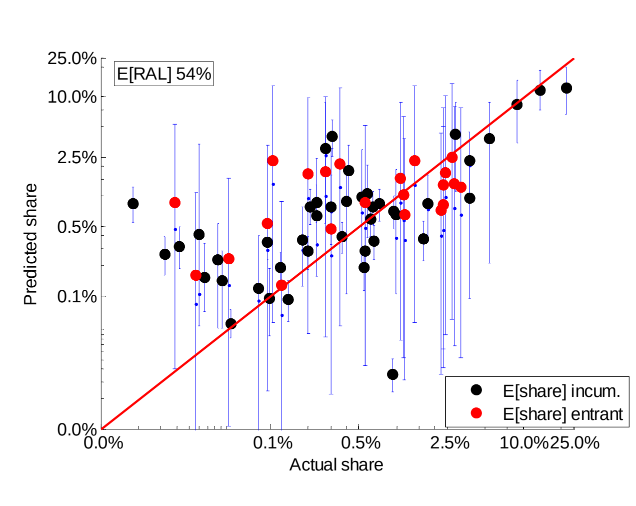
Figure 10 — Select trace plots of Bayesian-estimated correlated mixed logit model parameters

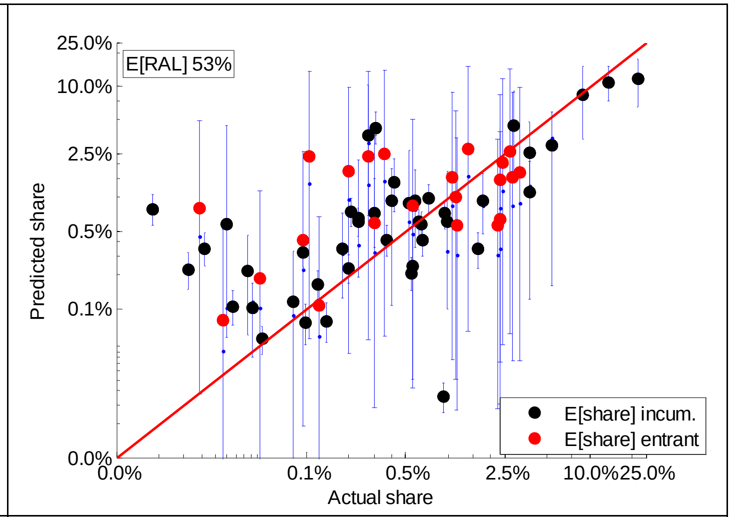
7.18 Appendix R

Table 34 — Actual versus predicted 2.5-97.5% interval of shares of 1-year-forward predictions for the Bayesian full and independent models using each of the ASC generation methods

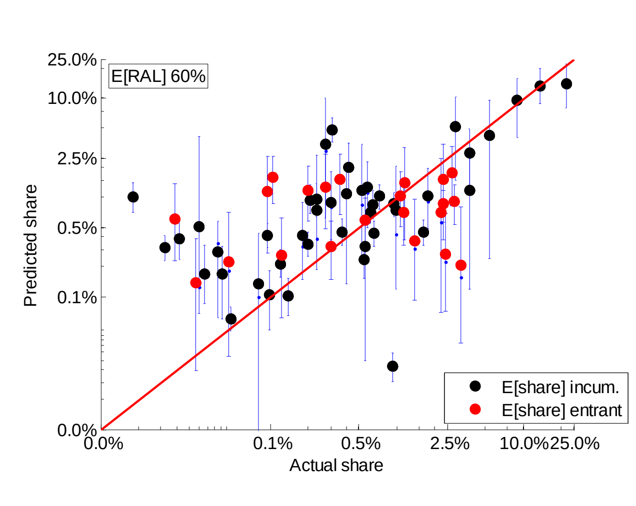
<table>
<tr><th></th><th>Bayes full</th><th>Bayes independent</th></tr>
<tr><td>Actual v. predicted Method 0 (None)</td><td>
</td><td>
</td></tr>
</table>

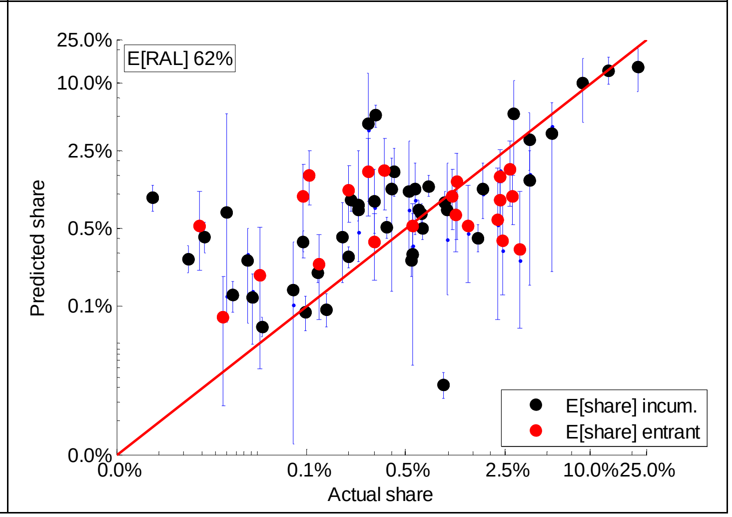
Actual versus predicted Method 1 (All)



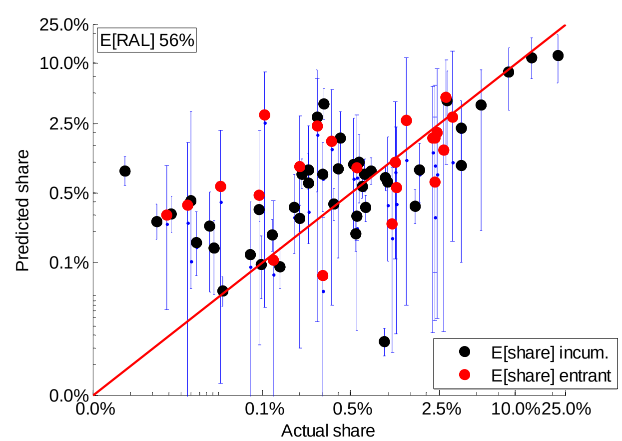


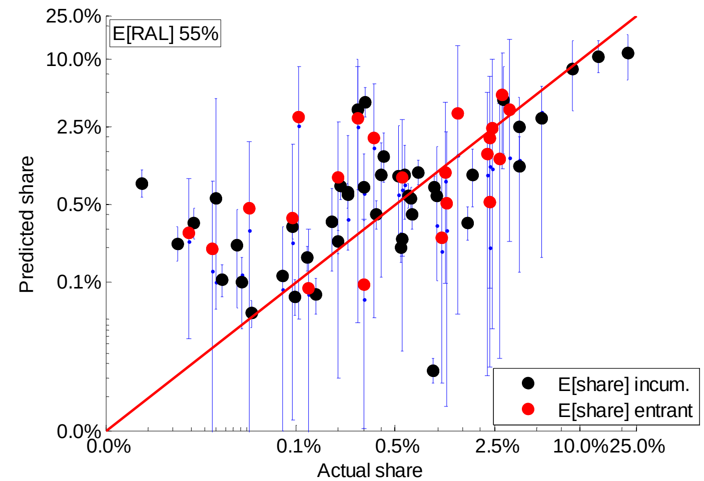
Actual versus predicted Method 2 (Nearest neighbor)



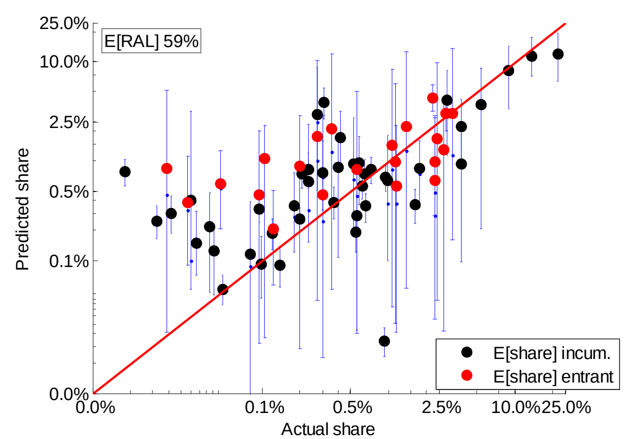


Actual versus predicted Method 3 (Brand)





Actual versus predicted Method 4 (Make-Model)



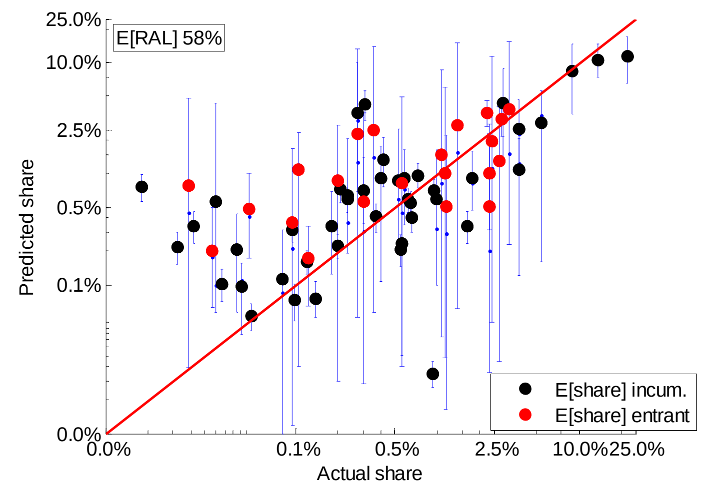
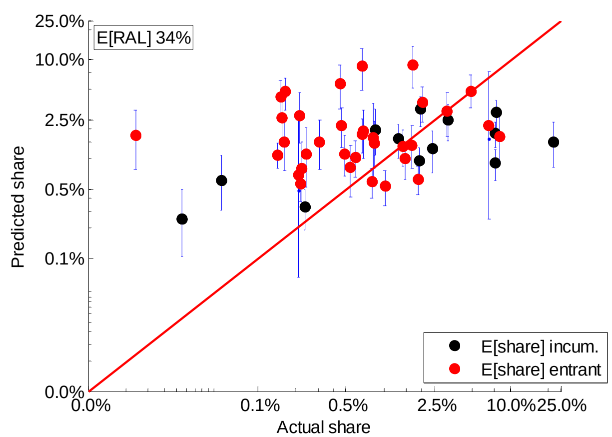
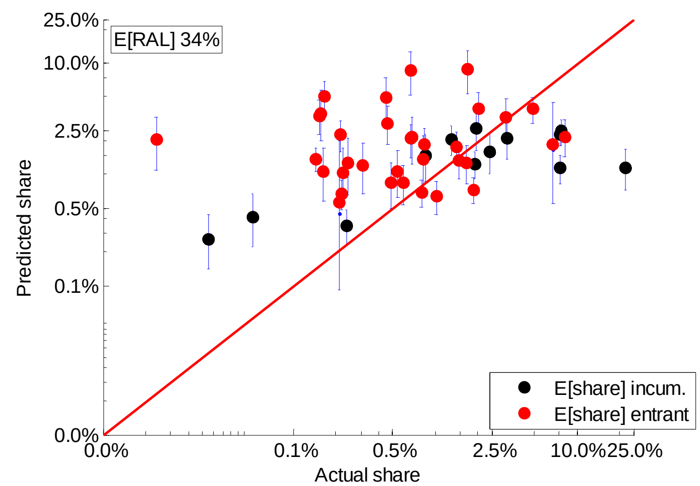
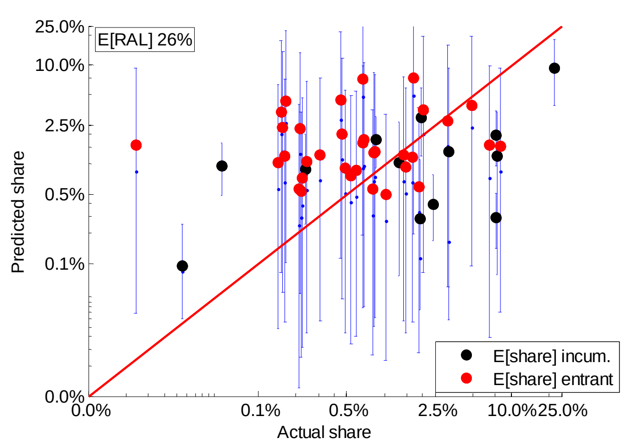
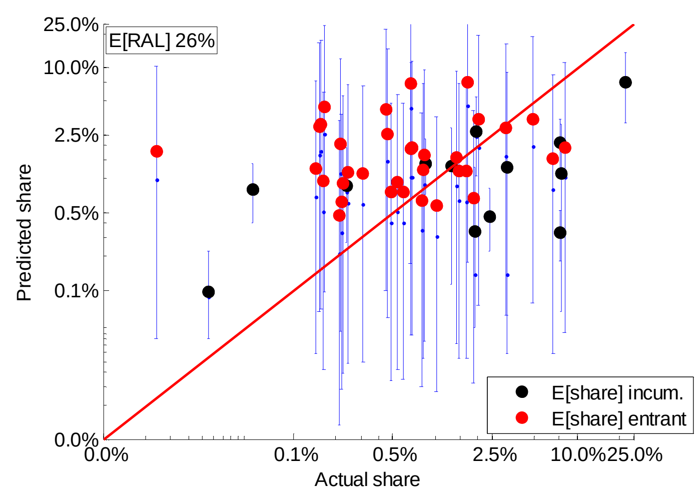
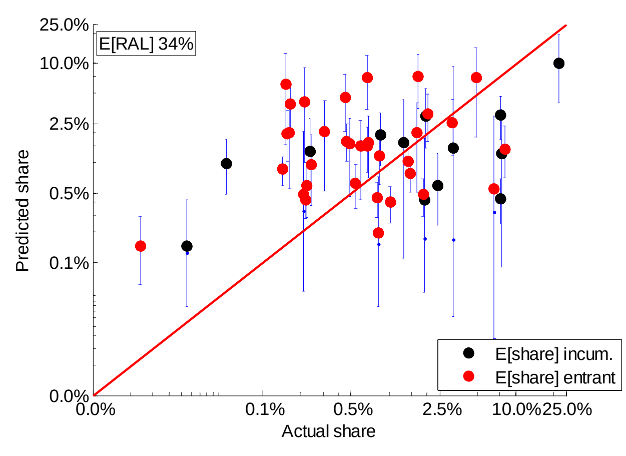
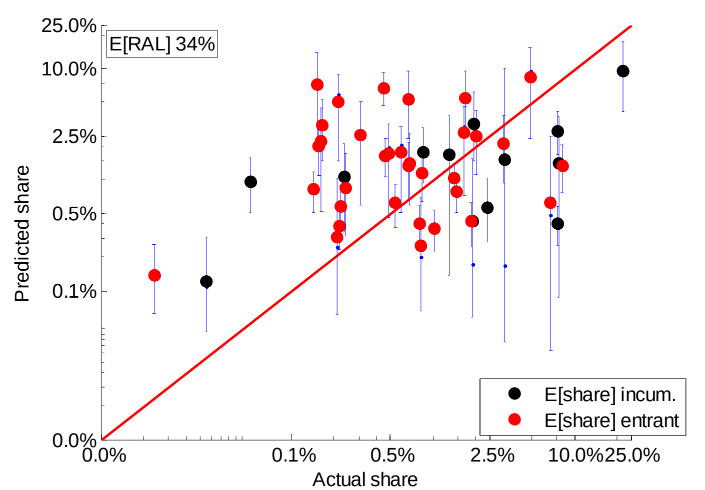


Table 35 — Actual versus predicted 2.5-97.5% interval of shares of 5-year-forward predictions for the Bayesian full and independent models using each of the ASC generation methods

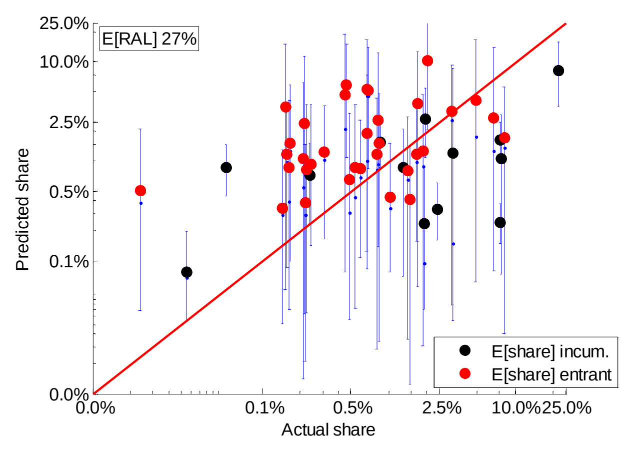
	Bayes full	Bayes independent
Actual versus predicted Method 0 (None)		
Actual versus predicted Method 1 (All)		

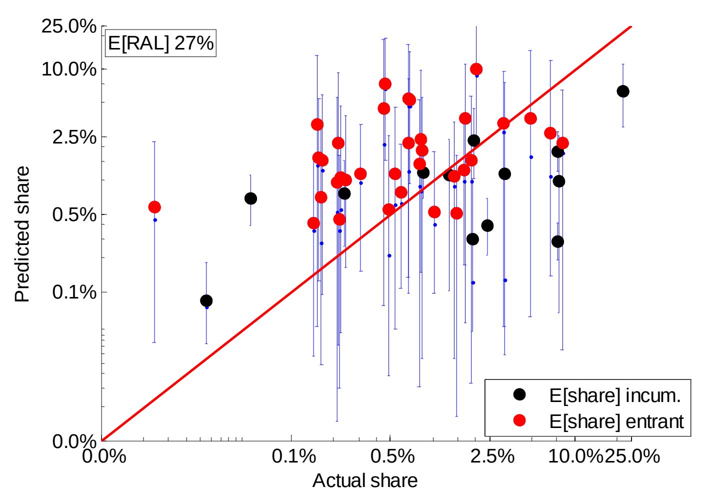
Actual versus predicted Method 2 (Nearest neighbor)





Actual versus predicted Method 3 (Brand)





Actual versus predicted Method 4 (Make-Model)