ESTIMATING DEMAND FOR AUTOMOBILE INDUSTRY IN THE U.S. MARKET: 2010 – 2013

A thesis presented

by

Xiaonan Qin

to

The Department of Economics

In partial fulfillment of the requirements for the degree of
Master of Arts

in the field of

Economics

Northeastern University
Boston, Massachusetts
December, 2014

ESTIMATING DEMAND FOR AUTOMOBILE INDUSTRY
IN THE U.S. MARKET: 2010 – 2013

by

Xiaonan Qin

ABSTRACT OF THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Economics
in the College of Social Sciences and Humanities of
Northeastern University
December, 2014

By applying market-level data on quantities, prices and vehicle characteristics in the year of 2010 and 2013, I conduct an estimation of automobile demand in the U.S. market using the conditional logit and nested logit model. By testing and applying the instruments variables, I address the exogeneity problem of prices and within-group market share. The demand for cars and light trucks are both elastic. The nested logit regression indicates that models within the same group/nest are better substitutes than models across the group. The results imply that consumers prefer the efficient vehicles. However, the electric vehicles and low displacement vehicles are not preferred over high displacement fuel vehicles by consumers.

TABLE OF CONTENTS

I Introduction

The discussion of automobile demand estimation has been a hot topic for decades. As one of the world's most important economic sectors that generate a large number of revenue for firms and jobs in countries like the U.S, Germany and China, the auto industry has been intensively studied. Company managers are eager to know about demand estimation because it helps them to forecast and decide next season's production capacity and sales revenue. Economists are keen to study and research automobile firms' market power, pricing strategy and demand elasticity since it provides a good example of an oligopolistic differentiated products market. Policy makers are interested in automobile demand estimation because it helps them forecast demand for highway infrastructure.

Other public policy issues include the development and introduction of environmentally friendly electric or new power vehicles, high trade tariff in the auto industry, and emerging auto markets in developing countries. From 2008 to 2010, the auto industry was heavily affected by the global financial crisis. The crisis affected several countries in Europe and Asia, and mostly the major auto companies in the United States. GM, Chrysler and Ford asked the U.S. government for a $50 billion bailout to avoid bankruptcy. After this slowdown, companies are now trying to learn from lessons and figuring a way to quickly recover (The Washington Times, 2012).

Consumers' preferences for automobiles are, in general, heterogeneous across different type of consumers. Sports-utility vehicles might be more preferred by households in Boston where heavy snows always fall during the winter. Other industries like auto repair and maintenance, insurance, and fuel also pay attention since their businesses are closely related to automobile's demand. Estimating and forecasting the sales trends of the automobile industry might also be a good way to analyze the global finance trend since the auto industry is always associated closely to financial and fuel crises.

This paper empirically analyzes the demand in the auto industry in the U.S. market by using a discrete choice model. In discrete choice modeling, individuals choose from a set of options to gain the highest utility. The utility depends on the product attributes and price. Further, this paper follows Berry's (1994) theoretical framework and applies a conditional logit and a nested logit model. I also construct two different nest structures and compare the results based on fitness and magnitude, and address the issue of endogeneity by using instrumental variables.

I combine the datasets from Ward's Automotive Yearbook and from Fuel Economy to obtain a panel data on market price, number of vehicles sold, and products specifications with 473 observations in total. The results of within group market share show stronger substitution patterns for models within the same group than models across different groups. That is, when facing an increase of model's price that the consumers choose to buy in the first place, consumers switch to models within the same group more often than models in other groups.

The organization of this paper is as follows. In section 2, a literature review is provided. Section 3 presents the theoretical framework and econometric methods. Section 4 shows the empirical results and section 5 concludes.

II Literature Review

Berry (1994) proposes a framework of supply-and-demand estimation using market-level data only and addressing the price endogeneity issue at the same time. In common with other previous literature on the discrete choice estimation, consumers choose from a set of products, and the utility of consumers depends on product characteristics, price, and unobserved product characteristics. Market shares are then derived as the aggregate outcome of consumer decisions. The variable price suffers from simultaneous bias as well as omitted variable bias coming from unobserved product quality, which can be addressed by utilizing instrument variables. However, in a discrete choice model, both prices and unobserved characteristics enter the equation in a nonlinear fashion which makes hard to implement traditional instrumental variable estimation. Berry (1994) proposes a method that inverts the market share to linearize and isolate the mean utility of each product, where mean utility can be estimated as a function of products characteristics and prices. The endogeneity issue of price can then be addressed by using instrumental variable techniques. Berry illustrates three examples specifically, the logit model, the nested logit model and the vertical differentiation model. I use his method of logit model and nested logit specifically to empirically analyze the demand in the automobile industry.

Berry et al. (2004) use a combination of micro and macro level data to analyze the demand of the auto market. They are able to obtain individual survey data from General Motors, especially on consumers' "second-choice." That is the purchase that consumers would have made if their preferred product are not available. This dataset help them to define the alternative/substitution patterns that can be measurements of product unobserved characteristics in consumers' mean utility function. Applying the theoretical framework of Berry (1994), they estimate auto market demand empirically, and also do two prediction exercises. They evaluate the potential demand of new "high-end" SUVs, and a major production decision of shutting down the Oldsmobile division of General Motors. For this paper, individual choice dataset is very hard to collect, but their empirical work inspires me to think of the proper sets of product characteristics as well as substitution patterns of utility.

Pertin (2002) follows Berry et al. (1995) and proposes a technique that augments the market share data relating to the demographics of consumers to the products specifications they purchase. By applying the new vehicle data on passenger cars, sport-utility vehicles, station wagons and minivans, and on consumer data from Consumer Expenditure Survey, he obtains the estimates that reflect demographic-drive differences in tastes for observe attributes. The paper also suggests that consumers gain far more benefit from the introduction of minivans than the costs of development and the profits obtained by the innovator.

Murry (2014) disagrees with Berry et al. (1995) on ignoring the distinction between manufacturing and retailing when estimating demand in the automobile market. He estimates a random coefficients logit model of demand for new cars using new car transaction data and local market advertising data in the state of Virginia. After applies the results from demand estimation, Murry discovers that manufacturer and dealer surplus depends not only on the markups, but also on the advertising spending. Through a set of counterfactual exercises, he also quantifies the pricing and advertising incentive problems in the industry and evaluates the effect of dealer-manufacturer regulations on welfare.

Deng and Ma (2010) use market-level data of cars on quantities, prices and automobile characteristics from 1995 to 2001 to conduct market analysis and demand estimation in the Chinese automobile industry. They apply a nested logit model and the discrete choice method on the demand side. On the supply side, they assume a Bertrand model to analyze the price and markups set by automobile manufacturers. Their results suggest a high markup of China's auto manufacturers, but the market power decline during the late 1990's. They also relax the assumption by allowing the within–group heterogeneity to varies across groups. By comparing price elasticities in the Chinese and the U.S. automobile market, they find the magnitudes are much higher in the Chinese market.

This paper extends previous literature by applying a more recent dataset, more numbers of models available in the market as well as the information of electric vehicles. By including light trucks models, I analyze the substitution patterns of light trucks versus regular cars. The paper also develops two nest structures based on the assumption of consumer behaviors.

III Data Description

The data is obtained from Ward's Automotive Yearbook 2010, 2011, 2013 and 2014, specifically from U.S. Light Vehicle Sales by Segment 2010 & 2013, '10 & '13 Model Car and Light Truck U.S. Specifications and Prices, and from the Fuel Economy MPG data as well. I construct a panel data with the hope that it contains more information than cross-section data. Limited by the availability of data sources and time, I pick only two years 2010 and 2013 and obtained 473 observations in total.

WardsAuto has covered the auto industry for more than 80 years, and provides a wide range of information on products. It provides auto industry data information not only on U.S., Canada and Mexico, but also on Asian and European countries. Ward's Yearbook divides vehicle models into different market segments based on the price, body style and model size. The whole market is defined as "light vehicles," which includes cars and light trucks. Cars include small, middle, and large cars, luxury cars and special cars[1]. Light trucks include Cross/ Utility vehicle[2], Sport/ Utility vehicle[3], vans and pickups.

The class defined for each model in WardsAuto is not "constant" in 2010 and 2013. For example, Hyundai Azera is defined as "lower luxury" in 2013 but "upper small model" in 2010 because of price difference. Also, WardsAuto gives too many detailed classes, which makes the analysis complicated. Thus, to assign each model to a class, I obtain a dataset from Fuel Economy that matches with the models in WardsAuto. This dataset provides for 14 different classes: six classes are for cars and eight are for light trucks.

Product characteristics obtained are drive type[4], body style, wheel base size, length, width, height, weight, displacement, horsepower (HP), torque, estimated miles per gallon (MPG) and the information on whether the model is electric vehicle or not.

---

1 Special cars are defined as 2-door, 4-passenger or 2+2 seating.

2 Cross/ Utility vehicle are typically wagon or hatchback body style with unibody construction, front- or all-wheel-drive and passenger vehicle quantities the dominant characteristic with limit off-road capability.

3 Sport/ Utility vehicle are off-road capability a strong characteristic, body-on-frame or unibody construction, offering standard or optional low-speed transfer case gearing or all terrain management system and minimum 7.5-in. (91-mm) ground clearance.

4 Drive Type: FWD is front-wheel drive; RWD is rear-wheel-drive; AWD is a type of 4-wheel-drive in which power is automatically transmitted to front or rear axels depending on traction conditions.

Data on price is the manufacture retail prices on base model at the beginning of the model year in U.S. currency. Price in 2010 is adjusted for inflation choosing 2013 as a base year. Some models entered the market after 2010, so they only have data for 2013. Likewise, a model may exit the market after 2010, so data is only available for 2010. Combining numbers of vehicle sold, price and specifications from WardAuto and classes from Fuel Economy, I obtain an unbalanced panel data treating each model/year as an observation. The total number of observations is 473.

Besides the characteristics I obtain, I construct the variables $size\ (m^2) = length \times width \times height$, and $HP\ /\ weight =$ horsepower per 100 lbs. Since the models include fuel vehicles and electric vehicles as well, estimated MPGe in the city for electric vehicles are very different from MPG of fuel vehicles and may influence the estimates. As suggested by Berry et al. (1995), I constructed MP$=estimated miles per dollar. For fuel vehicles, miles per dollar equals estimated miles per gallon divided by national annually average gasoline price in that year (either in 2010 or 2013)[5]. For electric cars (Nissan Leaf and Chevrolet Volt), I convert estimated MPGe to k.w.h per 100 miles[6], and then multiply k.w.h to the national average electricity bill which is about 12 cents, MP$=100 miles/price. MP$ is also adjusted for inflation.

The total numbers of vehicles sold is 15,532,232 for 2013 and 11,554,518 for 2010, with a 34.43% increase in total and nearly a 10.36% increase annually. The percentages of cars and light trucks relative to total quantity sold stay almost the same when comparing 2010 and 2013. The percentage of imported cars increases, while imported light trucks remains relatively

5 U. S. Energy Information Administration.

6 www.fueleconomy.gov

constant. The percentage of imported CUVs increase slightly, but imported SUVs drops about 5%. By assuming each household chooses one transportation method, and one vehicle is just enough for one household, the measure of market size is the total number of households in the U.S., and is taken from the United States Census Bureau.

Table 2 shows the descriptive statistics. All of the variables have positive skewness except for length. Table 3 shows the mean of numbers of vehicles sold and median of other variables grouped by model's body style. The available data mostly concentrate on 4-dr cuv and 4-dr sedan style. 2-dr convertible and 4-dr coupe have relative small sale quantity but relative high price, whereas 2-dr p.u. has the largest quantity. 2-dr style has relative high HP/weight mainly because of the less weight.

Table 4 shows the data summary grouped by year. The year of 2013 contains more numbers of observations. Numbers of automobiles sold increases in 2013 as well. Table 5 summarizes the data by drive type. Most of the models are FWD and RWD. FWD drive type has relatively very large quantity and low prices compared to other drive types.

Numbers of vehicles sold and price for each brand are listed in table 6. For prices over $30,000, sales are mostly under 25,000 except for Lexus and BMW. For prices under $30,000 but over $20,000, sales exceed 25,000 except for Mitsubishi.

Compared to Berry et al.'s (1995) dataset from 1971 to 1990, the data I obtain is more recent but only has two years. The dataset does not contain information on air conditioning, but it includes the information on electric vehicle. Additionally, Automotive News Market Data Book shares the information on numbers of vehicles sold from Wards. The models in Automotive News Market data have the same classes that include passenger cars, CUVs, SUVs, and vans. Within one year, the most observations Berry et al. (1995) have is only 150 since their data does not include CUVs, SUVs and vans, but only cars, whereas this data contains a larger number of observations within one year.

IV Econometric Model

I apply two different models, a logit model and a nested logit model to estimate automobile demand following Berry's (1994) method by using macro-level data and product attributes only in the equation.

4.1 Logit Model

Consumer chooses from products among 0,1…N. There are N+1 options in total, where N denotes the number of car models available in U.S. market and 0 is the outside good. The utility of consumer *i* for product *j* is given by

(1) $$u_{ij} = \beta X_j - \alpha p_j + \xi_j + \varepsilon_{ij}$$

where X_j are product attributes. p_j is the price of each model. X_j and p_j are observed, whereas ξ_j includes all unobserved product attributes such as inner design, the auto companies' and brand's reputation, etc. By assuming ε_{ij} is identically and independently distributed with type I extreme value distribution function, the market share of product *j* is

(2) $$s_j = \frac{exp(V_j)}{\sum_{k=0}^{N} exp(V_k)}$$

where

(3) $$V_j = \beta X_j - \alpha p_j + \xi_j$$

Normalizing V_0 of the outside good to zero, taking the log of s_j/s_0, and substituting V_j back into equation (3), I obtain the linear expression in terms of relative market share

(4) $$\ln(s_j) - \ln(s_0) = \beta X_j - \alpha p_j + \xi_j$$

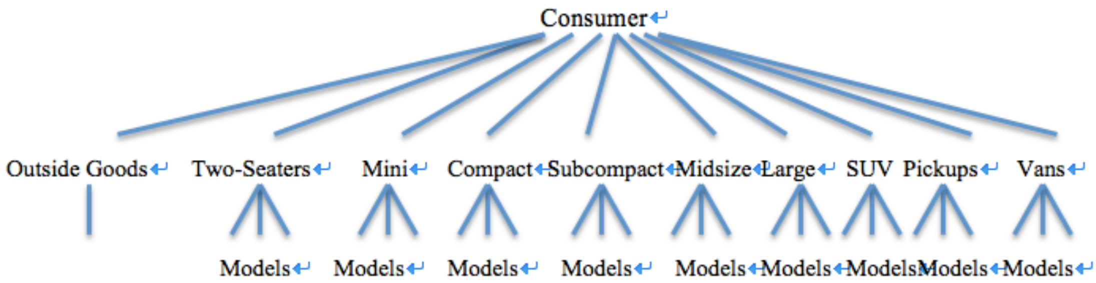
The logit model has a drawback, the assumption of independence of irrelevant alternatives, which gives that the cross-price elasticity between two products is $\eta_{jk} = \frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j} = -\alpha p_k s_k$, which only depends on the price and market share of product k, but not on other similar products. That is the cross-elasticity of any product with respect to product k are all equal, $\eta_{jk} = \eta_{lk}, \forall\, j, k, \mathrm{l} \in 1 \ldots N, j \neq l \neq k$, which is not very realistic. The own-price elasticity is $\eta_j = \frac{\partial s_j}{\partial p_j}\frac{p_j}{s_j} = \alpha p_j(1 - s_j)$.

4.2 Nested Logit Model

I apply and compare two nest structures to see which one fits demand estimation better.

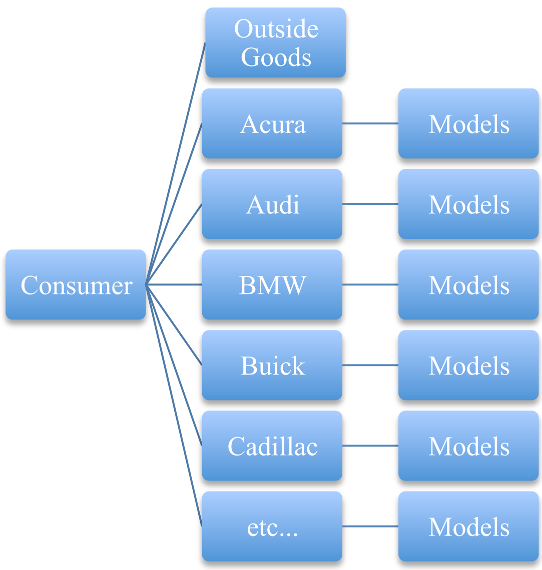
Assume that a consumer's decision to buy a car is as in Figure 1. The first stage is to choose between purchasing a certain class of car or an outside good. I combine and select nine classes in total (i.e., two-seaters, mini, subcompact, compact, midsize, large, SUVs, pickups and vans). At the next level, the buyer decides which model to buy given the class he has chosen. Compared with other nested logit structures (e.g. at the first stage, consumers decides which brand to buy), consumers make decisions based on family size first.

Figure 1



Another nest structure I use is as in Figure 2. The first stage is to choose which brand to buy (e.g., Acura, Audi, BMW etc.), or an outside good. An outside good in this case can be train, bus, walking etc. There are 36 groups in total. At the second level, the buyer decides which model to buy given the brand he or she has chosen. Compared with the first nested logit structure, consumers make decisions based on the preferences for each brand. They can choose one particular brand because of the reputation, the overall design style of that brand, or the development history and imported country of that brand's manufacturer.

Figure 2



The consumer utility becomes:

(5) $$u_{ij} = \beta X_j - \alpha p_j + \xi_j + \zeta_{ig} + (1 - \sigma_g)\varepsilon_{ij}$$

where $[\zeta_{ig} + (1 - \sigma_g)\varepsilon_{ij}]$ is an extreme value random variable [Cardell (1997)]. g is the class or group which the model corresponds to. I denote M_g as the set of car models in group g. Then the market share of model *j* in group g as a fraction of this total group share is:

(6) $$s_{j/g}(\mathrm{V}, \sigma) = \frac{exp[V_j/(1-\sigma_g)]}{\sum_{j \in M_g} exp[V_k/(1-\sigma_g)]}$$

Denoting the denominator as D_g, the probability of choosing one of the groups is:

(7) $$s_g(\mathrm{V}, \sigma) = \frac{D_g^{(1-\sigma_g)}}{\sum_g D_g^{(1-\sigma_g)}}$$

giving the market share of *j*

(8) $$s_j(\mathrm{V}, \sigma) = s_{j/g}(\mathrm{V}, \sigma)\overline{s_g}(\mathrm{V}, \sigma) = \frac{exp[V_j/(1-\sigma_g)]}{D_g^{\sigma_g}[\sum_g D_g^{(1-\sigma_g)}]}$$

where $\overline{s_g}$ represents the observed group share. The outside good is the only member of group zero. Then normalizing V_0=0, s_0=1/[$\sum_g D_g^{(1-\sigma_g)}$], and taking logs of relative market share, I obtain the linear equation for a nested logit demand estimation.

(9) $$\ln(s_j) - \ln(s_0) = \beta X_j - \alpha p_j + \sigma ln(\bar{s}_{j/g}) + \xi_j$$

The own- and cross-price elasticities are:

$$\eta_j = \frac{\partial s_j}{\partial p_j}\frac{p_j}{s_j} = -\alpha p_j\left(\frac{\sigma_g}{1-\sigma_g}s_{\frac{j}{g}} + s_j - \frac{1}{1-\sigma_g}\right), \mathrm{j} \in M_g$$

$$\eta_{jk} = \frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j} = -\alpha p_k\left(\frac{\sigma_h}{1-\sigma_h}s_{\frac{k}{g}} + s_k\right), \mathrm{j} \neq \mathrm{k}, \mathrm{j} \in M_g, \mathrm{k} \in M_h$$

When $g \neq h$, then the cross- group cross-price elasticity of demand becomes $-\alpha p_k s_k$.

4.3 Endogeneity Issue

Since the price is determined by the equilibrium of supply and demand, it is correlated with the error term and the coefficient form OLS estimation suffers from simultaneous bias and The OLS estimator also suffers from omitted variable bias because the price is correlated with unobserved product attributes. I use two-stage least squares instrumental variable to deal with this econometric issue. Possible valid instruments will be factors that are correlated with price, but are not correlated with ξ_j. As suggested by Berry et al. (1995), possible instruments include cost shifters, the characteristics of competing models, and characteristics of models produced by the same manufacturer. The cost shifters are size, weight, displacement, torque, horsepower, estimated MPG, wheel base length, body style dummies and drive type dummies. For example, if one of the characteristics is the size, then the instruments for product *j* include the size of the car *j*, the sum of size across own-firm models, and the sum of size across competing models of rival firms.

The within group market share is also endogenous. The error term ξ_j is the unobserved characteristics of models, which determine the consumer utility and market share as well. Thus, the unobserved characteristics are correlated with within group market share. I use instrumental variables to address this endogeneity source. The possible instruments include the number of models available within the group and characteristics of competitors within the group.

To test for instrument validity, I perform two tests to address the relevance and exogeneity of the proposed IVs. To test the relevance of the IVs, I regress the endogenous variable on all exogenous variables (control variables and IVs). Then I perform a F-test on all instruments excluding control variables/observed attributes. If the F-test is greater than 10, it indicates that the IVs are strong instruments. Otherwise, they are weak instruments.

Since I have more instruments than endogenous variables, I can do overidentifying restrictions tests to see whether the instruments are exogenous or not. Based on a regression-based Hausman test, I run the two-stage-least-square estimation and collect the residuals from the second stage. Then I regress the residuals on the instruments and control variables, which is shown in equation (10). Let F denote the homoscedasticity-only F-statistic to test the hypothesis that $\delta_1 = \cdots = \delta_m = 0$. The overidentifying restrictions test statistic is J=mF. The null hypothesis of instrument exogeneity implies that the instruments are not correlated with the errors. J is distributed as χ^2_{m-k}, where m-k is the degree of overidentification. It is the number of instruments minus the number of endogenous variables. If J exceeds the critical value in chi-square distribution, I reject the null hypothesis and conclude that at least one of instruments is not exogenous.

$$\hat{\mathcal{U}}_i^{TSLS} = \delta_0 + \delta_1 Z_{1i} + \cdots + \delta_m Z_{mi} + \delta_{m+1} W_{1i} + \cdots + \delta_{m+r} W_{ri} + e_i \tag{10}$$

I also apply a fixed effect model using this panel data.

$$\ln(s_{it}) - \ln(s_{0t}) = \beta X_i - \alpha p_{it} + \sigma ln(\bar{s}_{j/g}) + \xi_j + \lambda_t + \mathcal{U}_{it} \tag{11}$$

IVs also allow me to solve the endogeneity problem assuming the instruments are valid.

V Empirical Results

Table 7 shows the empirical results of estimating the logit demand model. Columns (1)-(4)

are regressions of relative market share on price, size, displacement, horsepower over weight, miles per dollar, and electric vehicle dummy, using OLS and IV estimation for equation (4). To begin with, the coefficients of price variable are -0.039 and -0.037 for OLS estimations: the absolute value in column (3) goes up to 0.052 for the IV model using all possible instruments. Applying tests for both relevance and exogeneity (to be discussed below), I pick the sum of size across own-firm models and sum of displacement across own-firm models to be the combination of most valid instruments. The absolute value of the price coefficient goes to 0.112 in column (4), which indicates the endogenous problem of OLS estimation.

To test for instruments' relevance, I first run a regression of price on all control variables as well as two instruments. For the model in column (4), the joint F-test on the instruments (sum of size across own-firm models and sum of displacement across own-firm models excluding control variables, size, HP/weight, MP$, and displacement) is 23.27, bigger than 10, which indicates that the sum of size and displacement across own-firm models are relevant to price, and therefore are strong instruments. Then I run the overidentifying restriction test. J equals 0.56, smaller than the 90% critical value. I fail to reject the null hypothesis, which means that the two instruments may be either both exogenous or both endogenous. Since I fail to reject the hypothesis (where rejection would indicate that at least one of the instruments is endogenous), the instruments are possibly to be both exogenous.

Table 8 shows the results of estimating the nested logit models. The columns (1) and (2) are for the nested logit model defining each class of cars as a group. The price coefficients change from -0.020 to -0.056 after applying IVs for both price and within-group market share $\ln(s_{j/g})$.

Columns (3) and (4) are the nested logit model denoting each brand as a group. The magnitude of the price coefficient after using IVs goes up to -0.068 in column (4). The magnitude of the price coefficients in nested logit model is not as large as the price coefficient of standard logit model under IVs estimation. The instruments I used in the model to address the endogeneity of price and within-group market share $\ln(s_{j/g})$ are the sum of own-firm models' size, the sum of own-firm models' displacement, the sum of rival-firm models' size and model's body type dummies.

The results of relative market share within group are all positive and significantly different from zero, which indicates the auto models within the same group are better substitutes for each other than autos across groups. Comparing the two nested logit models, the R squares are similar. The magnitude of within group share in column (4) is higher than in column (2). That is, when consumers consider what models to buy, they are more likely to consider which brands they prefer instead of what size they want.

The estimates of automobile characteristics show the size, horsepower over weight, and miles per dollar all positively impact the relative market share, except that the coefficient of horsepower over weight is negative in column (1) of table (8). The coefficients on the size variable are significantly different from zero in all eight models in table (7) and (8). Coefficients of horsepower over weight are also significantly different from zero in all IV estimation models. On the other hand, the displacement does not always positively impact the market share and is not significantly different from zero. The coefficients of electric dummy are negative indicating that customers prefer the fuel vehicles to the electric vehicles.

In the fixed effect model in Table 9, the price coefficient is -0.039, which is the same as the first OLS model in table 7. By adding a year dummy in column (2), the price coefficient goes up and becomes not significant. Column (3) shows results of applying both fixed effect and IV method. The absolute value of the price coefficient increases to 0.519, and is significantly different from zero. When applying fixed effect and instruments in the second nested logit model (each brand is a group) in column (4), the price coefficient becomes not significant, but the coefficient of relative market share within group is very close to 1. It implies that consumers have more tastes that are related to that nest. That is, consumers choose a model because he or she prefers more of that brand of vehicles.

Using the cross- and own-price elasticity equations in section 3.1, table 10 shows the price elasticity of several representative automobiles by applying the price coefficient from model (4) in table 7 (the standard logit with IVs). The first column shows that own-price elasticity increases as price increases. The other column shows the cross-price elasticity of any model with respect to the model under consideration. For example, if the price of the Toyota RAV4 increases by 1%, its own market share will drop by 2.69% and the market share of Acura MDX will increase 0.0023%.

The average own-price elasticity is -4.10, compared to -3.28 in Goldberg (1995), who uses a nested logit model and U.S. data from 1983 to 1987. The magnitude does not change much with over this 20-year gap. Table 11 shows the own- and cross-price elasticity of different class of cars using the price coefficient from the first nested logit model (column (2) in table 8) and sales weighted averages. The magnitudes of own-price elasticities are different across groups. The largest magnitude is for two-seater cars followed by large cars, which are both relatively expensive compared to others. Following these two expensive classes, the elasticities of mini and subcompact cars are -2.46 and -2.57 respectively. In contrast, Deng ad Ma's (2010) analysis of the Chinese market suggests that subcompact is most elastic followed by mini cars.

Mini cars have the largest magnitude of within-group-cross-elasticity followed by two-seaters. The within-group-cross-elasticity of large cars, SUVs and vans are bigger than the rest of classes. That is, comparing the situation when the price of a SUV model and a compact model increase by 1%, according to the magnitude of within-group cross-elasticity, the market shares of other SUVs increase by 0.084% whereas the market shares of other compact cars increase by 0.022%. The SUV class has the largest magnitude of cross-group-cross-elasticity. Compared with the within-group cross-price elasticities, η_{jk} across group have much smaller magnitude suggesting that models within the same group are better substitutes than models in other groups. However, both cross-price elasticities are smaller than the own-price elasticities.

Using coefficient of price in the second nested logit model to calculate the own-price elasticities of several selected brands in Table 12, I find that overall the higher the price, the higher own-price elasticities. According to information in WardsAuto, I assign each brand to a region: Asia, U.S. and Europe based on the location of the headquarter company. I find that the within group cross elasticity of European brands are relatively large. One reason is that the prices of European models are expensive. However, Mini Copper is not very expensive comparing to others. If the price of one Mini Copper model increases, consumer switch to other Mini Copper models more often than the impact on other Toyota models when the price of a certain Toyota model increases.

Brands like GMC, Jeep, Land Rover and Ram have models that are mostly light trucks (based on the available data). They do not have significantly high own-price elasticities except for Land Rover whose price is expensive. Combining the information in table 11, the own-price elasticities of light trucks' classes are not very large. It means that if the government wants to reduce the consumption of light trucks because of their heavy pollution, a policy to imply a tax on light trucks is not very effective if the light truck models are relatively cheap in the first place. Consumers will still buy these models.

VI Conclusion

This paper applies recent macro-level panel data on market share and price as well as product attributes to estimate the demand for automobile in the U.S. market. By applying a logit model, two different nest structures and the linearization method suggested by Berry (1994), I analyze the demand parameters and substitution patterns. The price of demand is elastic, and compared to Berry et al. (1995) with passenger cars classes only, the coefficient of price does not change dramatically because of including the light truck classes.

The paper uses different econometric methods including OLS, IV estimation, and fixed effects. Starting with Berry et al.'s (1995) suggestion, I apply all possible instruments and test the relevance and exogeneity of them to find the valid instruments. The magnitude of price coefficient doubles after using IVs to address the endogeneity problem.

By comparing the two different nest structures, the model that defines brands instead of classes as groups fits better to the demand estimation. The cross-price elasticities imply that models within the same group are better substitutes than models across the group. If the consumer has chosen a model and the model's price increases, the consumer is more likely to switch to other models within the nest (the brand or class he or she has chosen).

The results of price elasticities can be used in many policy questions including trading policy, merger policy and environmental problems. The model can also be used to analyze and forecast the change of demand when one of the determinants changes assuming others are held constant. The positive impact of miles per dollar on automobile demand implies an optimistic prospect for efficient vehicles. The electric vehicles and low displacement vehicles are not preferred over high displacement fuel vehicles by consumers. Additionally, the elasticities of light trucks are not very large. Thus, the consumption of light trucks will not be affected much by implying a tax on heavy pollutions vehicles. Government can push more "education" to people to purchase environmental- friendly cars instead of implement higher prices.

The paper can be further extended to construct a higher level of nest structure, and to apply data on a longer time period. The endogeneity issue can be further addressed and solved by continuing searching for good IVs. It can also be extended by doing an analysis on supply side and markups.

Appendix

Table 1 U.S. number of light vehicles sold in 2010 and 2013

	2010	2010		2013	2013	
		% total number of vehicles sold	% imported within segment		% total number of vehicles sold	% imported within segment
Total Number	11,554,518	100.00%		15,532,232	100.00%	
Cars		48.77%	20.83%		48.84%	28.39%
Light Trucks		51.23%	15.18%		51.16%	15.60%
CUVs		24.54%	24.42%		25.50%	25.67%
SUVs		6.93%	18.30%		6.67%	13.49%

Number of observations: 473
Light trucks include CUVs, SUVs, vans and pickups.

Source: Ward's Automotive Year Book

Table 2 Summary statistics: U.S. light vehicle data

Year 2010 & 2013

	Mean	Std. Dev	Min	Max	5%	25%	50%	75%	95%	Skewness	Kurtosis
Quantity	54610.45	80674.69	10	713960	1346	8330	23416	66146	218249	3.24	17.86
Price ($1,000)	37.04	22.99	12.8	200.4	16.62	22.81	30.50	42.87	84.56	2.71	14.71
Wheel Base (ins.)	110.61	10.83	73.5	206.9	98.0	104.4	109.6	114.8	126.4	3.14	29.16
Length (ins.)	186.89	18.25	95.7	240.6	160.6	178.3	189.1	197.0	213.1	-1.21	8.12
Width (ins.)	74.48	9.93	61.4	182.6	67.4	70.9	73.3	76.3	79.7	8.45	86.57
Height (ins.)	63.32	7.79	49.0	86.3	53.5	57.6	61.0	68.5	74.6	76.80	83.70
Size (m^2)	14.55	3.41	6.5	26.4	10.15	12.17	13.67	16.48	21.07	0.89	3.71
Weight (lbs.)	3849.70	861.25	1808	6000	2584	3223	3760	4367	5578	0.47	2.73
Displacement (CC)	3140.04	1220.98	999	8382	1591	1999	2995	3696	5552	0.89	3.82
Horsepower (HP)	242.03	90.33	70	640	121	173	240	301	402	0.83	4.41
Torque (Nm.)	334.90	165.52	92	1998	167	230	339	377	565	4.63	45.09
HP/Weight	6.26	1.94	2.2	19.1	4.12	5.13	5.90	6.97	9.56	2.52	14.18
MPG	23.47	8.54	13	115	16	19	22	26	30	6.00	60.15
MP$	6.72	2.54	3.6	30.2	4.78	5.38	6.28	7.41	8.82	4.58	35.77

Number of observations: 473

Source: Ward's Automotive Year Book

Table 3 Summary statistics of U.S. light vehicle data by body style, year 2010 & 2013

Door	Body style	Freq	Quantity	Price ($1,000)	Wheel (ins.)	Length (ins.)	Width (ins.)	Size (m^2)	Weigh (lbs.)	Liter	H/w	Torque (Nm.)	MPG	MP$
2-dr	convertible	14	3778	49.00	98.3	172.2	70.9	11.29	3308	2.6	7.30	298	23	6.28
2-dr	hatchback	17	19469	20.83	101.5	165.9	70.2	10.98	2950	2.0	5.73	240	26	7.76
2-dr	p.u.	15	157790	22.36	119.0	205.7	79.4	19.90	4502	3.6	4.55	353	20	5.68
2-dr	suv	2	124906	23.30	95.4	152.8	73.7	13.08	3760	3.6	7.58	353	18	5.54
2-dr	coupe	51	36119	38.75	106.3	176.8	72.6	11.31	3523	3.0	7.87	363	22	6.57
3-dr	van	6	48347	27.91	135.0	224.0	79.4	24.39	5226	4.6	4.23	388	16	4.53
4-dr	cuv	114	52926	31.70	110.0	188.6	74.4	15.05	4013	3.0	5.84	343	22	6.00
4-dr	p.u.	10	133015	32.03	130.0	221.3	79.1	21.44	5049	5.3	6.28	454	17	5.19
4-dr	suv	60	30377	40.84	113.6	194.2	77.2	17.40	5339	5.0	6.20	454	17	4.94
4-dr	van	12	62850	27.29	121.2	202.5	78.7	17.86	4337	3.5	6.09	346	20	5.86
4-dr	coupe	3	3767	72.10	113.2	194.5	74.1	13.18	4158	4.7	9.67	601	20	5.64
4-dr	hatchback	14	22642	28.91	102.4	168.9	69.6	12.11	3208	2.0	4.95	251	28	7.91
4-dr	sedan	140	70676	27.53	109.3	191.4	72.5	13.20	3435	2.5	5.79	251	25	6.96
4-dr	wagon	6	38789	16.61	99.6	156.7	67.4	11.13	2625	1.8	4.88	170	27	8.44
5-dr	hatchback	1	23094	39.99	105.7	177.1	70.4	11.50	3781	1.4	2.22	--	--	25.02
5-dr	van	2	109727	26.29	119.3	200.2	78.2	17.68	4275	3.5	6.22	332	21	5.86

Number of observations: 473

Source: Ward’s Automotive Year Book

Table 4 Summary statistics of U.S. light vehicle data by year

Year	Frequency	Quantity (Mean)	Price (Median)
2010	230	46,442	29.90
2013	243	62,342	30.87

Price in $1,000

Table 5 Summary statistics of U.S. light vehicle data by drive type

Year 2010 & 2013

Drive Type	Frequency	Sales (Mean)	Price (Median)
4WD	10	31,095	62.78
AWD	59	19,470	49.74
FWD	243	70,169	24.02
RWD	151	47,877	38.71

Price in $1,000

Table 6 Summary statistics of U.S. light vehicle data by brand

Year 2010 & 2013

Brand	Frequency	Sales (Mean)	Price (Median)
Smmart	2	7,596	12,491
Fiat	1	35,834	15,632
Scion	8	14,250	16,466
Kia	15	58,162	18,533
Suzuki	5	4,648	19,464
Mazda	13	36,773	20,115
Mini	4	28,037	20,374
Subaru	9	50,726	20,690
Hyundai	18	68,673	21,387
Mitsubishi	9	12,458	21,419
Honda	23	106,809	21,600
Jeep	10	74,971	21,811
Ram	3	182,616	22,018
Dodge	15	62,288	22,620
Ford	29	136,803	23,087
Volkswagen	18	36,740	23,597
Toyota	31	109,102	24,480
Chevrolet	30	109,435	24,715
Chrysler	7	71,415	25,995
Nissan	32	52,600	26,354
Mercury	4	23,290	26,845
GMC	13	60,252	29,359
Buick	8	41,117	29,995
Volvo	13	8,212	33,245
Acura	13	22,622	35,181
Lexus	16	31,426	38,125
Audi	19	13,384	38,625
Lincoln	11	15,229	41,695

Table 6 Summary statistics of U.S. light vehicle data by brand

Year 2010 & 2013 (Ctd.)

Brand	Frequency	Sales (Mean)	Price (Median)
Infiniti	10	10,720	42,810
BMW	19	27,555	45,580
Land Rover	10	8,551	48,100
Porsche	9	6,871	50,124
Mercedes-Benz	25	21,647	64,154
Jaguar	7	4,327	69,337
Number of observations: 473			
Source: Ward's Automotive Year Book			

Table 7 Demand estimation for logit model

	(1)	(2)	(3)	(4)
	OLS	OLS	IV	IV
Instruments used:			All[1]	Two[2]
Dependent Variable: $\ln(s_j)$-$\ln(s_0)$				
Price ($1,000)	-0.039***	-0.037***	-0.052***	-0.112***
	(0.004)	(0.006)	(0.007)	(0.016)
Size (m2)	0.194***	0.146***	0.211***	0.287***
	(0.042)	(0.054)	(0.046)	(0.079)
Displacement (L)	-0.187	-0.238	-0.138	0.121
	(0.115)	(0.128)	(0.127)	(0.209)
Hp/weight (hp/100 lbs)	0.122	0.071	0.210*	0.707***
	(0.064)	(0.063)	(0.062)	(0.177)
MP$	0.152***	0.015	0.164***	0.244***
	(0.045)	(0.040)	(0.047)	(0.073)
Electric	-1.929***			-0.237***
	(0.718)			(1.061)
Drive Style Dummy		Yes		
Body Type Dummy		Yes		
Brand Dummy		Yes		
Constant	-10.94***	-8.77***	-11.49***	-14.82***
	(0.886)	(0.956)	(1.030)	(1.917)
R^2	0.360	0.624		
F-test from first stage			50.06	23.26
Observations	441	437	436	441

Note: Robust standard errors in parenthesis; * significant at 10%; ** significant at 5%; *** significant at 1%.

[1]Instruments used in column (3) include cost variables that are size, wheel base size, weight, displacement, torque, horsepower, estimated MPG, drive type dummies and body style dummies. Instruments also include sum of the cost variable across own-firm products, sum of the cost variable across rival firms' products for each of the first seven cost variables (exclude dummy variables).

[2]Instruments used in column (4) are the sum of size across own-firm models, and sum of displacement across own-firm models.

Table 8 Demand estimation for nested logit model

	(1) OLS	(2) IV	(3) OLS	(4) IV
Instruments used:		Four[1]		Four[1]
Dependent Variable: $\ln(s_j)-\ln(s_0)$				
Price ($1,000)	-0.020***	-0.056***	-0.033***	-0.068***
	(0.003)	(0.009)	(0.003)	(0.008)
$\ln(s_{j/g})$[2]	0.617***	0.176***		
	(0.036)	(0.064)		
$\ln(s_{j'/g'})$[3]			0.514***	0.473***
			(0.043)	(0.088)
Size (m2)	0.088*	0.184***	0.187***	0.232***
	(0.037)	(0.046)	(0.031)	(0.046)
Displacement (L)	-0.111	-0.066	-0.072	0.067
	(0.092)	(0.129)	(0.092)	(0.131)
Hp/weight (hp/100 lbs)	-0.044*	0.252*	0.168**	0.445***
	(0.046)	(0.091)	(0.052)	(0.088)
MP$	0.008	0.127***	0.180***	0.222***
	(0.035)	(0.045)	(0.034)	(0.045)
Electric			-1.143*	-0.393***
			(0.544)	(0.693)
Constant	-5.93***	-10.47***	-10.58***	-12.46***
	(0.738)	(1.126)	(0.684)	(1.070)
R^2	0.698	0.482	0.548	0.439
Observations	410	410	441	441

Note: Robust standard errors in parenthesis; * significant at 10%; ** significant at 5%; *** significant at 1%.
[1]Instruments used for price and within-group share in column (2) and (4) are the sum of size across own-firm models, and sum of displacement across own-firm models, the sum of size across rival-firm models, and body-type dummy.
[2]Group is defined as the class of cars.
[3]Group is defined as the brand of vehicles.

Table 9 Demand estimation using fixed effects

	(1)	(2)	(3)	(4)
	OLS	OLS	IV	IV
Instruments used:			MP$	Three[1]
Dependent Variable: $\ln(s_j)$-$\ln(s_0)$				
Price	-0.039*	-0.025	-0.519*	-0.074
	(0.018)	(0.019)	(0.263)	(0.054)
$\ln(s_{j/g})$				0.947***
				(0.035)
Year Dummy		-0.194*	0.360	-0.300***
		(0.075)	(0.345)	(0.072)
Constant	-7.025***	-7.444***	10.49	-3.08
	(0.676)	(0.686)	(9.549)	(1.917)
F-test from first stage			11.97	
Observations	460	460	451	451

Note: Robust standard errors in parenthesis; * significant at 10%; ** significant at 5%; *** significant at 1%

[1]The instruments are miles per dollar, the sum of miles per dollar across own-firm models and the sum of it across rival firm models.

Table 10 Cross and Own Elasticity of Models

Model	η_j [1]	η_{jk} [1]
Nissan Versa 2013	-1.43	0.0015
Volkswagen Golf 2013	-2.10	0.0003
Ford Fusion 2010	-2.43	0.0046
Toyota RAV4 2013	-2.69	0.0051
Ford Econoline 2010	-3.14	0.0029
BMW 1-series 2010	-3.57	0.0041
Audi A4 2013	-4.21	0.0013
Acura MDX 2013	-4.95	0.0023
Jaguar XF 2010	-6.22	0.0004
Mercedes SL 2013	-11.92	0.0001

[1]Using coefficient from standard logit model

Table 11 Price and elasticity by class

Class	Price	η_j[1]	η_{jk} within group[1]	η_{jk} across group[1]
Two-Seaters	74.043	-4.92	0.115	0.0003
Mini	38.659	-2.46	0.166	0.0004
Subcompact	38.686	-2.57	0.055	0.0013
Compact	21.112	-1.41	0.022	0.0015
Midsize	26.431	-1.78	0.021	0.0024
Large	50.301	-2.74	0.078	0.0016
SUV	34.247	-2.31	0.084	0.0046
Pickup	23.685	-1.53	0.022	0.0017
Van	29.633	-1.92	0.094	0.0013

Note: Averages are sales weighted average
[1]Using coefficient from the first nested logit model

Table 12 Price and elasticity by selected brand

Brand	Price	η_j [1]	η_{jk} within group [1]	η_{jk} across group [1]
Toyota	23.209	-2.65	0.150	0.0028
Chevrolet	26.369	-3.00	0.189	0.0035
Ford	24.218	-2.72	0.202	0.0045
Nissan	23.190	-2.59	0.210	0.0020
Dodge	23.270	-2.60	0.211	0.0011
Honda	23.133	-2.56	0.239	0.0033
Hyundai	20.748	-2.25	0.252	0.0018
Volkswagen	21.668	-2.31	0.303	0.0012
Mazda	20.520	-2.16	0.322	0.0009
Mitsubishi	20.608	-2.16	0.333	0.0002
Subaru	22.124	-2.31	0.360	0.0012
Jeep	24.835	-2.64	0.364	0.0016
GMC	29.579	-3.17	0.402	0.0018
Audi	41.921	-4.63	0.429	0.0006
Suzuki	17.495	-1.64	0.471	0.0001
Acura	38.221	-4.11	0.502	0.0008
Cadillac	43.728	-4.75	0.529	0.0010
Chrysler	26.619	-2.68	0.532	0.0015
Buick	32.879	-3.44	0.534	0.0010
Lincoln	43.306	-4.69	0.539	0.0005
Mercedes-Benz	54.224	-6.00	0.550	0.0017
Lexus	41.307	-4.43	0.555	0.0016
BMW	45.392	-4.90	0.578	0.0017
Infiniti	44.274	-4.58	0.768	0.0005
Mini Copper	21.115	-1.76	0.786	0.0005
Land Rover	62.595	-6.75	0.806	0.0004
Ram	23.582	-1.66	1.188	0.0047
Jaguar	64.808	-6.57	1.260	0.0002

Note: Averages are sales weighted average. Blue is European companies. Green American companies. Orange is Asian companies.
[1]Using coefficient from the first nested logit model.

References

Berry, Steven. 1994. “Estimating Discrete-Choice Models of Product Differentiation.” The RAND Journal of Economics. 25(2). Pp, 242-262.

Berry, Steven., Levinsohn, James. & Pakes, Ariel. 1995. “Automobile Price in Market Equilibrium.” Econometica. 64(4). Pp, 841-890.

Berry, Steven., Levinsohn, James. & Pakes, Ariel. 2004. “Differentiated Products Demand Systems from A Combination of Micro and Macro Data: The New Car Market.” Journal of Political Economy. 112 (1). Pp, 68-105.

Cardell, N. Scott. 1997. “ Variance Components Structures for the Extreme Value and Logistic Distrubutions.” Econometric Theory. 13(2). Pp, 185-213.

Deng, Haiyan., Ma, Alsyson. 2010. “Markets Structure and Pricing Strategy of China’s Automobile Industry.” Journal of Industrial Economics. 58(4). Pp, 818-845.

Goldberg, Pinelopi K., 1995. “ Product Differentiation and Oligopoly in International Markets: the Case of The U.S. Automobile Industry.” Econometrica. 63(4). Pp 891-951.

Murry, Charles. 2014. “Advertising in Vertical Relationships: An Equilibrium Model of the Automobile Industry.” University of Virginia, Charlottesville, VA.

Petrin, Amil. 2002. “Quantifying the Benefits of New Products: The Case of Minivan.” Journal of Political Economy. 110 (4). Pp, 705-729.

The Washington Times. “Government recalls GM stock.” December 26th, 2012. Accessed November 11th, 2014. at <http://www.washingtontimes.com/news/2012/dec/26/government-recalls-gm-stock/>.