Car Ownership and the Impact of Built Environment, Demographics and Transport Systems

Rolf Moeckel1 and Di Yang2

1. Department of Civil, Geo and Environmental Engineering, TUM Institute for Advanced Study, Technical University Munich, 80333 Munich, Germany
2. Department of Civil and Environmental Engineering, University of Maryland, College Park, Maryland 20742, USA

Abstract: Auto ownership is one of the most important linkages between travel demand and land use. Residents in denser, urban or more transit accessible neighborhoods tend to own fewer cars. Car ownership influences almost all aspects of travel behavior, including travel frequency, travel distances, mode choice and time-of-day choice. At the same time, car ownership affects residential location choices, as households owning cars are less likely to choose urban neighborhoods than households without cars. This paper describes a new microscopic auto-ownership model that has been estimated with survey data. The model is fully integrated with a land use and a transportation model to capture: (1) how owning a car affects travel behavior and location choice; and (2) how the built environment and the transportation needs affect auto-ownership decisions. The model has been validated against census data and is fully operational.

Key words: Auto ownership, vehicle ownership, discrete choice, logistic regression, land-use/transport interactions.

1. Introduction

Auto ownership is a driving force for travel behavior. It influences the number of trips, the chosen destinations, the mode of travel and the time of day of out-of-home activities [1]. On the other hand, car ownership is influenced by the built environment. Ewing et al. [2] found that households in low-density areas own 0.5 cars more than households in high-density areas, even after controlling for income. This makes car ownership a crucial link between transportation and land use.

In addition, household characteristics play in key role in the decision to purchase a vehicle. Household income, size, number of workers and number of children are positively correlated with the number of cars owned by a household.

Car ownership also has a significant impact on economic opportunities. Gurley and Bruce [3] studied longitudinal surveys of current and former welfare recipients in Tennessee and found that access to car increases the probability of being employed and leaving welfare dependency. Car access also enabled survey respondents to find better-paying jobs.

This paper reviews the state of the art and develops the design of a new car-ownership model. A logistic regression is estimated and the interactions between transportation and land use through car ownership are explored in detail.

2. State of the Art

To the knowledge of the authors, the first computational auto-ownership model ever published was applied to Leeds, U.K. in the 1960s [4, 5]. Income and residential density were used as independent variables in a linear regression, of which the former explained 12% and the latter explained 54% of the variation. Nowadays, most models report income as the most influential variable. It is possible that in the 1960s, auto ownership was affordable for high-income households only (there were 84 autos per 1,000 households in Leeds in 1960, compared to 430
in 2011), which could explain the overwhelming importance of residential density in Beesley and Kain’s model.

In the late 1990s, Dargay and Vythoulkas [6] estimated an auto-ownership model using expenditure surveys in the UK. They developed a linear model to estimate the average number of autos, subject to income, household composition, transportation costs and area type. Despite finding fairly good model fits, the authors admit that income and car ownership are not linearly related to each other. In fact, Blumenberg and Pierce [7] found that car ownership rises quickly from very low-income households to low-income households, but an income increase from medium income to higher income has less effect on car ownership. Using a linear relationship between income and car ownership is likely to misrepresent this impact.

Bhat and Guo [8] developed a joint mixed multinomial logit-ordered response structure model to estimate residential location choice and auto ownership jointly in one model. They found that household demographics are more important to explain car ownership than the built environment variables, such as density or accessibility. They identified household income as the dominant factor in car ownership. While the authors explained theoretically that residential self-selection suggests benefits of estimating location choice and auto ownership in one model, they concluded that separate models are equally adequate if residential location choice accounts for sorting effects of demographics.

Zegras [9] estimated a multinomial logit model of household motor vehicle ownership for Santiago de Chile. He found that by far the most dominant independent variable for vehicle ownership was income, though he also found statistically significant built-environment variables, such as auto-to-bus access ratio, proximity to metro stations and the distance to the CBD (central business district).

Gurley and Bruce [3] emphasized the relevance of car access over car ownership. Particularly for households in lower income groups, it is not uncommon to have more workers than autos in a household, creating competition for the use of the car. The person not having access to the car has very similar travel behavior as a person in a 0-car household. Also, ride-sharing becomes much more relevant for households with fewer autos than workers.

A good overview of auto-ownership model types has been published by de Jong et al. [10], including models that could be used for World Bank investment decisions, tax revenue forecasting, energy consumption forecasts and transport demand. While this paper focuses on the latter only, it is notable that such models have been used in very different contexts. The authors define nine model types, and the model discussed below fits best into the type called “static disaggregate ownership models”. It should be noted, though, that this particular model is not static but fully integrated with a land use and a transportation model. De Jong et al. concluded that this particular model type is strong for long-term forecasts while using solid theory with modest data requirements, though the distinction of vehicles types is very limited when using this model type.

Most auto-ownership models described in the literature distinguish the impacts of household attributes and the impact of the built environment. Analyzing the effects of the built environment on auto ownership bears the problem that households may have selected a certain neighborhood to live because of its preference to own or not to own cars. In other words, a household may have moved into a dense urban neighborhood with good transit access because it prefers not to own a car. In this case, the built environment would not cause the auto choice. This is called self-selection [11] and blurs the analysis of the impact of the built environment on auto ownership. Mokhtarian and Cao [12] analyzed seven methods and their capabilities to account for self-selection effects.
Ultimately, they argue, conventional methods are unable to account for the impact of self-selection. Alternatively, they propose to either conduct true panel studies with residents who relocated or to set up natural experiments to test the impact of the built environment. Neither one was accessible for this research, which is why the methods presented can not truly account for a possible self-selection bias.

The majority of the models described in the literature estimate average number of cars. While this is sufficient for aggregate analyses, this paper develops a microscopic approach that was needed for full model integration. In microscopic auto-ownership models, the number of vehicles is modeled separately for each individual household, which allows accounting for detailed household attributes when choosing the number of vehicles owned. Models found in the literature that are capable of predicting discrete number of vehicles tended to work with a rather limited number of independent variables. Those models tended to focus on income and density at the household’s residence. While these are crucial factors to explain auto ownership, most models ignore transit accessibility. Zegras [9] included auto-to-bus access ratio, proximity to metro stations, but fell short of using potential accessibility [13] by transit. The model described here is fully integrated with a transportation model and a land use model, including both linkages to the built environment and to travel behavior.

3. Links between Land Use and Transport

The auto-ownership model developed here forms a crucial link between land use and transportation. Wegener [14] identified car ownership to be driven by activities a household wants to pursue and costs associated with travel (by all modes). Fig. 1 shows a selection of household attributes and neighborhood attributes that drive the car-ownership decision. At the same time, car ownership is one of the most crucial factors that drive travel behavior [1], and thereby affecting the overall transport system performance.

For this reason, both highway and transit congestion are heavily influenced by car ownership.

Both land use design and the transportation system heavily influence auto ownership, which in turn has substantial impact on congestion and urban growth patterns. By fully integrating the auto-ownership model with a land use model and a transportation model, most of these interactions are represented explicitly in the model described below.

To account for land use/transport interactions through auto ownership, the model described here has been integrated with a transportation model and a land use model. The MSTM (Maryland Statewide Transportation Model) [15] models travel behavior and the microscopic SILO (simple integrated land use orchestrator) land use model [16] simulates land use changes.

4. Model Design

As auto ownership is a major driver for travel demand and may influence household location choice, a model was developed that would model the number of vehicles subject to household and neighborhood characteristics. Given that SILO is designed as a microsimulation, it was necessary to simulate the auto ownership for each household individually as well.
The literature review described various auto-ownership models. Most applied some form of regression analysis. Such models are appropriate if the average number of vehicles shall be modeled, such as how many cars do households with a certain income level own in a given zone on the average. If the auto ownership is applied as a microsimulation, however, a discrete number of vehicles needs to be modeled for each household. In particular, there are three econometric methods suitable to estimate discrete values, namely discriminant analysis, contingency analysis and logistic regression. Contingency analyses require independent variables to be nominal, a restriction discriminant analysis and logistic regression do not carry. Discriminant and contingency analysis are strong at revealing linear relationships, while the logistic regression uses the s-shaped log function, which is assumed to better represent psychological processes and values of decision-making [17]. The discriminant analysis further requires: (1) independent variables to be normally distributed; (2) identical variance-covariance matrices in each case distinguished as a dependent variable. The logistic regression is more flexible and does not carry such limitations [18], and thus, was chosen as the preferred method of analysis for this auto-ownership model.

The 2007~2008 TPB (Transportation Planning Board)/BMC (Baltimore Metropolitan Council) household travel survey was conducted in the Washington/Baltimore metropolitan region. 14,365 households were surveyed, and both household characteristics and travel behavior were recorded. Fig. 2 shows the study area covered by this survey.

Based on the literature review and team brainstorming, several independent variables were defined that are expected to explain auto ownership. Those include both household level variables and neighborhood variables. In contrast to other models reported in the literature, this auto ownership model was expected to be used in an operational model to forecast auto ownership to 2040. As a consequence, only variables could be included that either can be forecasted or otherwise could be reasonably assumed to not change over time. Household variables considered include household income, household size and number of workers. Neighborhood characteristics tested include population density, employment density, auto accessibility and transit accessibility. Densities were included as they do not only represent the opportunities that can be reached by walking or biking in the neighborhood, which could reduce the need to own a car, but also because they serve as a proxy for scarcity of parking that is likely to affect auto ownership. Zegras [9] found a small but significant impact of street network density. This model, however, predicts car ownership for a large metropolitan study area and works with a simplified network, which precluded the authors from using street network density as an independent variable. Zegras also found that broadband internet access is positively correlated with auto ownership. While this might be a curiosity of Santiago de Chile, Zegras explains that broadband access may serve as a “lifestyle” proxy. While this is an interesting variable, it was not included in this analysis.
because internet access was not a variable surveyed here, nor could it be reasonably forecasted into the future.

The model has the following logistic form:

\[ u_{h,a} = \exp \left( c_a + \sum_k \left( \alpha_{k,a} \cdot v_{h,k} \right) \right) \]  \hspace{1cm} (1)

where:

- \( u_{h,a} \) = utility for household \( h \) to own \( a \) number of autos (1, 2 or 3+);
- \( c_a \) = constant for unincluded attributes for \( a \) number of autos;
- \( \alpha_{k,a} \) = parameter of attribute \( k \) for \( a \) number of autos;
- \( v_{h,a} \) = value of attribute \( k \) for household \( h \).

As households that own more than three vehicles are rare (1.4% in Maryland and 0.9% in Washington, D.C. according to census data), four auto-ownership levels are distinguished: 0, 1, 2 and 3+ vehicles. Utilities for owning 1, 2 or 3+ autos are calculated in comparison to owning no auto. Probabilities are calculated as follows:

\[ p_{h,a} = u_{h,a} \cdot \frac{1}{1 + \sum_{a=1}^{4} p_{h,a}} \quad \text{and} \quad p_{h,0} = 1 - \sum_{a=1}^{4} p_{h,a} \]  \hspace{1cm} (2)

where, \( p_{h,a} \) = probability of household \( h \) to own \( a \) number of cars.

In an earlier version of the model design, a positive coefficient was estimated for transit accessibility. This outcome was unexpected: The estimation suggested that households would own more cars in neighborhoods with higher transit accessibility, the opposite of what is proposed in the literature. At that time, density was used as a continuous (or metric) variable. Using multi-collinearity analysis, it was found that transit accessibility and density correlated with each other, leading to unrealistic coefficients. In a first attempt to solve this issue, different models were estimated for zones with transit access and zones without transit access. While this led to reasonable parameters in each model, parameters were quite different, and the sudden change in parameters between zones with and without transit access was deemed to be unrealistic. This led to the idea to use categories for transit accessibility and density. In other words, these variables were converted from metric variables into ordinal variables. Three density categories and three transit accessibility categories were defined. Estimating the model with these independent ordinal variables led to coefficients that had the expected sign (+ or −) and seemed to have reasonable values.

Yet, three categories for density and for transit accessibility appeared to be fairly coarse. Values for these variables would have to change a lot to have an impact on auto ownership. For example, in some cases, transit accessibility could double, but the category would not change, and therefore, probabilities for auto ownership would remain unchanged. Hence, transit accessibility was moved back into a metric variable and only density was applied as an ordinal variable. Parameters remained very reasonable. Finally, the number of density categories was increased systematically from 3 to 10 density categories. While 10 density categories produced the best model fit, a diminishing return was found for additional number of categories (Fig. 3). Therefore, 10 categories were used to account for density. The revised model is fully sensitive to even small changes in transit accessibilities.
While designing the auto-ownership model, it was considered to distinguish federal employees from other employees. Federal workers receive transit passes at lower costs. This leads to higher transit shares among federal employees, despite above-average incomes of this population segment. However, after careful consideration, this idea was abandoned. Even though the household travel survey distinguishes between federal employees and other employees, it would be very challenging to forecast the number and spatial distribution of the federal workforce over the next decades. While the federal government traditionally has been an economic engine of the Washington, D.C. region, ongoing calls for tax cuts and reduction of government make it uncertain how many workers will be employed by the federal government in the future. Multi-worker households would add the additional complication if some household members work for the federal government and others do not. It is acknowledged that not representing federal workers explicitly might slightly overstate car ownership in the Washington, D.C. area.

Blumenberg and Pierce [7] have shown that low-income households rapidly exchange rising income for higher auto-ownership rates, while higher-income households are less likely to spend additional income on auto ownership. Once a certain middle-income level has been reached, car ownership does not change much with rising income. If a household’s annual income rises from $15,000 to $20,000, this household is much more likely to spend additional money on transportation than another household whose income rises from $50,000 to $55,000. A linear relationship would misrepresent the impact of income on car ownership. Instead, the 12 income categories provided by the survey were used in the estimation directly as an ordinal variable.

### 5. Estimation Results

In this model estimation, only parameters that turned out to be significant at least at the 0.1% level were included. Table 1 shows the estimation results. Owning no car is used as the base case, and the three columns show the parameters to estimate the probabilities to own 1, 2, or 3 or more cars.

The impact of income on car ownership proposed by Blumenberg and Pierce [7] was confirmed with this model estimation. Income affects car ownership almost linearly below $60,000 (Fig. 4). For higher income groups, income has much smaller impact on auto ownership. The data suggest that an income of $60,000 is a turning point. Up to $60,000, additional income is likely to be invested in transportation. Zero-car households with higher income are unlikely to have opted against car ownership for financial reasons.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Model estimation results.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>1 car</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.421</td>
</tr>
<tr>
<td>Household size</td>
<td>−0.121</td>
</tr>
<tr>
<td>Number of workers</td>
<td>0.327</td>
</tr>
<tr>
<td>Transit accessibility to employment (scaled)</td>
<td>−0.022</td>
</tr>
<tr>
<td>Household income</td>
<td></td>
</tr>
<tr>
<td>Less than $10,000</td>
<td>−3.783</td>
</tr>
<tr>
<td>$10,000–$14,999</td>
<td>−3.134</td>
</tr>
<tr>
<td>$15,000–$29,999</td>
<td>−2.423</td>
</tr>
<tr>
<td>$30,000–$39,999</td>
<td>−1.964</td>
</tr>
<tr>
<td>$40,000–$49,999</td>
<td>−1.374</td>
</tr>
<tr>
<td>$50,000–$59,999</td>
<td>−0.811</td>
</tr>
<tr>
<td>$60,000–$74,999</td>
<td>−0.996</td>
</tr>
<tr>
<td>$75,000–$99,999</td>
<td>−0.913</td>
</tr>
<tr>
<td>$100,000–$124,999</td>
<td>−0.716</td>
</tr>
<tr>
<td>$125,000–$149,999</td>
<td>−0.271</td>
</tr>
<tr>
<td>$150,000–$199,999</td>
<td>−0.236</td>
</tr>
<tr>
<td>$200,000 or more</td>
<td>0</td>
</tr>
<tr>
<td>Employment density 1 (lowest)</td>
<td>2.678</td>
</tr>
<tr>
<td>Employment density 2</td>
<td>2.306</td>
</tr>
<tr>
<td>Employment density 3</td>
<td>1.412</td>
</tr>
<tr>
<td>Employment density 4</td>
<td>1.461</td>
</tr>
<tr>
<td>Employment density 5</td>
<td>1.121</td>
</tr>
<tr>
<td>Employment density 6</td>
<td>0.762</td>
</tr>
<tr>
<td>Employment density 7</td>
<td>0.895</td>
</tr>
<tr>
<td>Employment density 8</td>
<td>0.471</td>
</tr>
<tr>
<td>Employment density 9</td>
<td>0.678</td>
</tr>
<tr>
<td>Employment density 10 (highest)</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: Nagelkerke’s pseudo R square: 0.544; Cox and Snell’s pseudo R square: 0.500.
and thus, an increase in income is less likely to result in owning more vehicles. Such households have other reasons not to own a car, such as living in a neighborhood with sparse parking or personal travel preferences that do not include owning a car.

Fig. 5 visualizes the density parameters. Density curves are exponentially shaped (note the logarithmic x-axis). In rural areas, most households own at least one car. In denser areas, car ownership tends to be lower, which is in line with estimation results found elsewhere in the literature review [6, 9]. In other words, if a household moves from a somewhat rural neighborhood to a more rural neighborhood, the probability to own cars rises sharply. By contrast, households moving from a very urban neighborhood to a somewhat less urban show a much smaller increase in car ownership.

6. Model Validation

The model was estimated using the TPB/BMC household travel survey and is applied to the MSTM study area, which includes all of Maryland, Delaware, the District of Columbia, and parts of Virginia, West Virginia and Pennsylvania. Census data were used as independent data for model validation.

The number of autos per household is validated in Fig. 6. The model slightly overestimates households with one auto by 2% and underestimates households with three or more cars by 3%. Overall, the model replicates auto ownership well in the MSTM study area.

Fig. 7 compares the average number of autos per household modeled versus observed by jurisdiction. The model matches auto-ownership rates for the entire study area and for Maryland very well. Auto ownership is overestimated by 0.2 autos per household.
in Washington, D.C. and underestimated in Delaware and parts of Virginia that are covered by the MSTM study area by 0.1 autos per household.

A likely explanation for overestimating car ownership in the District of Columbia might be that race has not been represented explicitly in this estimation. Blumenberg and Pierce [7] found that African-Americans are less likely to purchase automobiles, even after controlling for income. In the District of Columbia, 49% of the population identified themselves as African-Americans [19], which is a much higher share than other jurisdictions shown in Fig. 7.

Another factor disregarded in this model is the costs for parking. Within the District of Columbia, parking is much more expensive and difficult to find than anywhere else in the study area. Tam and Lam [20] found that parking constraints have a significant impact on auto ownership. Unfortunately, parking capacity is difficult to measure and was not included in this research. However, density served as a proxy for parking costs. Finally, governmental workers who may have access to free transit passes were not represented explicitly. This simplification may have affected the District of Columbia to a larger degree than other jurisdictions in the area.

7. Conclusions

Auto ownership is one of the key drivers for travel behavior. Households owning a car tend to make more trips that are longer in distance. They also have lower shares of transit and non-motorized travel. To assess travel demand, it is crucial to capture the right level of auto ownership. This is particularly relevant because of the spatial pattern of higher car ownership in suburban and rural areas than in urban areas. This model has been validated against observed census data and tested in the state-of-practice Maryland statewide model, where it helped to improve model results.

A challenge of auto-ownership modeling is separating household self-selection from the impact of the built environment on car ownership. In this modeling suite, a land use model has been integrated with a transportation model to model self-selection. The survey data, however, only shows the resulting car ownership, not allowing the analyst to spell apart the impact of self-selection versus built environment. Panel surveys that interview the same households before and after household relocation are needed to identify the impact of the built environment on auto ownership.

Acknowledgements

With the support of the Technische Universität München Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement No. 291763. This research was partly funded by the Maryland State Highway Administration. Highly appreciated is also important feedback on model design and performance given by Rick Donnelly and Leta Huntsinger (both Parsons Brinckerhoff), Fred Ducca (University of Maryland) and Subrat Mahapatra (Maryland State Highway Administration).

References

Car Ownership and the Impact of Built Environment, Demographics and Transport Systems

Transportation Research Board 2320: 28-36.


