A Vehicle Ownership (Car Shedding) Model as a Pre-Step of Travel Demand Modeling

Sadegh Sabouri [1](Sadegh.Sabouri@utah.edu), Guang Tian[2], Reid Ewing[1], Keunhyun Park [3]

Poster ID # 20-02556

[1]Department of City and Metropolitan Planning, University of Utah, [2] Department of Planning and Urban Studies, University of New Orleans,[3] Department of Landscape Architecture and Environmental Planning, Utah State University

Key Points

* Vehicle ownership models are used by policy makers to identify factors that affect vehicle miles traveled, and therefore address problems related to energy consumption, air pollution, and traffic congestion.
* While not always treated as such, vehicle ownership forecasting is a step in conventional travel demand forecasting process, and is also always part of activity-based modeling.
* The most critical limitation of the vehicle ownership models, especially in the conventional process, is that they are often related mainly to sociodemographic variables, not so much to built environmetnal variables.
* We pool regional household travel survey data from 32 diverse regions of the United States and generate consistent measures for all regions.
* Next, we test both count regression models (i.e., quasi-Poisson and Poisson) and the more commonly used multinomial logit (MNL) model to estimate vehicle ownership.
* The study results provide practical implications for state and local planning and transportation agencies with better accuracy and generalizability.

State-of-the-Practice in Vehicle Ownership Modeling

To understand the gap between academic research and practical implementation, we conducted a survey of 25 Metropolitan Planning Organizations (MPOs) in the U.S. in mid-2018 focusing mostly on large regions since we assume that their MPOs are leaders in using new travel modeling techniques.

MPO Name	Major City/ State	Population (2010)	Is Vehicle Ownership Modeled?	Method
Brunswick MPO	Brunswick, GA	79,626	No	-
Roanoke Valley MPO	Roanoke, VA	227,507	No	-
Lincoln Area MPO	Lincoln, NE	285,407	No	-
North Front Range MPO	Fort Collins, CO	433,178	No	-
Chattanooga-Hamilton County/North Georgia Transportation Planning Organization	Chattanooga, TN, GA	436,669	Yes	Multinomial Logit Model (MNL)
Augusta Regional Transportation Study	Augusta, GA, SC	440,134	No	-
Des Moines Area MPO	Urbandale, IA	475,855	No	-
Stanislaus COG	Modesto, CA	514,453	No	-
Community Planning Association of Southwest Idaho	Meridian, IA	550,359	No	-
Association of Monterey Bay Area Governments	Marina, CA	732,667	No	-
Capital District Transportation Committee	Albany, NY	823,239	No	-
Fresno COG	Fresno, CA	930,885	Yes	Multinomial logit model
Memphis Urban Area MPO	Memphis, TN, MS	1,077,697	No	-
Wasatch Front Regional Council	Salt Lake City, UT	1,561,348	Yes	Multinomial logit model
METROPLAN Orlando	Orlando, FL	1,837,385	No	-
Mid-America Regional Council	Kansas City, MO, KS	1,895,535	Yes	Series of binary logit models
Ohio-Kentucky-Indiana Regional COG	Cincinnati, OH, KY, IN	1,981,230	Yes	Nested Logit Model
East-West Gateway COG	St. Louis, MO, IL	2,571,253	Yes	Multinomial logit model
Boston Region MPO	Boston, MA	3,159,512	Yes	Multinomial logit model
Southeast Michigan COG	Detroit, MI	4,703,593	No	No in the current model, but MNL in the ABM
The National Capital Region TPB	Washington, DC, MD, VA	5,068,540	Yes	Multinomial logit model
Houston-Galveston Area Council	Houston, TX	5,892,002	No	No in the current model, but MNL in the ABM
North Central Texas COG	Arlington, TX	6,417,630	No	-
North Jersey Transportation Planning Authority	Newark, NJ	6,579,801	No	-
Chicago Metropolitan Agency for Planning	Chicago, IL	8,444,660	Yes	Multinomial logit model

Methodology

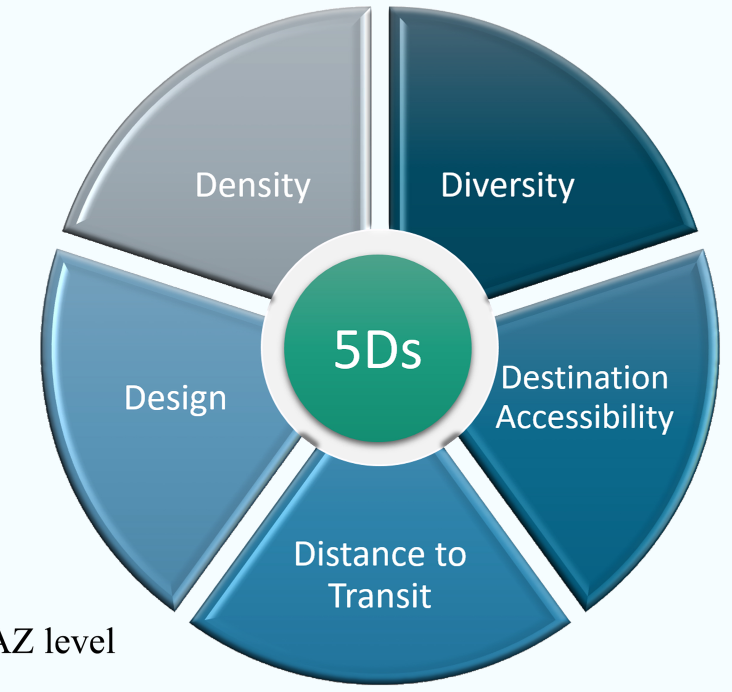
*** Data:**
* Household travel survey in 32 regions

*** Dependent Variable:**
* Number of vehicles owned by household

*** Models:**
* Count Regression Models:
  * Multi-level Poisson
  * Multi-level Quasi-Poisson
* Discrete Choice Models:
  * Multi-level Ordered Logit
  * Multi-level Multinomial Logit

All models are fixed at region level and random at TAZ level

*** Model Evaluation:**
* 5-fold cross-validation using Root Mean Square Error (RMSE)



Nesting Structure of The Data + Variables

1.92 Average Vehicle Ownership

91,979 Households

14 Predictors

883,695 Trips

Dependent Variables

vehicle (count models)	vehicle_cat (logit models)
Actual number of vehicles owned by household	Vehicles owned by households as 0, 1, 2, and 3+

Level 1: Household

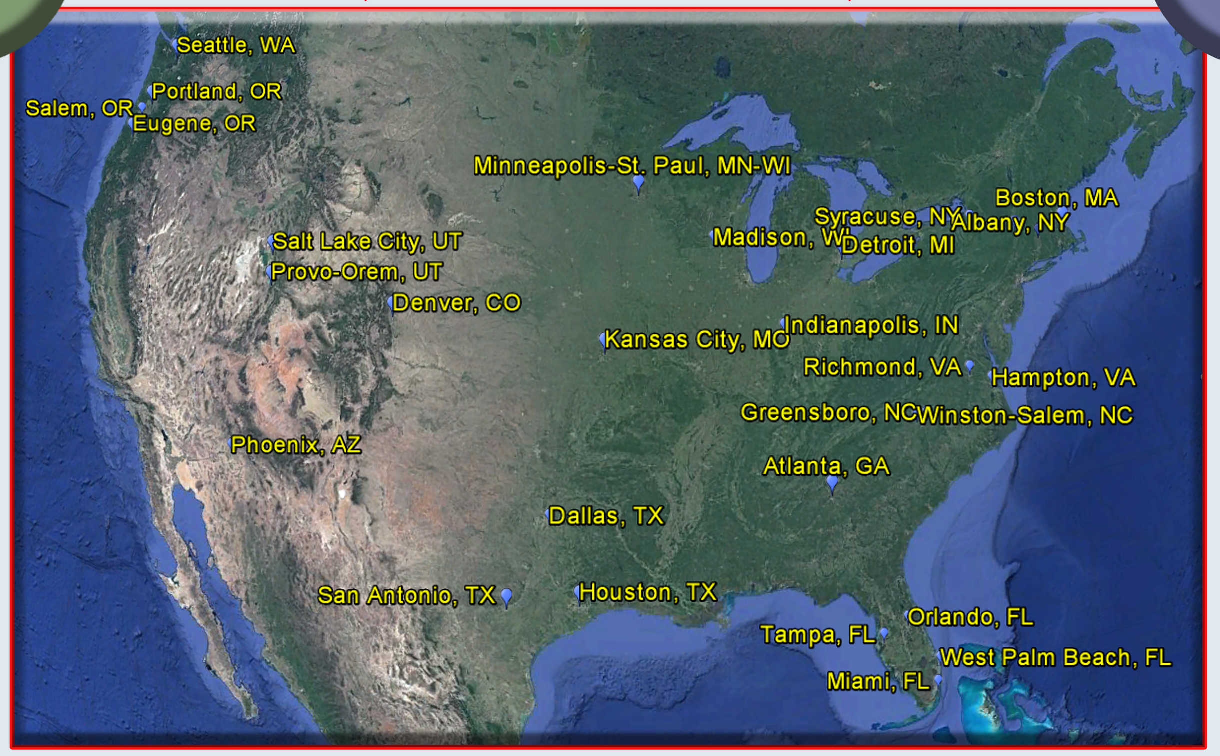
hhsize_cat	employed_cat	lnincome
Household size of 1, 2, 3, 4, and 5+	Number of employed persons in household: 0, 1, 2, and 3+	Natural log of household income

Level 2: Traffic Analysis Zone

actden	jobpop	entropy	intden	pct4way	pctemp10a	pctemp20a	pctemp30a	pctemp30t	transitden
activity density within TAZ (pop + emp per square mile in 1000s)	job-population balance within TAZ	land use entropy	intersection density within TAZ	percentage of 4-way intersections within TAZ	percentage of regional employment within 10 minutes by auto	percentage of regional employment within 20 minutes by auto	percentage of regional employment within 30 minutes by auto	percentage of regional employment within 30 minutes by transit	transit stop density

Level 3: Region

region
a set of regional dummy variables



Results of count regression and discrete choice models

	Multi-level Quasi-Poisson		Multi-level Poisson		Multi-level Ordered Logit		Multi-level Multinomial Logit (Ref: vehicle = 0)					
							Vehicle = 1		Vehicle = 2		Vehicle = 3	
Variable[1]	Coeff	t ratio	Coeff	t ratio	Coeff	t ratio	Coeff	t ratio	Coeff	t ratio	Coeff	t ratio
(Intercept)	-1.26948	-41.21***	-1.26200	-24.29***	-3.94745	-30.22***	-6.20828	-16.59***	-6.20828	-16.59***	-18.9133	-44.84***
hhsize_cat2[2]	0.43787	85.93***[4]	0.43970	50.21***	2.07109	94.55***	0.23988	4.72***	2.71152	48.59***	2.74445	39.35***
hhsize_cat3	0.49698	82.26***	0.49870	48.13***	2.50299	93.78***	0.15161	2.03*	2.61276	33.12***	3.31191	36.95***
hhsize_cat4	0.48770	77.90***	0.48870	45.55***	2.36481	79.22***	0.0531	0.51	3.05236	28.92***	3.31464	28.95***
hhsize_cat5	0.53972	78.10***	0.54140	45.79***	2.56507	74.25***	0.03781	0.31	3.13177	25.44***	3.58724	27.32***
employed_cat0[3]	-0.43852	-59.69***	-0.44080	-35.17***	-2.77728	-61.65***	-0.38789	-1.57	-0.94938	-3.97***	-3.46337	-14.42***
employed_cat1	-0.37146	-59.44***	-0.37320	-35.07***	-2.49424	-59.34***	0.09486	0.39	-0.31776	-1.34	-2.66589	-11.25***
employed_cat2	-0.29222	-49.01***	-0.29390	-28.95***	-1.99007	-48.07***	-0.0452	-0.18	0.30747	1.27	-1.79052	-7.42***
lnincome	0.18714	75.22***	0.18600	43.97***	0.95166	89.77***	0.96673	43.89***	1.86261	75.26***	2.24246	79.36***
actden	-0.00467	-19.17***	-0.00486	-11.83***	-0.01426	-29.78***	-0.0064	-9.87***	-0.02943	-27.02***	-0.03748	-18.65***
entropy	-0.07059	-10.02***	-0.06747	-5.90***	-0.41287	-13.18***	-0.67677	-8.85***	-0.92502	-11.29***	-1.23155	-13.86***
intden	-0.00053	-16.38***	-0.00053	-9.96***	-0.00261	-22.06***	-0.00152	-7.1***	-0.00308	-12.56***	-0.00573	-18.55***
pct4way	-0.00039	-3.55***	-0.00038	-2.16*	-0.0033	-7.45***	-0.00575	-6.27***	-0.00953	-9.31***	-0.0086	-7.46***
pctemp10a	-0.00065	-2.83***	-0.00061	-1.66	-0.00561	-5.36***	-0.00954	-3.41***	-0.01426	-4.79***	-0.01508	-4.76***
pctemp30a	-0.00103	-9.69***	-0.00099	-5.76***	-0.00543	-11.04***	-0.003	-1.93*	-0.00507	-3.07***	-0.01042	-6.1***
pctemp30t	-0.00072	-5.63***	-0.00066	-3.25**	-0.00481	-7.92***	-0.02234	-9.8***	-0.02837	-12.01***	-0.03066	-12.64***
transitden	-0.00082	-9.90***	-0.00082	-6.04***	-0.00309	-14.2***	-0.00226	-7.46***	-0.00522	-11.02***	-0.00657	-9.71***
Threshold parameters for probabilities												
Mu(01)					3.90428	170.87***						
Mu(02)					6.93591	257.65***						
Var(cons)-TAZ	0.26168		0.00004		0.37583		0.30164		0.12062		0.31331	
Model Evaluation												
Log Likelihood (LL(β))	N/A		-107289		-68393		-66107					
AIC/N	N/A		2.733		1.743		1.443					
McFadden R2	N/A		0.154		0.2826		0.3065					
Cor(Veh, Pred)[5]	0.6769		0.6536		0.6527		0.6536					
Cor(Veh, IntPred)[6]	0.6230		0.6008		0.6039		0.6065					
RMSE	0.8147		0.8347		0.9083		0.8964					

Notes:
1 Fixed-effect variables for regions were included in the models, but are not shown here.
2 Household size equal to 1 is the reference category
3 Number of employed persons equal to 3 is the reference category.
4 *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05
5,6 Correlation between Veh = actual number of vehicles and Pred = predicted values, and IntPred = rounded predicted values

Model Evaluation

The Results of 5-Fold Cross-Validation

Fold	Quasi-Poisson	Poisson	Ordered logit	Multinomial Logit
1	0.820	0.852	0.922	0.918
2	0.824	0.847	0.922	0.909
3	0.826	0.833	0.910	0.902
4	0.840	0.832	0.941	0.930
5	0.813	0.839	0.933	0.917
Overall RMSE	0.825	0.841	0.926	0.915

The Results of Our Multi-Level Poisson Model and WFRC's Multinomial Logit Model

Goodness-of-Fit Measures	Best-Fit Model	WFRC Model
RMSE for all TAZs	0.5274	1.1431
Correlation (predicted vs. actual) for all TAZs	0.6557	0.0276
RMSE for TAZs with 10 or more households	0.2293	0.9243
Correlation (predicted vs. actual) for TAZs with 10+ households	0.8506	0.0882

Conclusions

* Household vehicle ownership has positive relationships with socio-demographic variables and negative relationships with several built environmental variables.
* Although the elasticities of built environmental variables are smaller than the elasticities of the socioeconomic variables (specifically household income), all are highly significant.
* For urban planning and design practices, this study suggests that car shedding occurs as built environments become more dense, mixed, connected, and transit-served.
* This finding has important implications in the policy and planning practice, where decision makers seek solutions to deal with VMT, emissions, obesity, and other health and environmental concerns.
* Lastly, based on the results of this study, we would recommend using count models (quasi-Poisson and Poisson) over discrete or categorical models (ordered logit and multinomial logit).