Transportation Research Part D 47 (2016) 136–148

Contents lists available at ScienceDirect

Transportation Research Part D

journal homepage: www.elsevier.com/locate/trd

Small area estimation of vehicle ownership and use

Yangwen Liu [a], Cinzia Cirillo [b,*]

[a] Neustar, Inc., New York, USA
[b] Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20740, USA

ARTICLE INFO

Article history:

Keywords:
Vehicle ownership
VMT
Discrete–continuous model
Transferability
NHTS
ACS

ABSTRACT

A national model of vehicle ownership and use is developed for the USA. Decisions about the number of cars owned by households and the annual miles traveled are jointly modeled using a discrete–continuous probit model, which has been estimated on the 2009 National Household Travel Survey (NHTS) data. The model system covers four Census Regions (Northeast, Midwest, South and West) and three area types (urbanized area, urban clusters and rural). Models' estimates have been applied to data extracted from the American Community Survey (ACS) to forecast household vehicle demand at county level. Results show that the national models are transferable to small areas with different geographical and socio-demographic characteristics.

Introduction

In the transportation and energy related literature a consistent number of studies are dedicated to the vehicle ownership problem; applications exist for rural and urban areas (Dargay and Vythoulkas, 1999), large cities (Potoglou and Kanaroglou, 2008) and metropolitan areas (Potoglou and Susilo, 2008), regional (Cirillo and Liu, 2013) and national level (Hensher and Ton, 2002), and for developed (Bhat and Pulugurta, 1998b) and developing countries (Dargay et al., 2007; Li et al., 2010). The majority of these studies are based on data extracted from household travel surveys or dedicated data collections of stated preference type. Vehicle ownership models developed for the U.S.A. are mainly based on the National Household Travel Survey (NHTS), which was last conducted in 2009. The NHTS data contains a wealth of nation's daily travel information, including the number of cars in the households, their type and vintage and the annual miles traveled. However, the NHTS is designed at the national level, and the sample size is not large enough to produce design-based (direct survey weighted) estimates with acceptable precision at the state or finer level (e.g. county, municipality). In general, the NHTS data are not recommended for analysis of categories smaller than the combination of Census division, Metropolitan Statistical Area (MSA) size, and the availability of rail. According to Hu et al. (2007) extrapolating NHTS data within small geographic areas could risk developing and subsequently using unreliable estimates. Even though some States participate to the add-on program that collects supplementary sample for the States participating, many lack the necessary resources to collect local data and do not have the technical capabilities to estimate their own models.

Statistical techniques, known as Small Area Estimation (SAE), have been developed and used to obtain estimates for cases where the number of area-specific sample observations is not big enough to produce reliable direct estimates (Rao, 2003). The term "area" in SAE refers to any subpopulation or domain of interest, such as geographical domains (e.g. state or county), sociodemographic groups (e.g. income, race, age), land use characteristics (e.g. density) (Vaish et al., 2010). These techniques

* Corresponding author.
E-mail addresses: yangwen.liu@gmail.com (Y. Liu), ccirillo@umd.edu (C. Cirillo).

http://dx.doi.org/10.1016/j.trd.2016.05.002
1361-9209/

have been used for a number of applications and the demand for such data small areas has greatly increased during the last two decades. This increase is due to the usefulness of these data in government policy and program development, allocation of various funds and regional planning (Hidiroglou, 2007). For example, Statistics Canada applied them to obtain estimation of health statistics, of average weekly earnings, of under-coverage in the census, and of unemployment rates (Hidiroglou, 2007). The World Bank has developed a technique that combines information from household surveys (which contain comprehensive information) and censuses (which allow fine disaggregation). This statistical inference method is used to create local welfare estimates and detailed poverty maps; the derived information on poverty is sufficiently disaggregated to capture heterogeneity (Elbers et al., 2003). Examples of major small area estimation programs in USA include: the Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program (see the SAIPE web site at http://www.census.gov/hhes/www/saipe/ for more information); the Bureau of Labor Statistics' Local Area Unemployment Statistics (LAUS) program (see the LAUS web site at http://www.bls.gov/lau/); the National Agricultural Statistics Service's County Estimates Program, which produces county estimates of crop yield (USDA 2007, see also at http://www.nass.usda.gov/); and the estimates of substance abuse in states and metropolitan areas (refer to http://www.samhsa.gov/ for details) (Rahman, 2008).

In transportation, very few studies develop comprehensive model systems for small area estimation of vehicle ownership and/or vehicle miles traveled. The barriers include the difficulties to capture demand levels for different population segments and different land use, and the limited data availability for small geographical areas. SAE techniques have been applied by Vaish et al. (2010) to produce small area estimates of the percentage of persons among different age groups having high daily person miles of travel. In Vaish et al. (2010) the authors use NHTS to estimate the model parameters, then the model is applied to predict person level probabilities. The desired small-area-level estimates are obtained by aggregating the person-level predicted probabilities using the appropriate population count projections for each of the 50 States in the USA.

In the context of transferability of transportation related model predictors, Hu et al. (2007) combined the 2001 NHTS data and 2000 census data to provide estimates of regional or local travel, including vehicle trips (VT), vehicle miles of travel (VMT), person trips (PT), and person miles of travel (PMT) by trip purpose and a number of demographics (Hu et al., 2007). In their report, transferability refers to the process of using statistical analysis on survey data sampled at one level (in the case of NHTS, the Census Division-MSA Size-Rail level) to estimate travel statistics at finer levels, such as state or local. Specifically, data was considered "transferable" in the case that estimates resulting from the transferability process were statistically valid (Hu et al., 2007).

This study develops a model system that estimates jointly vehicle ownership and use on national survey data (NHTS 2009). Twelve sub-models are estimated for respectively four Census Regions (Northeast, Midwest, South and West; Fig. 1) and three area types (urbanized area, urban clusters and rural; Fig. 2). Regions and area types were determined

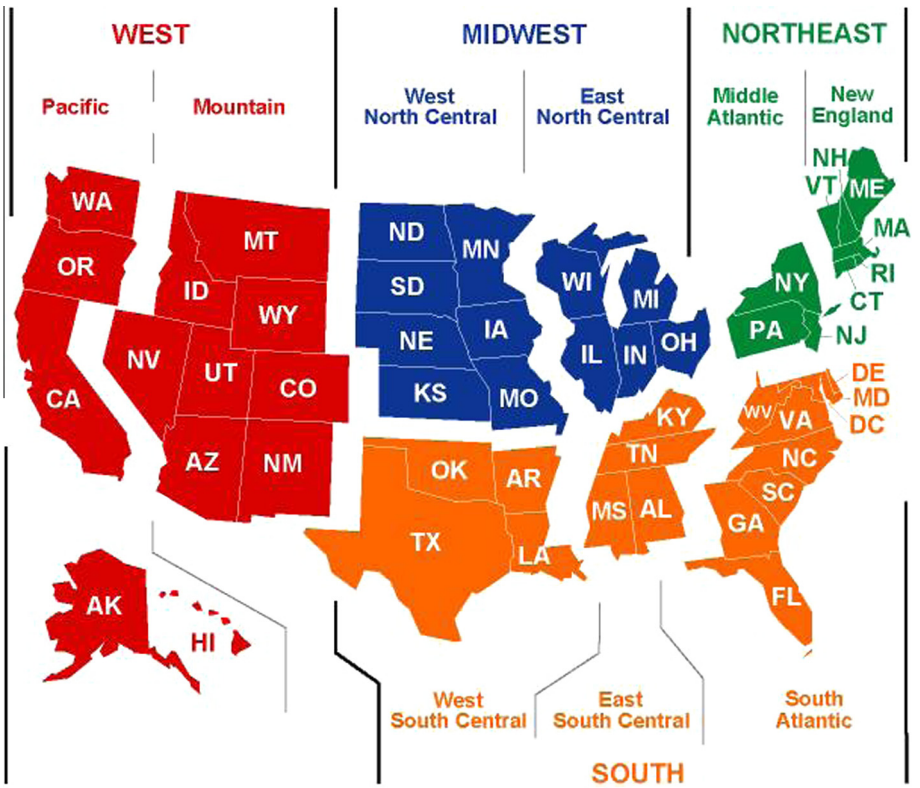

Fig. 1. United States regions (Census Bureau).