A Comprehensive Analysis of Household Vehicle Make/Model/Vintage and Usage Decisions

Sudeshna Sen
Chandra R. Bhat
University of Texas at Austin

Seminar at ETH, Zurich

August 4, 2006

Overview

- Background & Motivation
- Literature Review
- Research Objectives
- Model Structure
- Data
- Sample Characteristics
- Estimation Results
- Application of the model
- Summary and Conclusions

Background & Motivation

- Increasing dependence on automobiles
- Wide-ranging impacts of automobile dependency
  - Household level
  - Community level
  - Regional level

- A widely used indicator of automobile dependency is vehicle holdings and use
  - 92% of US households owned at least one motorized vehicle in 2003 (compared to 80% in the early 1970s)
  - Household VMT has increased 300% between 1997-2001 relative to a population increase of 30% during same period
- Important to examine vehicle holdings and usage
  - Travel demand forecasting
  - Transportation policy analysis

Literature Review

- Key issues related to vehicle holdings and use modeling
  - Dimensions used to characterize vehicle holdings and use
  - Determinants of vehicle holdings and usage decisions
  - Modeling methodology

- Dimensions used to characterize vehicle holdings and usage
  - Number of vehicles owned by the household
  - Vehicle type of the most recently purchased/most driven vehicle of the household
    - Vehicle type defined by body style, vintage, type of fuel and vehicle acquisition type
    - Vehicle type characterized by vehicle makes/models
  - Vehicle transaction type (addition, replacement or disposal of the vehicle)
  - Duration of vehicle ownership by vehicle type between two successive transactions
  - Vehicle type and/or usage of all the vehicles owned by the household

- Determinants of vehicle holdings and usage decisions
  - Household and individual demographic characteristics
  - Vehicle attributes
  - Fuel costs and travel costs
  - Built environment characteristics (land-use and urban form attributes) of the residential neighborhood

- Modeling methodology
  - Standard discrete choice models (multinomial logit, nested logit or mixed logit) are the most commonly used models in predicting vehicle ownership and/or vehicle type
    - These models are not equipped to handle multiple discreteness or satiation effects
    - These model cannot represent a choice situation characterized by joint choice of
      - Multiple alternatives from a set of mutually exclusive alternatives, and
      - Single alternative from a set of mutually exclusive alternatives

Research Objectives

- Examine several dimensions of household vehicle holdings and usage decisions
  - Number of vehicles owned
  - Vehicle body type
  - Vehicle age (i.e., vintage)
  - Vehicle make and model
  - Vehicle usage

Research Objectives

- Incorporate a comprehensive set of determinants of vehicle holdings and usage decisions
  - Household demographics
  - Individual characteristics
  - Vehicle characteristics
  - Built environment characteristics
- Develop a comprehensive econometric model to analyze the many dimensions of vehicle holdings and use that accommodates for
  - Multiple discreteness
  - Satiation effects

Model Structure

- Random Utility Model Structure

$$\tilde{U} = \sum_{k=1}^{K} \left[\exp\left(\max_{l \in N_k} \{W_{lk}\} \right) \right] (m_k + 1)^{\alpha_k}$$

Annual mileage of use for vehicle type *k*

Rate of diminishing marginal utility from using a particular vehicle type *k*

Where, random utility of the make/model *l* of vehicle type *k*,

$$W_{lk} = \beta' x_k + \gamma' z_{lk} + \eta_{kl}$$

$\beta' x_k$ is the overall observed component utility of vehicle type *k*

z_{lk} is an exogenous variable vector influencing the utility of vehicle make/model *l* of vehicle type *k*

γ is a corresponding coefficient vector to be estimated

η_{kl} is an unobserved error component specific to make/model *l* of vehicle type *k*

N_k is the set of makes/models *l* within vehicle type *k*

- Kuhn-Tucker Conditions

$$\left.\begin{array}{l} H_k = H_1 \quad \text{if} \quad m_k^* > 0 \\ H_k < H_1 \quad \text{if} \quad m_k^* = 0 \end{array}\right\} \quad (k = 2,3,\ldots K)$$

$$H_k = \underset{l \in N_k}{Max}\left\{\beta' x_k + \gamma' z_{lk} + \eta_{lk}\right\} + \ln\alpha_k + (\alpha_k - 1)\ln(m_k^* + 1),\ k \geq 1$$

Econometric Model – Basic Structure

$$\eta_{kl} = \lambda_k + \lambda_{kl}$$

λ_k is a common unobserved utility component shared by all vehicle make/model alternatives of vehicle type k

λ_{kl} is an extreme value term distributed identically with scale parameter θ_k

$$H_k = \beta' x_k + \lambda_k + \underset{l \in N_k}{Max}\{\gamma' z_{lk} + \lambda_{lk}\} + \ln\alpha_k + (\alpha_k - 1)\ln(m_k^* + 1)$$

$$= \beta' x_k + \theta_k \ln \sum_{l \in N_k} \exp\left(\frac{\gamma' z_{lk}}{\theta_k}\right) + \ln\alpha_k + (\alpha_k - 1)\ln(m_k^* + 1) + \varepsilon_k \ ,$$

Econometric Model – Basic Structure

Marginal probability that the household uses first Q of the *K* vehicle types ($Q \geq 1$) for annual mileages $m_1^*, m_2^*, \ldots m_Q^*$, using Multiple Discrete-Continuous Extreme Value model (MDCEV derived by Bhat, 2005):

$$P(m_1^*, m_2^*, \ldots m_Q^*, 0, 0, 0, \ldots, 0) = \left[\prod_{k=1}^{Q} r_k\right]\left[\sum_{k=1}^{Q} \frac{1}{r_k}\right]\left[\frac{\prod_{k=1}^{Q} e^{V_k}}{\left(\sum_{k=1}^{K} e^{V_k}\right)^Q}\right](Q-1)! \quad \ldots\ldots\ldots\ldots \quad (1)$$

where,

$$r_k = \left(\frac{1-\alpha_k}{m_k^* + 1}\right)$$

$$V_k = \beta' x_k + \ln \alpha_k + (\alpha_k - 1)\ln(m_k^* + 1), \qquad \text{if } k \notin B, k \geq 1$$

$$= \beta' x_k + \theta_k \ln \sum_{l \in N_k} \exp\left(\frac{\gamma' z_{lk}}{\theta_k}\right) + \ln \alpha_k + (\alpha_k - 1)\ln(m_k^* + 1), \quad \text{if } k \in B, k \geq 1$$

Econometric Model – Basic Structure

Conditional probability that vehicle make/model *l* will be used for an annual mileage $m_k^*(l \in N_k, k \in B)$, given that $m_k^* > 0$

$$P(l \mid m_k^* > 0;\ l \in N_k) = \frac{\exp\left(\frac{\gamma' z_{lk}}{\theta_k}\right)}{\sum_{g \in N_k} \exp\left(\frac{\gamma' z_{gk}}{\theta_k}\right)} \quad \text{..... (2)}$$

Unconditional probability that household uses vehicle make/model *a* of vehicle type 1 for annual mileage, m_{1a}^* , make/model *b* of vehicle type 2 for m_{2b}^* ……. Make/model *s* for vehicle type *S* for m_{sS}^* is

$$P(m_{1a}^*, m_{2b}^*, m_{3c}^*, \ldots m_{Ss}^*, m_{S+1}^*, m_{S+2}^* \ldots m_Q^*, 0, 0 \ldots 0)$$

$$= P(m_1^*, m_2^*, \ldots m_Q^*, 0, 0, \ldots 0) \times P(a \mid m_1^* > 0) \times P(b \mid m_2^* > 0) \ldots P(s \mid m_S^* > 0)$$

Econometric Model – Mixed MDCEV-MNL model

Unconditional probability of vehicle holdings and usage:

$$P(m_{1a}^{*}, m_{2b}^{*}, m_{3c}^{*}, \ldots m_{Qq}^{*}, 0, 0, 0, \ldots 0)$$

$$= \int_{\beta} \int_{\gamma} \left\{ P(m_{1}^{*}, m_{2}^{*}, \ldots m_{Q}^{*}, 0, 0, \ldots 0) \times P(a \mid m_{1}^{*} > 0) \times P(b \mid m_{2}^{*} > 0) \right.$$

$$\left. \ldots \times P(q \mid m_{Q}^{*} > 0) \mid (\beta, \gamma) \right\} \phi(\beta) \phi(\gamma) \, d\beta \, d\gamma$$

Data

- 2000 San Francisco Bay Area travel survey (BATS)
  - Designed and administered by MORPACE International Inc.
  - 2-day survey of 15000 households
  - Information on vehicle fleet mix of households, individual and household socio-demographics, individual characteristics and activity episodes
- Data on vehicle make/model attributes from secondary data sources
  - Consumer Guides
  - EPA Fuel Economy Guide
- Land use/Demographic coverage data from MTC of San Francisco Bay area
- GIS layer of bicycle facilities from MTC of San Francisco Bay area
- Census 2000 Tiger files

Sample Characteristics

- Final sample: 8107 households
- 10 motorized vehicle types
  - Coupe
  - Mini/Subcompact Sedan
  - Compact Sedan
  - Mid-size Sedan
  - Large Sedan
  - Hatchback/Station Wagon
  - Sports Utility Vehicle (SUV)
  - Pickup Truck
  - Minivan
  - Van
- 2 vintages considered for each motorized vehicle type
  - New vehicles (age of the vehicle less than or equal than 5 years)
  - Old Vehicles (age of the vehicle is more than 5 years)
- Twenty-one vehicle types/vintages studied including
  - 20 motorized vehicle type/vintages
  - Non-motorized form of transportation

Classification of Vehicle type/vintage

Distribution of Vehicles

Number of vehicles owned by the household	Total No. of households	% of households
1	4459	55%
2	2918	36%
3	644	8%
4 or more	86	1%

Descriptive Statistics of Vehicle Type/Vintage Holdings

Vehicle type/vintage	Total number (%) of households owning	Annual Mileage	No. of households who own (%)	
			Only Vehicle type/vintage	Vehicle type/vintage and other Vehicle type/vintages
New Coupe	389 (5%)	7763	132 (34%)	257 (66%)
Old Coupe	1024 (13%)	7766	374 (37%)	650 (63%)
New Subcompact Sedan	292 (4%)	7838	127 (43%)	165 (57%)
Old Subcompact Sedan	513 (6%)	9570	238 (46%)	275 (54%)
New Compact Sedan	767 (9%)	8321	342 (45%)	425 (55%)
Old Compact Sedan	1175 (14%)	9614	495 (42%)	680 (58%)
New Midsize Sedan	987 (12%)	7688	361 (37%)	626 (63%)
Old Midsize Sedan	1543 (19%)	9342	636 (41%)	907 (59%)
New Large Sedan	250 (3%)	7418	71 (28%)	179 (72%)
Old Large Sedan	377 (5%)	8339	151 (40%)	226 (60%)
New Station Wagon	242 (3%)	7869	80 (33%)	162 (67%)
Old Station Wagon	728 (9%)	8248	254 (35%)	474 (65%)
New SUV	707 (9%)	8920	245 (35%)	462 (65%)
Old SUV	711 (9%)	9813	213 (30%)	498 (70%)
New Pickup Truck	578 (7%)	8887	153 (26%)	425 (74%)
Old Pickup Truck	1198 (15%)	8679	301 (25%)	897 (75%)
New Minivan	459 (6%)	9156	115 (25%)	344 (75%)
Old Minivan	480 (6%)	9890	130 (27%)	350 (73%)
New Van	39 (1%)	10640	8 (21%)	31 (79%)
Old Van	122 (2%)	8203	33 (27%)	89 (73%)
Non-Motorized form of transportation	201 (3%)	2695	-	201 (100%)

Empirical Results

- Variables considered
  - Household socio-demographics
    - Household income, presence of children in the household, presence of a senior adult in the household, household size and number of employed people in the household
  - Household location attributes
    - Area type variables (central business district, urban zone, suburban zone and rural zone), residential density and employment density variables
  - Built environment characteristics of the residential neighborhood
    - Percentages and absolute values of acreage in residential, commercial/industrial, and other land-use categories; fractions and number of single family and multi-family dwelling units, and fractions and number of households living in single family and multi-family dwelling units, bikeway density, street block density, highway density
  - Characteristics of the household head
    - Age (classified into less than 30 years of age, 31 to 45 years of age and greater than 45 years of age), gender and ethnicity (primarily, Caucasian, African-American, Hispanic, Asian and Other)
  - Vehicle Characteristics
    - Purchase price, fuel cost, seating capacity, luggage volume, engine size, number of cylinders, front headroom space, front legroom space, rear headroom space, rear legroom space, standard payload capacity (for pickup trucks only), wheelbase, length, height, width, horse power, vehicle weight, type of fuel used, amount of greenhouse gas emissions (tons/year), types of drive wheels, type of vehicle make

- MDCEV model – Effects of Household Demographics
  - Medium income (35-90K) and high income (>90K) households have a high baseline preference for new SUVs as compared to low-income households and a low preference for old vans
  - High income households have a lower baseline preference for old vehicles compared to low/middle income households
  - High income households less likely to undertake activities using non-motorized forms of transportation
  - Households with very small children (less than 4 years of age) are more likely to use compact sedans, mid-size sedans, and SUVs than other households
  - Households with kids between 5 and 15 years of age have a high baseline preference for minivans than other households
  - Households with senior adults (greater than 65 years) are more likely to use compact, mid-size, and large sedans relative to coupes and subcompact sedans
  - As the size of the household increases, the household is more likely to use mid-size sedans, large sedans, station wagons, SUVs, pickup trucks, minivans and vans
  - Household with more number of employed members have a high baseline preference for new vehicle types such as subcompact sedans and compact sedans while a low baseline preference for large sedans and minivans

- MDCEV model – Effects of Household Location Characteristics
  - Households residing in the suburban zones are less likely to own and use old vehicles relative to households in urban zones
  - Households residing in the suburban and rural zones are more likely to own and use pickup trucks relative to urban households

- MDCEV model – Effects of Built Environment Characteristics of the Residential Neighborhood
  - Households located in highly residential/commercial areas are less likely to prefer large vehicle types such as pickup trucks and vans, irrespective of the age of the vehicle
  - Households located in a neighborhood with high bike lane density have a high baseline preference for non-motorized modes of transportation
  - Households located in a neighborhood with high street block density are more likely to prefer smaller vehicle types (such as subcompact and compact sedans), and older vehicles, relative to new vehicles

- MDCEV model – Characteristics of the Household Head
  - Older households (*i.e.*, households whose heads are greater than 30 years) are generally more likely to own vehicles of an older vintage compared to younger households (*i.e.*, households whose heads are less than or equal to 30 years of age)
  - Older households are more likely to own minivans and old vans, and travel by non-motorized forms of transportation
  - Households have higher baseline preference for older and larger vehicles if the male is the oldest member (or only adult) in the household relative to households with the female being the oldest member (or only adult)
  - Asians more likely to own sedans and new minivans, and less likely to own pickup trucks, than other races.

- MDCEV model – Random Error Components/Coefficients
  - Households preferring old coupes due to unobserved factors also prefer new coupes
  - Intangible unobserved factors that affect utilities of all old vehicles

MNL model for Vehicle Make/Model Choice

Variable	Parameter	t-stat
Purchase Price (in $)/Income (in $/yr) [x 10]		
Mean Effect	- 0.173	- 5.71
Standard Deviation	- 0.064	- 4.44
Fuel Cost (in $/yr) /Income (in $/yr) [x 10]	- 0.003	- 1.61
Seat Capacity * Household Size less than equal to 2 dummy variable	- 0.075	- 5.11
Luggage Volume (in 10s of cubic feet)	0.023	3.54
Standard Payload Capacity (for Pickup Trucks only) (in 1000 lbs)	0.196	5.13
Horsepower (in HP) /Vehicle Weight (in lbs) [in 10s]	1.102	4.89
Engine Size (in liters)	- 0.045	- 2.42
Dummy variable for All-Wheel-Drive (base: rear-wheel-drive)	- 0.214	- 3.81
Dummy Variable for Vehicle Make - Chevy	- 0.149	- 1.25
Dummy Variable for Vehicle Make - Ford	0.716	5.37
Dummy Variable for Vehicle Make - Honda	1.444	5.37
Dummy Variable for Vehicle Make - Toyota	0.752	5.29
Dummy Variable for Vehicle Make - Cadillac	0.880	4.36
Dummy Variable for Vehicle Make - Volkswagen	0.374	2.55
Dummy Variable for Vehicle Make - Dodge	0.699	4.96
Amount of Greenhouse Gas Emissions (in 10s of tons/yr)	- 0.429	- 2.71
Dummy variable for Premium Fuel (base: regular fuel)	- 0.552	- 5.01

Satiation Effects

- All the satiation parameters are very significantly different from 1
- Middle and High income households are more likely to get satiated with the increasing use of any vehicle type/vintage compared to low income households
- Low income households are least likely to get satiated with the increasing use of old subcompact sedans, new and old compact sedans, and old midsize sedans
- Satiation effect is highest for non-motorized mode of transportation compared to all vehicle type/vintage categories

Logsum Parameters

- Indicate the presence of common unobserved attributes that affect the utilities of all makes/models corresponding to old SUV, old minivan, new minivan, old van, and new van vehicle type/vintage categories

Application of the Model

Vehicle Type	Impact of a 25% increase in bike lane density		Impact of a 25% increase in street block density		Impact of a 25% increase in fuel cost	
	% change in holdings of vehicle type	% change in overall use of vehicle type	% change in holdings of vehicle type	% change in overall use of vehicle type	% change in holdings of vehicle type	% change in overall use of vehicle type
Compact Car	-	-2.2%	8.5%	3.4%	1.3%	-0.9%
Midsize and Large Sedan	-2.2%	-2.1%	-	-0.8%	-	-0.6%
SUV	-0.6%	-0.4%	-	-	-	-
Pickup Truck	-1.4%	-0.4%	-2.1%	-1.7%	-5.7%	-2.3%
Minivan and Van	-	-0.7%	-	-0.6%	-2.6%	-
Non-motorized modes of transportation	7.4%	13.9%	-4.0%	-3.3%	1.5%	0.8%

Summary and Conclusions

- Comprehensive analysis framework
  - Examined several dimensions used to characterize vehicle holdings and use
  - Incorporated a comprehensive set of explanatory variables
- Comprehensive modeling framework
  - Estimated a comprehensive joint MDCEV-MNL model
    - Accommodates for heteroscedasticity and/or error correlation
    - Incorporates random coefficients
    - Accommodates for multiple discreteness and satiation effects
- Data drawn from 2000 San Francisco Bay Survey
- Empirical results provide important insights into the determinants of vehicle holdings and usage decisions of households
- Application of the model gives useful insights on the impact of different transportation policies

Thank You!