Transportation Research Part B 84 (2016) 182–210

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Forecasting light-duty vehicle demand using alternative-specific constants for endogeneity correction versus calibration

C. Grace Haaf[a], W. Ross Morrow[b], Inês M.L. Azevedo[c], Elea McDonnell Feit[d], Jeremy J. Michalek[a,c,*]

[a] *Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[b] *Ford Motor Company, Ford Research and Innovation Center, 3200 Hillview, #200, Palo Alto, CA 94304, USA*
[c] *Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[d] *Marketing, Drexel University, Philadelphia, PA 19104, USA*

ARTICLE INFO

Article history:
Received 23 February 2015
Revised 10 November 2015
Accepted 19 November 2015
Available online 13 January 2016

Key words:
Automotive demand
Discrete choice model
Demand forecast
Forecast uncertainty

ABSTRACT

We investigate parameter recovery and forecast accuracy implications of incorporating alternative-specific constants (ASCs) in the utility functions of vehicle choice models. We compare two methods of incorporating ASCs: (1) a maximum likelihood estimator that computes ASCs post-hoc as calibration constants (MLE-C) and (2) a generalized method of moments estimator that uses instrumental variables (GMM-IV) to correct for price endogeneity. In a synthetic study we observe significant coefficient bias with MLE-C when the price-ASC correlation (endogeneity) is large. GMM-IV successfully mitigates this bias given valid instruments but exacerbates the bias given invalid instruments. Despite greater coefficient bias, MLE-C yields better forecasts than GMM-IV with valid instruments in most of the cases examined, including most cases where the price-ASC correlation present in the estimation data is absent in the prediction data. In a market study of U.S. midsize sedan sales from 2002 – 2006 the GMM-IV model predicts the 1-year-forward market better, but the MLE-C model predicts the 5-year-forward market better. Including an ASC in predictions by any of the methods proposed improves share forecasts, and assuming that the ASC of each new vehicle matches that 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 may be invalid.

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 market. In particular, the automotive literature employs DCMs to understand drivers of purchase behavior (Allcott & Wozny, 2014; Copeland et al., 2011; Li et al., 2011; Dasgupta et al., 2007; Sudhir, 2001; Lave & Train, 1979) and predict future vehicle market shares (Greene et al., 2004; Greene et al., 2005; Duvall & Knipping, 2007; Balducci, 2008; Lin & Greene, 2010; U.S. Energy Information Administration, 2011). Alternative models of household vehicle choice can be used to forecast demand, but we focus exclusively on DCMs. DCM specifications include popular multinomial logit, nested logit, mixed

* Corresponding author. Tel.: +1 412 268 3765; fax: +1 412 268 2908.
E-mail addresses: grace.haaf@gmail.com (C.G. Haaf), wmorro13@ford.com (W.R. Morrow), iazevedo@cmu.edu (I.M.L. Azevedo), efeit@drexel.edu (E.M. Feit), jmichalek@cmu.edu (J.J. Michalek).

http://dx.doi.org/10.1016/j.trb.2015.11.012

logit, and probit models (Train, 2009), as well as variants of these models, such as the generalized multinomial logit model (Fiebig et al., 2009).

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, for which the modeler selects a set of attributes related to product choice and designs hypothetical products for respondents to choose. These studies can avoid issues such as omitted variables, endogeneity, and multicollinearity. However, such studies typically rely on the respondent to make hypothetical choices that do not necessarily reflect real purchase choices in a market context. In contrast, revealed preference data (often aggregate market sales data) track 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 and/or endogeneity issues, depending on whether the correlated attribute is observed by the modeler (Swait et al., 1994). Different sets of econometric assumptions are needed for the stated and revealed preference modeling approaches. 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 likely be missing information about some of the attributes that determine consumer choices. In order to address the utility not captured by explanatory variables, modelers frequently include an alternative-specific constant (ASC) in the utility function. For example, following Lave and Train (1979):

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

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 treated as a random variable. Consumer-specific attributes like income or family size can also be included in the utility function, but are not included in this study since we use aggregate sales data where this information is not available.

An interpretation of the ASC introduced into the automotive demand context by Lave and Train (1979)[1] and popularized by Berry et al. (1995) is that it represents the mean cumulative effect of all product attributes that consumers use to evaluate a product but that are unknown to the researchers. Alternative terms for the ASC when it is used to represent omitted variables include the unobservable (Allcott & Wozny, 2014; Sudhir, 2001; Berry et al., 1995; Berry, 1994; Berry et al., 1999), the unobserved product characteristic or attribute (Berry et al., 2004; Beresteanu & Li, 2011), market-level disturbance (Petrin, 2002), and demand shock (Dubé et al., 2012; Knittel & Metaxoglou, 2012). However, the ASC need not necessarily be viewed as a representation of unobserved attributes but rather can be included as a purely mathematical construct to improve model fit (Greene et al., 2004; Greene et al., 2005), sometimes referred to as a calibration constant (Bunch et al., 2011). Indeed, 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 to measure 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 address potential sources of coefficient inconsistency and bias — especially price endogeneity (Berry et al., 1995). Inconsistency arises if the ASC is correlated with an observed attribute, such as price. 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[2] 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 estimated price coefficient bias (measured as the difference between estimates when endogeneity is ignored versus when it is corrected for) is in the expected directions (Berry et al., 1995; Villas-Boas & Winer, 1999; Chintagunta, 2001). The explanatory literature implements estimation techniques that mitigate endogeneity bias — 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 (2014). Valid instruments require that they are correlated with the endogenous observed vehicle attribute(s), uncorrelated with the unobserved attribute(s), and do not affect the dependent model variable (market share) except through the observed attributes (Wooldridge, 2010). 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 generate models that underperform those

[1] Lave and Train (1979) 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 (2007), 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.

[2] 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.” (Wooldridge, 2006)