Research in Transportation Business & Management 7 (2013) 14–26

Contents lists available at SciVerse ScienceDirect

Research in Transportation Business & Management

The influence of spatial and household characteristics on household transportation costs

Peter Haas *, Stephanie Morse, Sofia Becker, Linda Young, Paul Esling

Center for Neighborhood Technology, 2125 W. North Ave., Chicago, IL 60647, United States

A R T I C L E I N F O

Article history:
Received 31 May 2012
Received in revised form 31 January 2013
Accepted 11 March 2013
Available online 28 April 2013

Keywords:
Household transportation behavior
Location efficiency
Transportation cost

A B S T R A C T

It is widely understood that location affects the cost of housing. Until now, the influence of location on household transportation costs has remained elusive, despite transportation costs being the second highest household expense. The Center for Neighborhood Technology (CNT) has developed a model that relates spatial and household variables to auto ownership, auto use, and transit use. A cost is then applied to each of these components to calculate the average household transportation cost for a neighborhood. This paper will focus on the multi-dimensional regression analysis used to relate the independent spatial variables (household density, block size, access to transit and employment, among others) and independent household variables (income, size, workers per household) to the three dependent variables (auto ownership, auto use, transit use). This model is used to estimate the transportation cost variation for a typical household in metropolitan areas, as featured on the website http://htaindex.cnt.org. This paper shows that variation in household transportation cost is related more strongly to the characteristics of a neighborhood than household. It is important to examine this phenomenon at small geographic level, i.e. the neighborhood.

1. Introduction

Transportation costs are a major expense for most US household budgets. In 2010, transportation costs accounted for 16 percent, on average, of household expenditures (BLS, 2010), second only to housing costs. The combined expenditures on housing and transportation exceeded 50 percent of all household expenditures (39% of pre-tax income) (Ibid). However, the standard measure of housing affordability begins and ends with the cost of housing. If a household spends 30% or less of its income on a home, it is considered affordable. This leads people to pass over city neighborhoods or inner-ring suburbs to live in outer-ring communities where housing costs are cheaper.

Although average transportation costs are available in the aggregate for US households and help reveal that transportation is a major cost component for households, the information does not provide clarity on the potential tradeoffs households make regarding their transportation costs when they base their location decisions on housing costs alone. The built characteristics of a city neighborhood can vary dramatically from an exurban community. How might those characteristics affect transportation costs? More specifically, the question driving this research is:

> *"What neighborhood characteristics influence household transportation behavior?"*

While the concept of energy efficiency is a familiar term, locations can be efficient too. Compact neighborhoods with walkable streets, access to transit and jobs, and a wide variety of services and amenities have high location efficiency. They require less time, money, and greenhouse gas emissions for residents to meet their everyday travel requirements.

One of the first significant studies related to location efficiency was a survey of 32 large cities around the world which found that gasoline consumption varied as a function of density both within and across cities; American cities consumed significantly more gasoline than their Australian, European, and Asian counterparts (Newman & Kenworthy, 1989). The authors of the study, Peter Newman and Jeffrey Kenworthy, are credited with coining the phrase "automobile dependency." Follow-up research found that income alone is not a reliable predictor of automobile dependence in different cities, but that urban form, particularly high density, is associated with lower levels of automobile ownership and use, and higher transit use (Kenworthy & Laube, 1999).

In an attempt to garner support for Location Efficient Mortgages (LEMs), Holtzclaw measured the reductions in automobile use and household transportation costs that result from different neighborhood characteristics. The study looked at the effect of residential density, neighborhood shopping, transit accessibility, and pedestrian accessibility on auto ownership and use in 27 neighborhoods in four California metro areas. Density and transit access were statistically significant predictors of vehicle miles traveled (VMT) and density was related to auto ownership rates. Doubling residential density lowered auto ownership and VMT by 16%; VMT was reduced an additional 5% by doubling public transit service (Holtzclaw, 1994).

* Corresponding author. Tel.: +1 773 269 4034; fax: +1 773 278 4877.
E-mail address: pmh@cnt.org (P. Haas).

2210-5395/$ – see front matter
http://dx.doi.org/10.1016/j.rtbm.2013.03.004

Using travel data from the 1990 San Francisco Bay Area Travel Surveys, Kockelman incorporated built environment characteristics into models of travel behavior. Measures of job accessibility, land use mixing, and land use balance had a statistically significant impact on VMT, auto ownership, and mode choice after controlling for demographic characteristics. Density only had a significant impact on auto ownership after controlling for accessibility (Kockelman, 1996).

In San Francisco, Chicago, and Los Angeles, 1990 Census data on vehicles available and 1990–95 odometer reading data were fit to socio-economic and built environment variables thought to explain the observed variation in auto ownership and VMT. With regards to autos per household, the variables with the most explanatory power were net residential density (households per residential acre), per capita income, household size, and transit access. The presence of local shopping was found to be strongly correlated with density and transit and its inclusion did not affect the significance of the model once these variables were accounted for. Density, income, household size, and transit access were also strong predictors of VMT; pedestrian and bicycle friendliness contributed to a lesser extent. Combining the datasets from the three regions produced results that were similar, but not as strong, suggesting that other important variables may not have been identified (Holtzclaw, Clear, Dittmar, Goldstein, & Haas, 2002).

In 2006, researchers at the Center for Neighborhood Technology began to study the relationship between transportation costs and built environment variables. To expand on the work of Holtzcalw et al., a statistical transportation cost model was created for the largest US metros in which three separate multiple regression analyses were conducted to predict auto ownership, auto use and transit use. Independent variables related to the built environment – density, job access, transit connectivity, neighborhood services and walkability – were used in the model and household income and size were held constant. Models were calibrated to measure auto ownership and transit use in the pilot region, Minneapolis-St. Paul, and VMT per household at the block group level using data from the National Household Transportation Survey (Haas, Makarewicz, Benedict, & Bernstein, 2008). This research led to the development of the Housing + Transportation (H+T®) Affordability Index.

In 2008 the H + T Index was expanded to 55 US urban areas, and in 2012, the research team again expanded, updated, and improved the model. The model now covers 180,000 Census block groups, uses 2009 American Community Survey 5-Year Estimates, and is better able to predict household transportation costs.

As a theoretical framework for this research, CNT's working hypothesis is that compact, efficient and convenient neighborhoods allow for a wealth of nearby destinations (jobs, shopping, services, recreation and public transit). These nearby assets shorten trips and afford households alternatives to driving (walking, cycling & transit), reducing auto ownership and driving and increasing use of public transit.

The remainder of this paper will describe the methods and data used to develop the regression analysis relating household transportation behavior with household and spatial variables. The formulae and the final coefficients are included, as is a discussion of the relative importance of the various inputs. Finally, the concept of location efficiency – similar to energy efficiency in buildings and fuel efficiency in vehicles – is described and defined.

2. Data and methods

The household transportation model used is constructed to estimate three dependent variables (auto ownership, auto use, and transit use) as functions of 11 independent variables (median income, per capita income, average household size, average commuters per household, residential density, gross density, average block size, intersection density, transit connectivity, transit access shed, and employment access). To hone in on the built environment's influence on transportation costs, the independent household variables (income, household size, and commuters per household) are set at fixed values to control for any variation they might cause, in order to examine only the efficiency inherent in various neighborhoods.

2.1. Geographic unit and scope

The household transportation model covers the Metropolitan and Micropolitan Areas in the United States, or the Core Based Statistical Areas (CBSAs), as defined by the Office of Management and Budget (OMB). The 2009 American Community Survey 5-year estimates serve as the primary dataset; of the 953 CBSAs as defined in 2008, 936 had ACS data that is usable for this study.

The model is constructed at the Census block group level. Census block groups are used instead of Census tracts in order to focus on smaller neighborhoods with more homogenous attributes.[1] 2009 TIGER/Line shape files from the US Census provide all block groups in the United States (211,668), but the model excludes several thousand block groups[2], including all block groups located outside CBSAs. Despite the elimination of these block groups, the remaining 146,262[3] block groups make for a large dataset to calibrate the model.

2.2. Basic structure of the model

The household transportation model is based on a multidimensional regression analysis, in which formulae describe the relationships between independent household and local environment variables and three dependent variables (auto ownership, auto use, and transit use). Neighborhood level (Census block group) data on household income (both median and per capita), household size, commuters per household, household density (both residential and gross), street connectivity (as measured using average block size and intersection density), transit access, and employment access are used as the independent or predictor variables.

To construct the regression equations, each predictor variable is tested separately; first to determine the distribution of the sample and second to test the strength of the relationship to the criterion variables. For this paper, the regression analysis is conducted in a comprehensive way, thus ignoring the distinction between the local environment variables and the household variables in order to obtain the best fit possible from all of the independent variables. The predicted result from each model is multiplied by the appropriate price for each unit – autos, miles, and transit trips – to obtain the cost of that aspect of transportation. Total transportation costs are calculated as the sum of the three cost components as follows:

$$\text{Household T Costs} = \left[C_{AO}{}^{*}F_{AO}(X)\right] + \left[C_{AU}{}^{*}F_{AU}(X)\right] + \left[C_{TU}{}^{*}F_{TU}(X)\right]$$

[1] There are between 1 and 9 block groups per census tract, with an average of 3.2 block groups per tract. The block groups have on average 540 households in them and an average population of 1445 people, verses tracts that have 18761 households in them and an average population of 4614 people.

[2] The ACS had problems identifying data from block groups (and tracts) in several counties. These "Issue Counties" are in Alabama, Illinois, Iowa, Maryland, Massachusetts, and Texas and not included in the model (United States Census Bureau (2009). Due to incompatible and insufficient data, all block groups in Puerto Rico are excluded. The model eliminates block groups with no households or other missing independent variables. This analysis uses 936 or the 953 CBSAs because of these issues.

[3] Finally, due to confidentiality reasons, the ACS suppressed certain statistics for many block groups. Vehicle availability information is suppressed in many block groups, leaving 146,262 where the model can be calibrated.