Vehicle Ownership in the New York Metropolitan Area

Robert Checco

SUNY Farmingdale

1. Introduction

This paper addresses the following question: Does the type of residential dwelling influence the household vehicle ownership decision in the New York Metropolitan Area? In this paper, the type of dwelling is classified as either a single-family residence or a multi-family residence. A single-family dwelling includes detached homes that are zoned to be inhabited by a single-household. Multi-family housing can be thought of as an apartment building, where there are multiple units that are rented out to different families. My hypothesis is that after controlling for socioeconomic and spatial variables, households who live in single-family dwellings are less likely to own fewer vehicles such as zero or one, and more likely to own multiple vehicles such as two or three-plus compared to a household in multi-family housing.

Understanding the factors that influence vehicle ownership are important for at least three reasons. First, while the ownership of automobiles does promote a high standard of living for society, there are negative externalities associated with higher levels of car ownership. Three externalities of car ownership are: environmental pollution, costs from auto collisions, and congestion of roadways. In light of these social costs, public policy can be used to induce modal switch, that is, getting citizens to switch from driving cars to an alternative transportation mode. The policy prescriptions implemented by lawmakers should be made in accordance with the most important factors related to vehicle ownership. Second, the number of vehicles owned by society is correlated with more vehicle miles traveled (VMT), and understanding the factors leading to car ownership can be used in models to forecast transportation demand along certain roadways. Third, mass transit projects are notoriously expensive to build, and transportation mode choice models are enhanced by having the model explain vehicle ownership rates from a disaggregate perspective.

This paper contributes to the existing literature by making use of a more recent dataset. The most recent study of vehicle ownership in the New York City area was published in 2009 and used the previous version of the dataset used in this paper. A more recent dataset will allow transportation researchers and officials to develop more robust models for estimating transportation demand. Transportation demand has became a hot topic in 2019 resulting from congestion pricing for Manhattan being passed into law.

This paper finds that households who lives in single-family housing are more likely to own more vehicles and less likely to own fewer vehicles. This is consistent with other research that has been done, and it aligns with the intuitive understanding of the topic.

2. Literature Review

Prior to the 1970's, a majority of transportation demand modeling used aggregate models to estimate transportation demand. (Mcfadden, 1973) developed a disaggregate model based on the probability of a household choosing an alternative. His modeling approach used the theory of random utility and linked it in a way to make sense of discrete choice problems.

Since the early 1990s, there has been a revival in the research of car ownership models. One reason for this is because of the more available computing power allows larger scale models to be connected. For example, a travel demand model may include a model for one of the variables to make it more robust. Following this approach may allow models to make more accurate estimates. Second, humans have become more aware of the environment and the dangers of global warming. Increased levels of car ownership would be likely be associated with increased car consumption that can lead to more air pollution.

Discrete choice models are used in the car ownership choice literature. This specification is used due to the dependent variable, the number of cars owned (per household), is a discrete choice. There are a variety of choice models that have been used to estimate the level of car ownership. These include, the multinomial varieties of logit and probit as well as the ordered varieties of logit and probit. A review of the literature reveals that researchers were split between using the ordered and unordered response mechanisms. This debate was researched by running both models on four different datasets and comparing the results. The result was that unordered class of models should be used for car ownership modeling. In the same paper, high income households were found to have higher levels of auto ownership. The elasticity effects of annual income for the Boston data are negative for levels of zero and one cars owned, and positive for levels of two, three, and four cars (Bhat and Pulugurta 1998).

In Hamilton, Canada, the authors used a Multinomial Logit Model to examine the influence of family structure and socio-economic characteristics on the number of cars owned by the household (Potoglou and Kanaroglou, 2008). The paper included a dummy variable for the single-family house attribute to serve as a proxy for parking-space availability nearby the home. The effect of single-family house attribute was found to be statistically significant, suggesting that for increasing levels of vehicle ownership, having more available parking from living in a single-family house increased utility. Medium range income levels were found to be statistically significant for a household’s decision to own one vehicle and high income was found to have a positive effect for households’ decision to own two vehicles at the 10% significance level. In general, the paper suggests that given a rise in income, there is a positive influence in the likelihood of households owning cars.

The car ownership choice problem is inherently related to the residential location choice problem. Where members of a household decide to reside may reflect their preference for car ownership. For example, a household who elects to live in the suburbs may choose to do so because they enjoy driving their car. That is, the “cost” of operating a vehicle may be lower than a household who elects to not own any vehicles and live in a dense urban center. Some models combine the choice of car ownership with residential location. One instance of this approach is seen using data from Mexico City and it was found that a 10% increase in household income would lead to a 8.4% increase of probability households owning two or more vehicles (Guerra 2015). The model used in this analysis was a mixed logit model to estimate the joint choice made by the household. When members of the household had a college degree, it was found to disrupt the strong relationship between income and car ownership.

Joint choice modeling can extend to multiple decisions, such as trying to estimate a model of the choices of car ownership, commute mode and residential location. The data used for the analysis was survey data from 1997 to 1998 from New York City. To model this joint choice, the researcher used seven multinomial logit models. The results of the research found that population density has a significant effect on car ownership, but warns that this result may be due to the lack of a variable for parking cost. An increase in income was found to decrease the probability of owning zero or one car, and increase the probability of owning two-plus cars (Salon 2009).

There has also been work using Structured Logit Models to estimate the joint decision of car ownership and mode choice. In one instance, data from the San Francisco Bay Area from 1975 was used to learn more about the commuting and auto ownership of commuters on the Bay Area Rapid Transit rail service (BART). The number of persons in the household was found to influence the household's decision for cars. Specifically, as the population of the household increased there was a higher probability for owning three cars to two, two cars to one, and one car to none. Another way to describe this phenomenon is that the estimated coefficients for household population is increasing with higher levels of car ownership (Train, 1980).

My research fits in with the larger scope of the literature because I am using newly available data. I have not seen any other paper use this data for vehicle ownership choice. Also, this topic of car ownership is closely related to transportation demand and that has been a hot topic in the NYC Metro area over congestion pricing plans.

3. Data

The data used for this analysis was solely obtained from the 2010/2011 Regional Household Travel Survey (RHTS). The survey was sponsored by two of the federally sanctioned Metropolitan Planning Organizations (MPOs) in the New York/New Jersey/Connecticut metropolitan area. The two MPOs are the New York Metropolitan Transportation Council (NYMTC) and the North Jersey Transportation Planning Authority (NJTPA). The survey was conducted from September 2010 through November 2011 by NuStats based in Austin, Texas. NuStats is a full service survey research consultancy specialized in large-scale social research studies. At various stages of data collection NuStats was assisted by GeoStats and Parsons Brinckerhoff.

The data was collected to be used for transportation planning, research, analysis and policy. The predecessor to the RHTS was the 1997/98 Regional-Travel Household Interview Survey (RT- HIS). At the time of the RHTS, the previous travel survey was over 12 years old and contained information that no longer reflected the travel profile of the region. The data collection method for the RHTS relied on the willingness of area residents to complete diary records of their daily travel over a 24-hour period. To generate a random sample of households, a phone interview was conducted to inform the household of the survey, its purpose and the requirement to fulfill a travel diary. The data diary was created from a single day, involves observations of many households, and examines many different characteristics from each household. The data is cross sectional data and the unit of observation for this analysis is the household. In total 31,156 households were recruited to participate in the survey and of these, 18,965 households completed travel diaries. All households were located in the 28-counties that constitute the New York/New Jersey/Connecticut metropolitan area. Later after data cleaning, the number of observations dropped to 17,537.

The following tables display the summary statistics of the dependent and independent variable. Only one of the independent variables is continuous, and all of the rest are categorical. For the variables that are categorical, instead of including the mean and standard error, I included a frequency table to understand the quantity spread throughout the observations.

Summary Statistics

HHVEH	Freq	Percent	Cumulative
0	3526	20.11	20.11

1	5413	30.87	50.97
2	5726	32.65	83.62
3+	2872	16.38	100
Total	17537	100	

Income Range	Freq	Percent	Cumulative
0-30k	3,630	20.7	20.7
30-75k	5,514	31.44	52.14
75-100k	2,320	13.23	65.37
100k+	6,073	34.63	100
Total	17,537	100	

NYC Dummy	Freq	Percent	Cumulative
Not NYC (0)	12,145	69.25	69.25
NYC (1)	5,392	30.75	100
Total	17,537	100	

HHLIC	Freq	Percent	Cumulative
0	1,685	9.61	9.61
1	5,977	34.08	43.69
2+	9,875	56.31	100
Total	17,537	100	

HHWRK	Freq	Percent	Cumulative
0	3,710	21.16	21.16
1	7,336	41.83	62.99
2+	6,491	37.01	100
Total	17,537	100	

Residence Type	Freq	Percent	Cumulative
Multi-Family (0)	7265	41.43	41.43
Single-Family (1)	10,272	58.57	100
Total	17537	100	

Variable	Observations	Mean	Std. Err.	Min	Max
HHSIZ	17537	2.31	1.25	1	10

4. Model/Econometric Specification

The model that is used in this paper is the multinomial logit model (ML). As mentioned in the literature review, the multinomial logit model is the suggested model for car ownership choice research (Bhat and Pulugurta, 1998). A brief explanation of the ML is as follows. The ML model is based on random utility modeling, and for each level of car ownership, the ML estimates a utility function. This utility function is comprised of observed utility, as well as unobserved utility in the form of the error term. The model is estimated with maximum likelihood estimation. Marginal effects permit interpretation of a change in probabilities with respect to a change in one the independent variable. For further reading about discrete choice theory, I suggest (Ben-Akiva and Lerman, 1985, Train 2003).

The first step in the specification of a discrete choice model involves describing the components of the model. The choice set in a discrete choice model outlines the household's choices among a set of alternatives. In this paper the four alternatives are choosing to own zero, one, two, or three-plus vehicles. In order to be correctly specified, the set of alternatives, commonly referred to as the choice set must exhibit three characteristics: mutual exclusivity, exhaustivity, and the number of alternatives must be finite. (Train 2009). The alternatives in the choice set are mutually exclusive because if a household chooses to own a particular level of autos, it precludes them from owning a different level of autos. For example, households can not simultaneously own zero vehicles and two vehicles. The choice of owning zero automobiles explicitly implies that the household is unable to own any other level of autos. The choice set is exhaustive because all possible alternatives are present in the choice set. The fewest number of cars that a household can own are zero vehicles, and the most vehicles that a household can own are three-plus. The choice set in this paper spans the entire set of all possible levels of auto ownership. Third, the number of alternatives is finite, because the choice set is composed of 4 different alternatives.

The dependent variable of interest is the number of automobiles that are owned by the household, denoted HHVEH. This variable represents the four possible levels of car ownership that a household can choose to have. Since each choice is a distinct outcome, the response can not be a continuous variable. Also, the number of autos owned by the household is a computed variable. The way it was computed was by bundling all car outcomes equal or greater than 3 and creating the 3+ category. One reason that I use the range, rather than the count data from the households is because the frequency after 3+ vehicles quickly decreased. 95% of the observations in the survey were of households who owned [0, 1, 2, 3] automobiles.The idea of using the classification system of 0, 1, 2, 3+ is consistent with other papers in the literature.

Private ownership of vehicles requires a place to park the car when it is not in use. Availability of parking should influence car ownership. For example, suppose that an apartment building does not provide parking, then the household is presented with a decision to pay for monthly parking, or to pay in terms of time from having to park far away and change sides of the streets during alternate side parking. Alternate side parking is when the streets are cleaned and all cars must switch to the other side of the street. However, the available data on residential parking availability is not complete. Some apartment buildings have parking agreements with parking garages and it is hard to determine how much of the parking is utilized by members of the apartment building. As a result, previous researchers have relied on a proxy representing parking supply. Holtzclaw et al., 2002 found that development density and dwelling type can be used to effectively model the effect of parking availability of car ownership. Households who reside in multi-family housing may not be guaranteed parking in their rental agreements. In New York City, where there are no residential parking permits for on street parking. Without a pricing mechanism, the parking is filled by first-come first-serve basis. Similarly Giuliano and Dargay (2006) used residence type in the form of a dummy variable to serve as a proxy for parking in their models for car ownership choice. Single-family homes are privately owned and homeowners have the choice to substitute some of their property to serve as a parking spot. For example, building a driveway on their property is not an option of those living in multi-family housing, where the needs of the community overrule the desires of the individual. Other papers on car ownership Potoglou and Kanarogou (2008), Hess and Ong (2002) and Chu (2002), used a dummy variable for residence type in Hamilton, Canada, Portland, OR, and New York City respectively.

An income variable is used to control for the difference in household ability to purchase vehicles. The income variable is coded in the following ranges {Less than $30k, $30k-74,999, 75-100k, 100k+}. The household income level is encoded as a dummy variable because the income levels are interval-based, mutually exclusive, discrete levels. Therefore households can not have income levels that exist in two different income ranges I coded this as a dummy variable. During exploratory analysis, it was found that as income increased, the percentage of households who owned more vehicles decreases and the percentage of households who did not own a car decreased. This supports the intuition that households who have a higher income may choose to own more vehicles.

In order to control for the variability in the size of households, I allow a continuous variable called "HHSIZE" to enter the model. This variable is meant to serve as a control variable. As the household gets larger, I expect that members of the household will compete for use of the vehicle. Also, as the household gets larger, the probability that a household has dependents such as children increases as well. Households who have young children may be more likely to own vehicles so they can be more mobile. For example, it is not easy to drop your child off at daycare if you ride a scooter to work.

The number of members within the household who have a valid drivers license should influence the number of vehicles the household chooses to own. It is coded as a dummy variable and it can take on the values of zero, one, and two-plus. I selected the base value for the dummy variable to be having one member of the household with a valid drivers license. The first reason for this variable is that households with zero licensed drivers would likely incur disutility for owning cars that they can not legally drive. Also, for each member in the household that has a driver's license, it serves as a signal for independence among members. Driver's licenses require passing exams, and learning how to drive, and any rational member of a household would not incur these costs unless they had to drive themselves somewhere regularly.

The geographic location of a household may influence the car ownership decision. To capture a potential effect of location, I include a dummy variable "NYC" that takes the value of one if the household is located in New York City (5 boroughs), and zero otherwise. The places where the value will take on a zero is primarily suburban areas within the New York Metropolitan area. It is well understood that the development density is higher closer to the city center than out in the suburbs. One explanation for this is that the price of land generally tends to increase as one gets closer to the city center. In response to increasing land prices, developers choose to substitute land for capital, choosing to build taller buildings closer to the CBD. The flip side of this decision making is that developers constructing buildings far away from the CBD, such as on Long Island, are more likely to substitute away from capital and towards using more land. As a result, housing, shopping, and employment tend to be geographically farther apart, requiring increased automobile dependence. Models that seek to understand the factors of car ownership have found that increased development density of the land (i.e residential density and employment density) is correlated with lower rates of car ownership (Chen et al., 2008).

Suburban residential communities tend to have more strict zoning laws, with residential developments and dedicated commercial districts. The result is that the zones tend to be homogeneous, meaning that an area zoned for residential development is composed of almost exclusively residential housing. This consequently leads to a less mixing of residential and employment within the same zone, and (Zegras, 2010) found that as the mixture of employment and housing increases, the probability of a household owning a car decreases.

Higher population density in urban areas makes mass transit a viable alternative. Mass transit is more viable in densely populated areas because more people using the service causes a self reinforcing change in the direction of an extreme outcome. For example, as more households opt to riding a bus, the bus must increase frequency in order to accommodate the increase in demand. As a result, the time waiting for the bus decreases, this lowers the cost of taking the bus and encourages more ridership.

The number of workers in the household "HHWRK" should be important for car ownership decisions. One reason for this is that when workers commute to work by car, they often park their car at work, which eliminates the possibility of sharing the car between other members of the family. This variable is coded as zero, one, or two plus as a dummy variable. Households with several workers may be a proxy for a household with many independent people. Usually we expect that as a size of a household gets larger, the amount of dependents would increase. This variable "checks" the instance where there are many independent workers in the household and suggests to the model that they may exhibit high levels of car ownership.

5. Results

When using a Multinomial Logit Model, the first set of results are the parameter estimates and their respective coefficients. For each of the possible outcomes of the dependent variables, there will be a separate equations.

Table of results on the next page

```
Iteration 0:   log likelihood = -23624.671
Iteration 1:   log likelihood = -15604.818
Iteration 2:   log likelihood = -15253.917
Iteration 3:   log likelihood = -14866.647
Iteration 4:   log likelihood = -14849.448
Iteration 5:   log likelihood = -14849.363
Iteration 6:   log likelihood = -14849.363

Multinomial logistic regression                 Number of obs     =     17,537
                                                LR chi2(27)       =   17550.62
                                                Prob > chi2       =     0.0000
Log likelihood = -14849.363                     Pseudo R2         =     0.3714
```

hhveh_r	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
0	(base outcome)					
1						
1.resty	1.989387	.0809048	24.59	0.000	1.830816	2.147957
hhsiz	.0399512	.0315483	1.27	0.205	-.0218824	.1017848
hhwrk						
0	.0196372	.0695619	0.28	0.778	-.1167017	.155976
2	-.4807401	.0919554	-5.23	0.000	-.6609693	-.3005109
hhlic						
0	-4.16696	.1464131	-28.46	0.000	-4.453925	-3.879996
2	.5536615	.0857545	6.46	0.000	.3855857	.7217373
incom_r						
2	.6275638	.0700059	8.96	0.000	.4903548	.7647728
3	.746278	.1027757	7.26	0.000	.5448412	.9477147
4	.5188598	.0891195	5.82	0.000	.3441888	.6935307
_cons	.0115516	.0807606	0.14	0.886	-.1467362	.1698395

2						
1.resty	3.754911	.0915763	41.00	0.000	3.575425	3.934397
hhsiz	.009264	.0368212	0.25	0.801	-.0629042	.0814321
hhwrk						
0	-.3900181	.0960324	-4.06	0.000	-.5782382	-.2017981
2	-.176876	.0990893	-1.79	0.074	-.3710874	.0173354
hhlic						
0	-5.466854	.7149489	-7.65	0.000	-6.868128	-4.06558
2	3.233603	.1021553	31.65	0.000	3.033382	3.433824
incom_r						
2	1.017021	.1033631	9.84	0.000	.8144332	1.219609
3	1.368967	.132103	10.36	0.000	1.110049	1.627884
4	1.21451	.1156653	10.50	0.000	.98781	1.44121
_cons	-3.265401	.1222216	-26.72	0.000	-3.504951	-3.025851
3						
1.resty	5.034441	.1225361	41.09	0.000	4.794275	5.274608
hhsiz	.1966625	.039305	5.00	0.000	.1196261	.2736989
hhwrk						
0	-.399063	.1214561	-3.29	0.001	-.6371127	-.1610134
2	.3106514	.1075022	2.89	0.004	.099951	.5213518
hhlic						
0	-4.322728	.7220948	-5.99	0.000	-5.738007	-2.907448
2	3.153412	.1271194	24.81	0.000	2.904262	3.402561
incom_r						
2	1.218437	.1437188	8.48	0.000	.9367532	1.500121
3	1.544391	.1685891	9.16	0.000	1.213962	1.87482
4	1.670979	.1510081	11.07	0.000	1.375008	1.966949
_cons	-6.209322	.1857753	-33.42	0.000	-6.573435	-5.845209

The way that I interpret these initial results is to begin at the top of the table and to the left where there is a bold number. This bold number represents which level of auto ownership the corresponding coefficients belong to. I chose to use zero cars owned as the base case because it made the interpretation a little more straightforward. For each level of auto ownership, there are the independent variables and their corresponding coefficient for that level of auto ownership. For each level of ownership, there exists an observed utility function that is a function of the independent variables and their associated coefficients. One important thing to note is that the sign of the coefficient is able to be interpreted within each utility function. A positive sign suggests that the variable contributes utility to the household where a negative sign would decrease the utility at that level of car ownership. Also, the magnitude of the coefficient can only be interpreted within the utility function for a single level of car ownership. For example, a comparison of magnitudes of coefficients between different levels of car ownership does not have a practical interpretation. The magnitude of the coefficients can be compared in relative terms to the other coefficients within a single utility function.

The convergence of the function from ~ -23,000 to ~ -15,000 suggests that the variables were able to achieve some headway in minimizing the maximum likelihood function. Similarly, a pseudo-r^2 of 0.37 is reasonably high and the p-value of the chi2 is statistically significant at the 0.01. These suggest that the results carry weight and that there is a relationship between the variables and the level of car ownership.

For the dummy variable “resty” which represents the type of dwelling, the coefficient across all levels of car ownership is positive. This is consistent with the initial hypothesis that a single-family dwelling would provide additional utility across all levels of ownership. Also, the statistical significance of the type of dwelling is highly significant and appears to be more significant as the levels of car ownership increase.

Another dummy variable “HHLIC” which represents the number of licensed drivers that are members of the household. The dummy variable can take on the values of 0, 1, and 2+. The base level is 1. When the variable takes on the level of 0, that can be interpreted that the household does not have any licensed drivers. For all car levels that are not zero, I expect that all of the signs should be negative. This would reflect the disutility of owning a car that can not be legally operated since no members of the household can legally drive. This result is confirmed in the table of initial coefficients.

Next, I present the table of marginal effects for the variable of interest, “resty”, the type of dwelling.

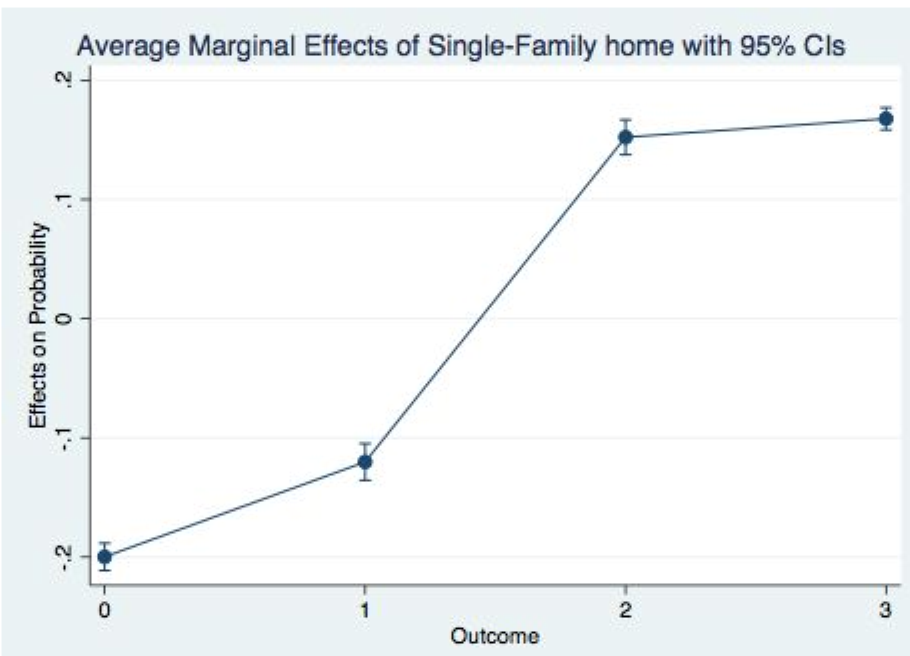
```
Average marginal effects                        Number of obs     =     17,537
Model VCE    : OIM

dy/dx w.r.t. : 1.resty
1._predict   : Pr(hhveh_r==0), predict(pr outcome(0))
2._predict   : Pr(hhveh_r==1), predict(pr outcome(1))
3._predict   : Pr(hhveh_r==2), predict(pr outcome(2))
4._predict   : Pr(hhveh_r==3), predict(pr outcome(3))
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
0.resty	(base outcome)					
1.resty						
_predict						
1	-.1997775	.0058544	-34.12	0.000	-.211252	-.188303
2	-.1201576	.0078588	-15.29	0.000	-.1355607	-.1047546
3	.1522204	.0074652	20.39	0.000	.1375888	.1668519
4	.1677148	.004741	35.38	0.000	.1584225	.177007

Note: dy/dx for factor levels is the discrete change from the base level.

The top of the table tells what variable the change in probability is with respect to. In this instance, it is in response to the “resty” type of dwelling variable. Just below that, there is a list of one, two, three, four, and these specify which levels of auto ownership are associated with each number. In this case, the number one corresponds to the event where the number of cars owned by the household is zero. The base outcome for the dummy variable is that the type of dwelling is multi-family housing. The way I interpret this table is, for a level of car ownership of zero cars, given a change in type of dwelling from multi- to single-family housing, I would expect a 20% reduction in the probability of owning zero cars. For the outcome of owning three-plus vehicles, I interpret the results that the probability of owning three-plus vehicles would increase by 16.7% from a family residing in single-family given the base case of living in multi-family housing. One important aspect is the signs of the coefficients of the marginal effects. The signs are negative for the levels of zero and one, and the signs are positive for two and three-plus. The way that I interpret this is, given a household living in single-family home, relative to the multi-family base outcome, I would expect households to be less likely to own zero or one cars, and more likely to own two or three cars. One easy way to see this is in a plot of the marginal effects. The x-axis represents the outcome, and the y-axis represents the change in the probability of living in a single-family home relative to the base case of living in multi-family housing. It is clear that there is a positive slope, and for the first derivative this means that for higher levels of car ownership, the probability of that outcome is becoming more likely. The p-value for this marginal effect is very statistically significant. A null hypothesis of a two tailed test that there is no relationship between dwelling type and car ownership is rejected at the 0.01 level.

This result is consistent with the previous literature that single-family housing can serve as a proxy for parking availability. Households who live in single-family homes are more likely to own more vehicles. As mentioned, this may be because of the more spread out of the built environment. Areas with more single-family homes are often more spread out and less mixed, and this may nudge households into having an always available form of transportation.



6. Conclusion

Car ownership is a choice that households make whether they realize it or not. After closely following the recommended models from the literature, and including variables of interest that have been used in previous literature, I was able to find statistically significant results. Households who live in single-family housing are more likely to own higher levels of cars and less likely to own fewer levels of cars.

One limitation of this paper is that I was not able to incorporate all of the variables that I would have liked. For example, one way to improve this paper would be to extract the commuting trips from the travel survey and use them in the model. For instance, given that a household lives close to a train station, and that they work in the city, they would be more likely to own fewer cars. Also, within the same dataset, there is personal information such as age, gender, and occupation. This information could be used to enhance the model. It seems likely that households with older populations are more likely to drive than walk, active transportation, or public transit. A deeper dive would include building a joint model that would simultaneously solve for residential choice, mode choice, and car ownership choice with the new data.

Works Cited

Baldwin Hess, D., & Ong, P. M. (2002). Traditional neighborhoods and automobile ownership. *Transportation Research Record*, 1805(1), 35-44.

Ben-Akiva, Moshe E., Steven R. Lerman, and Steven R. Lerman. *Discrete choice analysis: theory and application to travel demand*. Vol. 9. MIT press, 1985.

Bhat, Chandra R., and Vamsi Pulugurta. "A comparison of two alternative behavioral choice mechanisms for household auto ownership decisions." *Transportation Research Part B: Methodological* 32.1 (1998): 61-75.

Chu, You-Lian. "Automobile ownership analysis using ordered probit models." *Transportation Research Record* 1805.1 (2002): 60-67.

Wang, H., Fu, L., Zhou, Y., & Li, H. (2008). Modelling of the fuel consumption for passenger cars regarding driving characteristics. *Transportation Research Part D: Transport and Environment*, 13(7), 479-482.

Giuliano, Genevieve, and Joyce Dargay. "Car ownership, travel and land use: a comparison of the US and Great Britain." *Transportation Research Part A: Policy and Practice* 40.2 (2006): 106-124.

Guerra, E. (2015). The geography of car ownership in Mexico City: a joint model of households' residential location and car ownership decisions. *Journal of Transport Geography*, 43, 171-180.

Holtzclaw, John, et al. "Location efficiency: Neighborhood and socio-economic characteristics determine auto ownership and use-studies in Chicago, Los Angeles and San Francisco." *Transportation planning and technology* 25.1 (2002): 1-27.

McFadden, Daniel. "Conditional logit analysis of qualitative choice behavior." (1973).

Potoglou, Dimitris, and Pavlos S. Kanaroglou. "Modelling car ownership in urban areas: a case study of Hamilton, Canada." *Journal of Transport Geography* 16.1 (2008): 42-54.

Salon, Deborah. "Neighborhoods, cars, and commuting in New York City: A discrete choice approach." *Transportation Research Part A: Policy and Practice* 43.2 (2009): 180-196.

Train, Kenneth. "A structured logit model of auto ownership and mode choice." *The Review of Economic Studies* 47.2 (1980): 357-370.

Train, K., & Sonnier, G. (2003). Mixed logit with bounded distributions of partworths. *Applications of Simulation Methods in Environmental Resource Economics, edited by A. Alberini and R. Scarpa. New York: Kluwer Academic*.

Train, Kenneth E. *Discrete choice methods with simulation*. Cambridge university press, 2009.

Zegras, C. (2010). The built environment and motor vehicle ownership and use: Evidence from Santiago de Chile. *Urban Studies*, 47(8), 1793-1817.