University of California Transportation Center
UCTC-FR-2010-05 (identical to UCTC-2010-05)

The impact of residential density on vehicle usage and fuel consumption

Jinwon Kim and David Brownstone
University of California, Irvine
March 2010

The impact of residential density on vehicle usage and fuel consumption*

Jinwon Kim and David Brownstone
Dept. of Economics
3151 SSPA
University of California
Irvine, CA 92697-5100
Email: dbrownst@uci.edu

January 31, 2010

Abstract

This paper investigates the impact of residential density on vehicle usage and fuel consumption. The empirical model accounts for both residential self-selection effects and non-random missing data problems. While most previous studies focus on a specific region, this paper analyzes national level data from the 2001 National Household Travel Survey. Comparing two households that are equal in all respects except residential density, the household residing in an area that is 1000 housing units per square mile denser (roughly 50% of the sample average) will drive 1500 (7.8%) less miles per year and will consume 70 (7.5%) fewer gallons of fuel than the household in the less dense area. The effect of the contextual density measure (density in the context of its surrounding area) is quantitatively larger than the sole effect of residential density. A simulation moving a household from suburban to urban area reduces household annual mileage by 15%.

* Financial support from the University of California Transportation Center is gratefully acknowledged. Jan Brueckner provided many useful comments on earlier drafts. Of course the authors bear full responsibility for any errors or opinions expressed in this paper.

I. Introduction

How does urban sprawl affect household travel behavior? This paper addresses this question by investigating the impact of land use density on household annual mileage traveled and fuel consumption. Following previous studies (Brownstone and Golob, 2009, Boarnet and Crane, 2001), I use land use density as the measure of urban spatial structure (urban sprawl). Although urban sprawl is not simply low density, land use density is highly correlated with almost all measures of urban sprawl (see Badoe and Miller 2001). Most of the previous studies that attempt to measure the influence of urban spatial structure on vehicle usage focus on specific regions in order to guarantee geographic homogeneity (Boarnet and Crane, 2001, Bhat and Guo, 2007, Brownstone and Golob, 2009). This study analyzes national level data, which requires that geographic heterogeneity is controlled for. I use urban/rural dimension dummies (the contextual density measure) and rail transit dummies as the geographic control variables.

The most important econometric issue is residential self-selection effect, which occurs if residents of high-density areas differ in some unobservable characteristics that influence travel behavior. Unless this effect is controlled for, the estimated influence of land use density on travel behavior may be spurious. I follow the same methodology as in Brownstone and Golob (2009) to correct for the self-selection bias by specifying a simultaneous equation model where residential density, household mileage traveled, and fuel consumption are jointly endogenous. These three endogenous variables are assumed to be influenced not only by other endogenous variables but also by a rich set of socio-demographic variables.

The other econometric issue that may result in biased coefficients is the non-random missing data in the key endogenous vehicle fleet characteristics (needed to compute fuel use). To correct for the bias caused by this problem, weights are estimated to compensate for the higher probability of missing data for households owning many vehicles and weighted estimation is used. The wild bootstrap method is used to estimate standard errors that are robust to heteroskesdascity.

I use national level data from the 2001 National Household Travel Survey (NHTS). The big advantage of using national level data is that with the increased sample size we can specify a larger, more accurate model. We can also check whether travel behavior in a specific region is unique or not by comparing the results of the specific region and national level data. I provide comparisons between the results of this paper and the California analysis of Brownstone and Golob (2009) using the same specification. Additionally, we can investigate the impact of certain geographic conditions such as the supply of rail transit on travel behavior. A shortcoming of analyzing national data is that there exists heterogeneity in geographic conditions such as climate, fuel and vehicle prices, and access to public transit. To control for these conditions, I use various geographic control variables such census region, MSA category,

urban/rural dimension and rail transit dummies.

The results indicate that residential density has a statically significant but economically modest influence on vehicle usage and fuel consumption, which is similar to that in other previous studies. However, the effect of the urban/rural dimension variable (contextual density measure) is much greater than the sole effect of residential density. A simulation moving a household across the urban/rural dimension affects household annual mileage and fuel consumption significantly. Compared to the California subsample result of Brownstone and Golob (2009), the influence of residential density on mileage traveled is slightly higher. However, the impact of residential density on vehicle type choice, i.e. tendency toward more fuel efficient vehicle choices for households in denser area, is not statistically significant, unlike the California subsample result.

Literature Review

Studies of the effects of land use density (or other measures of urban spatial structure) on vehicle usage can be divided into aggregate and disaggregate studies. Aggregate studies use spatially defined averages for all variables. One of the most cited papers is Newman and Kenworthy (1999), where the authors implemented a global survey of 37 cities to assess automobile dependence cost. The results indicate that cities with more car use, road provision, and urban sprawl have higher automobile dependence, which causes direct and indirect costs in terms of higher road expenditure, more time spent on commuting, and higher external costs from road deaths and emissions.

Disaggregated studies use household observations of vehicle usage and either city-wide, zonal, or neighborhood averages for urban form variables. Bento et al. (2005) specify disaggregate models of commute mode choice, automobile ownership and annual vehicle miles traveled (VMT). They construct diversified measures of urban form and transit supply: measures of city shape, density of the road network, spatial distribution of population, jobs-housing balance, and bus route and rail miles supplied. Using the 1990 National Personal Transportation Survey, they find that the impacts of any of the urban form measures on travel behavior are frequently insignificant and small in magnitude.

Although disaggregate studies (Bento et al., 2005) that include a rich set of socioeconomic control variables are less subject to residential self-selection bias, it is still possible that residents in high density areas differ in some unobservable characteristics that influence their travel behavior. The only way to deal with this possibility is to construct a joint model of residential density and travel behavior. One of the first to do this is Boarnet and Crane (2001). They specify a demand function for travel in which the

number of trips of different travel modes are influenced by the relative time costs (price of travel) and various socio-demographic "taste" variables. By comparing models where land use density is endogenous and exogenous, they find that the measured influence of land use on travel behavior is very sensitive to the model specifications.

Bhat and Guo (2007) specify a joint mixed multinomial logit model of residential location and number of household vehicles. Their model allows for residential self-selection effects (correlation between the error terms in their equations), but after controlling for a rich set of covariates they do not find any significant effects of residential self-selection. This result implies the necessity of a rich set of socio-demographic variables to control for the residential self-selection. Using San Francisco Bay Area data, they find statistically significant but quantitatively small impacts of built environment measures (street block density, transit availability, and transit access time) on vehicle ownership.

Finally, Brownstone and Golob (2009) directly model the joint choice of residential density and vehicle usage to control for potential residential selectivity. Unlike other previous studies, they also explicitly model vehicle fuel consumption to account for the possibility that residents of high density areas choose more fuel efficient vehicles. Additionally, by adopting a weighting approach, they correct for the bias caused by systematic missing data problems. Using the California subsample of the 2001 NHTS, they find a statistically significant but quantitatively small impact of residential density on household vehicle usage and fuel consumption.

This paper extends Brownstone and Golob's (2009) study using national level data from the 2001 NHTS. The empirical model accounts for both residential self-selection effects and non-random missing data problems. I include various geographic control variables that are necessary in analyzing national data. I find that residential density still has a statistically significant and economically modest influence on vehicle usage. Additionally, I find that the urban/rural dimension dummies have considerable influences on household annual mileage and fuel consumption. I provide some simulation results that move a household across the urban/rural dimensions.

This paper is organized as follows. Section 2 discusses the data used in the study. Section 3 describes the empirical model and the procedure to correct for the self-selection bias and the non-random missing data problem. Section 4 gives estimation results and section 5 concludes.

II. Data

The National Household Transportation Survey (NHTS) contains information on household travel behavior and various socio-demographic variables. It is based on the household survey conducted by the U.S. Department of Transportation every 5 years. The latest survey was conducted and released in 2001.[1] It contains four levels of data set: household, person, vehicle and daily travel. The 2001 NHTS consists of 69,817 household observations. Of these households, 26,083 are in the national sample and 43,734 are from nine add-on areas. The add-on sample is added to help in-depth research for selected states. Among 26,083 households in the national sample, 21,350(82%) households have full information on the key endogenous vehicle fleet characteristics (needed to compute fuel use). Of these households with full information, 4,992 observations are randomly chosen for analysis.[2]

Vehicle Ownership, Mileage Traveled, and Fuel Consumption

Household annual mileage and fuel consumption are the key vehicle fleet characteristics that are used as the joint endogenous variables in our model. Household annual mileage is defined as total mileage per year of household vehicles. Household vehicles are defined as all vehicles owned by all household members. Annual vehicle mileage traveled is recorded and self-reported based on odometer data. Fuel usage computation is based on information about the make, model and vintage of all household vehicles.

Because household annual mileage and fuel consumption are calculated by summing up all household vehicle information, there is no need to explicitly use vehicle ownership as a variable. However, vehicle ownership is closely related to missing data on the key endogenous variables. For example, for households owning one vehicle, 93.5% have full information on the key endogenous variables. However, for households with two, three, four, and five or more vehicles, 87.5%, 66.9%, 51.4%, 29.6% have full information on the key vehicle fleet characteristics, respectively. We can observe that the probability of having full information is a decreasing function of vehicle ownership. Since vehicle ownership is closely related to the key endogenous variables, this means that the sample of households with complete energy information is not a random sample and this can lead to biased results. The methodology to correct for the bias caused by this non-random missing data problem is presented in section 3.

---

[1] The 2001 NHTS is described in detail on the NHTS website. Currently a new survey is being conducted and will be available in January 2010.

[2] Since the random sample is chosen from households with full information, the random sample may not match the population where all possible cases are considered in getting univariate distribution.

Land Use Density

The 2001 NHTS offers the land use density variables in the form of population and housing density at the census and block level. Percentage of renter occupied-housing units is provided at both the block group and tract group level and jobs per square mile is provided at the tract level. As expected, the density variables are all highly correlated.

Table 1 presents percentage of households residing in each density group (housing density at the census and block level) by different geographic regions. As expected, more people reside in denser areas in big cities such as New York and Chicago. We can also find that the urban/rural dimension variable, which categorizes geographic regions into urban, second city, suburban, town, and rural, is highly correlated with residential density. Residential density for rail cities is slightly higher than for non-rail cities.

Table 1: Percentage of households residing in density groups by different geographic regions (national sample, 26038 observations)

Housing units per square mile in Census block group	0 to 50	50 to 250	250 to 1K	1K to 3K	3K to 5K	over 5K
All households	17	16	22	30	8	7
Households residing in New York	1	9	20	27	9	34
Households residing in Chicago	2	7	21	38	13	20
Households residing in Urban	0	0	4	30	25	41
Households residing in Second City	1	6	24	51	12	6
Households residing in Suburban	0	4	31	52	9	3
Households residing in Town	8	36	37	36	1	0
Households residing in Rural	65	24	8	3	0	0
Households residing in Rail city	2	10	20	34	13	21
Households residing in Non-rail city	21	18	23	30	6	2

Urban/Rural Dimension and Rail Transit Dummies

Since there exists heterogeneity in geographic conditions such as climate, vehicle and fuel prices, and access to public transit, travel behavior may be influenced by these variables. For example, travel behavior in New York may be much different than in California. However, we do not know whether this difference comes from residential density or other geographic conditions. To control for these conditions, I use various geographic control variables such as census region, MSA category, urban/rural dimension and rail transit dummies.

The urban/rural dimension variable was devised by Claritas, Inc. to establish objective classifications of geographic locations that were less-boundary dependent and more life-style relevant. It defines five major classifications: urban, suburban, second city, town, and rural. In turn, each of the nation's 226,399 block groups was assigned to one of these categories. The classification is based on density in the context of its surrounding area, yielding a “contextual density measure”. The contextual density measure was based on population density, not of specific block group, but of the larger geographic area not constrained by boundary definitions. Thus, the measure ensures relevance to the density experienced by people. A heuristic approach was taken to determine meaningful breaks for defining each category. See Miller and Hodge (1994) for more detailed description. Table 2 shows average land use densities and selected demographics by the urban/rural classification.

The rail transit dummy indicates that the household lives in an MSA where rail transit is supplied. Among all national households, 24% have access to rail transit. 61% of households in urban areas have access to rail transit while only 3% in rural areas are offered rail transit. The big MSAs dummy indicates that the household lives in one of several big cities including New York, Chicago, Boston, Washington, and Philadelphia. Among all national households, 24% reside in these big cities, and 35% of households in urban area live in big cities.

Table2: Average land use densities and selected demographics by urban/rural dimension classification (national sample, 26038 observations)

Urban/Rural dimension	Urban	Second City	Suburban	Town	Rural
Housing units per sq mile - Block group	4087	2050	1828	742	170
Housing units per sq mile - Tract level	3959	1766	1669	539	85
Population per sq mile - Block group	14095	5125	4873	1622	354
Population per sq mile - Tract level	13472	4325	4336	1082	182
Workers per square mile living in Tract	3611	1747	1747	486	68
Household Annual Mileage	16042	19323	22687	25416	27923
Household Annual Fuel Consumption	714	888	1057	1207	1336
Household income ($1,000)	5.48	5.19	7.15	6.08	4.59
MSA has rail	0.61	0.20	0.36	0.16	0.03
MSA is big city[1)]	0.35	0.13	0.23	0.12	0.02

Notes: 1) Big cities include New York, Chicago, Boston, Washington and Philadelphia MSAs.

Vehicle Usage and Land Use Density

Table 3 presents vehicle usage characteristics sorted by residential density groups. Average household annual mileage decreases as residential density increases. Average household annual fuel consumption is also lower in denser areas. The negative correlation is slightly greater for fuel consumption than mileage, which suggests that households in denser area use more fuel efficient vehicles.

Average vehicles per household reveal similar patterns to mileage and fuel consumption, but the pattern for average number of drivers is less clear than others. Average household incomes in medium density areas are higher than those of denser areas. Table 4 gives descriptive statistics of all variables used in the analysis.

Table 3: Vehicle Usage by Residential Density (average across estimation sample, 4,992 observations)

Housing units per square mile in Census block group	0 to 50	50 to 250	250 to 1K	1K to 3K	3K to 5K	Over 5K
Annual mileage of all household vehicles	25786	24553	22218	18897	14982	13431
Annual fuel consumption in gallons	1308	1217	1067	895	717	599
Vehicles per household	1.95	1.97	1.86	1.63	1.34	0.79
Average number of drivers	1.79	1.81	1.78	1.59	1.40	0.85
Household income ($10,000)	4.57	5.56	6.30	5.50	4.67	4.42

Table 4: Descriptive Statistics

Variables	National			Subsample (N=4992)	
	Observations	Mean	Std. Dev.	Mean	Std. Dev.
Household fuel usage per year in gallons	21370	1073	974	937	964
Housing units per square mile in units of 1,000 - census block group	26029	1.52	1.66	1.88	1.94
Total Mileage per year on all household vehicles	23837	22989	21172	19323	20425
Annual household income in units of $10,000	26038	5.77	4.70	5.32	4.64
Number of children	26038	0.65	1.05	0.59	1.03
Number of workers	26038	1.32	0.97	1.18	0.94
1-worker household	26038	0.33		0.34	
2-worker household	26038	0.36		0.33	
3-worker household	26038	0.08		0.06	
Number of drivers	26038	1.82	0.81	1.59	0.87
1-driver household	26038	0.28		0.31	
2-driver household	26038	0.55		0.49	
3-driver household	26038	0.13		0.09	
Respondent has college degree	26038	0.43		0.42	
Respondent has postgraduate degree	26038	0.11		0.11	
Respondent is retired	26038	0.28		0.30	
Single-person household	26038	0.12		0.15	
Race is Asian	26038	0.02		0.02	
Race is Hispanic	26038	0.03		0.04	
Race is Black	26038	0.07		0.09	
Race is mixed White & Hispanic	26038	0.03		0.03	
MSA has rail	26038	0.24		0.29	
Household resides in Big MSAs[1)]	26038	0.15		0.21	
Household resides in Rural[2)]	26029	0.22		0.19	
Household resides in Second City	26029	0.18		0.19	
Household resides in Suburban	26029	0.24		0.22	
Household resides in Town	26029	0.23		0.22	
Household resides in Urban	26029	0.12		0.18	

Notes: 1) Big MSAs include New York, Chicago, Boston, Washington and Philadelphia. 2) The urban/rural dimension variable divides urban form into 5 categories: urban, second city, suburban, town, and rural. The classification is based on the contextual density measure (density in the context of its surrounding area). 3) Variables with missing Std. Dev. are dummy variables.

III. Model

Model Specification

I follow the same methodology as in Brownstone and Golob (2009) to measure the influence of residential density on vehicle and fuel usage by specifying a simultaneous equation model with three endogenous variables and many exogenous variables. The three endogenous variables are housing units per square mile in the census block group (residential density, D), total annual mileage per year of all household vehicles (M), and total household annual fuel consumption per year (F). I also include various socio-demographic and geographic control variables (X). The system can be written as below:

$$\begin{aligned} M_i &= A_{1,3}D_i + B_1X_i + \varepsilon_{1,i} \\ F_i &= A_{2,1}M_i + A_{2,3}D_i + B_2X_i + \varepsilon_{2,i} \qquad (1) \\ D_i &= B_3X_i + \varepsilon_{3,i} \end{aligned}$$

The model above can be written as a matrix form:

$$\begin{aligned} y_i &= Ay_i + BX_i + \varepsilon_i \\ Cov(\varepsilon_i) &= \Omega \end{aligned} \qquad (2)$$

where y_i is the matrix of the three endogenous variables, A and B are coefficient matrices, and ε_i is a vector of residuals with an unrestricted correlation structure. To identify the simultaneous equation, the model is assumed to be a recursive system by restricting the A matrix in (2). I also impose enough restrictions on the B matrix to identify the system, but these restrictions are based on removing several insignificant variables (see Table 5).

The key feature of this model is that it accounts for residential self-selections of households. Residential self-selection occurs if residents of high-density areas differ in some unobservable characteristics that influence travel behavior. This model assumes that household characteristics that influence household residential decision also influence household travel behavior. In the context of the model, there exist residential self-selections if the errors in the residential density equation are correlated with the errors in other equations. However, if we include all relevant variables that influence both the residential density and household travel behavior, the error correlations may be zero even in the presence of residential self-selection.

The model is first estimated using 3SLS with the restriction that error correlations are zero once a rich set

of socioeconomic and geographic exogenous variables are controlled for.[3] Then, all of the over-identifying restrictions (including the restriction that error correlations are zero) pass the specification test described later in this section. It implies that the results cannot reject the null hypothesis that all of the error correlations are zero, consistent with the findings of other studies (Brownstone and Golob, 2009, Bhat and Guo, 2007). However, it does not mean that there is no residential self-selection. Rather, it suggests the necessity of a large set of socio-demographic variables and geographic control variables because these exogenous variables capture the residential self-selection effects.

The recursive system in effect assumes that the household first chooses residential location (D) and then decides annual mileage and fuel consumption. Household annual mileage (M) is assumed to be influenced by the household residential location and various exogenous variables. Household fuel consumption (F) is determined by the household residential location, household annual mileage and various exogenous variables.

There are three paths of influences of residential density on mileage and fuel consumption. First, residential density affects annual mileage because miles per vehicle will be greater due to the separation of households and activity sites and households in lower density areas will choose to have more vehicles, controlling for socioeconomic and demographic differences. Second, residential density directly affects fuel usage in that households that choose to live in denser areas also choose to own more fuel efficient vehicles, partly due to higher costs of maintaining larger vehicles in dense area or the relative difficulty of maneuvering and parking large vehicles. Finally, there is a direct link from mileage to fuel usage. These endogenous effects define a recursive system, so there are no identification problems in the absence of error term correlations.

Weighted Estimation Methodology

In section 2, we observed that the probability of having full information on the key endogenous variables is a decreasing function of household vehicle ownership. Since the number of vehicles in the household is closely related to the endogenous variables in our model, this means that the estimation sample is effectively stratified on the endogenous variables. This non-randomness of missing data may cause biased results and inferences.

There can be a couple of approaches to correct for the bias caused by this non-random missing data

[3] This is actually identical to the OLS estimation by each equation.

problem: the structural approach[4] and the weighting approach. Following Brownstone and Golob (2009), I use the weighted estimation methodology. The weighted estimation is always inefficient, but it does not rely on functional form assumptions that are hard to justify. It also allows easy implementation of error heteroskedasticity.

To implement the weighted estimation, the weights are estimated in a way that the weights compensate for different probabilities of having complete energy information. Since probabilities of having full information for households owning more vehicles are lower, the households with more vehicles must be compensated more and consequently have higher weights. The weights are calculated as the inverse probability of having full information. For example, since 93.5% have full information on the key endogenous variables for households owning one vehicle, the inverse probability (the weight) is 1.07. Likewise, since 29.6% have full information for households owning 5 or more vehicles, the weight is 3.38. The weighted estimator is defined by:

$$\arg\min_{A,B} \sum w_i((I-A)y_i - BX_i)'\Omega^{-1}((I-A)y_i - BX_i) \quad (3)$$

where the weights are denoted by w_i. The covariance of the weighted estimator above is given by:

$$\begin{aligned} V &= \Psi^{-1}\Lambda\Psi^{-1} \\ \Psi &= -E(\partial^2 w_i L_i(\theta \mid y_i, X_i)/\partial\theta\partial\theta') \qquad (4) \\ \Lambda &= E((\partial w_i L_i(\theta \mid y_i, X_i)/\partial\theta)(\partial w_i L_i(\theta \mid y_i, X_i)/\partial\theta')) \end{aligned}$$

where L_i is the log likelihood function, and θ is the parameter vector.

Usual computer programs provide the weighted estimation solution. However, these programs use Ψ^{-1} to estimate the covariance of the estimator, and it is clearly biased. Following Brownstone and Golob (2009), I use the wild bootstrap method (Horowitz, 2002) to generate standard errors for the weighted estimates. This bootstrap works by taking vectors of estimated residuals, e_i, for each observation and multiplying by $(1-\sqrt{5})/2$ with probability $(1+\sqrt{5})/(2\sqrt{5})$ and $(1+\sqrt{5})/2$ with probability $1-(1+\sqrt{5})/(2\sqrt{5})$.

[4] The structural approach proposed by Heckman (1978) specifies a separate binomial probit model of whether the household has complete energy information. Then, the specified equation is added to the original structural system and the equation system is estimated. However, Heckman's method is known to be very sensitive to model specifications because it relies on a joint normality assumption among error terms, which is very strong assumption.

This implies that across the bootstrap repetitions the residuals will have mean equal to e_i and covariance equal to $e_i e_i^{'}$, which is the same approximation used to derive White heteroskedastic-consistent standard errors. This bootstrap procedure has an advantage in that it will yield consistent standard errors even if the errors in the model are actually heteroskedastic. I used 200 iterations to generate standard errors for the weighted estimates.

As it is mentioned earlier in this section, the structural model is imposing restrictions on the coefficients and the covariance matrix. One drawback of using weighted estimation is that, since they are not equivalent to maximum likelihood estimation, standard likelihood ratio tests for over-identifying restrictions cannot be used. Following Brownstone and Golob (2009), a bootstrap test for overidentifying restrictions (including the restrictions on the covariance matrix) is implemented by bootstrapping the variance of the difference between the restricted and the unrestricted reduced forms. The unrestricted reduced form is given by:

$$y_i = C_U x_i + \mu_i \quad (5)$$

The over-identifying (or structural) restrictions are given by:

$$\begin{aligned} C_R &= (I - A)^{-1} B \\ Cov(\mu_i) &= (I - A)^{-1'} \Omega (I - A)^{-1} \end{aligned} \quad (6)$$

The test statistic is given by:

$$(C_R - C_U)^{'} \sum^{-1} (C_R - C_U) \quad (7)$$

where C_R are the restricted reduced form coefficients, are C_U the unrestricted reduced form coefficients, and Σ is the bootstrap variance estimate of $(C_R - C_U)$. If the restrictions are correct, this statistic follows a Chi-squared distribution with degrees of freedom equal to the number of restrictions. The final model presented in the next section cannot reject the null hypothesis of no overidentifying restrictions.

I also implemented a Hausman (1978) test to check whether the assumption of a diagonal covariance matrix is acceptable in cases that all of the over-identifying restrictions (including the restriction that error correlations are zero) did not pass the specification test described above. The Hausman test compares the 3SLS estimates without imposing the restriction of zero error correlations and the estimates from the restriction. Although most model specifications passed the Hausman test, I decided to be conservative and present the model that passes both the bootstrap over-identification test and the Hausman test.

Finally, I implemented another Hausman test of whether the weights are exogenous. This test compares weighted estimates and standard maximum likelihood estimates ignoring the weights. When applied the final model, the test statistic does not reject the null hypothesis that weights are exogenous at 5% confidence level,[5] but this result is very sensitive to slight changes in model specification. While the weighted estimation passes the bootstrap test for overidentifying restriction described above, standard maximum likelihood estimation does not pass the test. I therefore decided to present the weighted estimation for the final results.

IV. Results and Interpretation

1. Estimation Results

The structural equation model was estimated using weighted OLS (i.e., 3SLS with the restriction of a diagonal covariance matrix) and heteroskedascity robust standard errors are achieved by the wild bootstrap method. The best model uses housing density at the census block level as the measure of urban spatial structure, although other six land use density variables also produce similar results. The urban/rural dimension variable is used to control for geographic heterogeneity. Note that the model is estimated under the assumption that the structural errors are uncorrelated. The best model using housing density and the urban/rural dimension variable passes the bootstrap test for over-identifying restrictions including the restriction on the residual correlation matrix at any usual level of confidence.

I also used the MSA category variable[6] and the census region variable[7] to control for geographic heterogeneity. The results are not much different from the result using the urban/rural dimension variable. Nevertheless, they do not pass the bootstrap test for over-identifying restrictions although they pass the Hausman test of whether the diagonal residual correlation matrix is acceptable or not. Thus, I present the results of the best model using housing density and the urban/rural dimension variable. The results for the structural model are given in Table 5.

[5] P-value (prob.>chi2) is 0.9798.

[6] This variable is categorized as 1=MSA of 1 million or more with rail, 2=MSA of 1 million or more and not in 1, 3=MSA less than 1 million, and 4=Not in MSA.

[7] The census region variable is offered by the Census Bureau and it divides states into four groups: northeast, mid-west, south, and west.

Table 6 gives the restricted reduced form coefficients corresponding to the structural model in Table 5. The restricted reduced form coefficients are calculated from the formula $C_R = (I - A)^{-1} B$ in equation (6), where A and B are the structural model coefficients. The reduced form coefficients give the total effects of the exogenous variables on the endogenous variable. The effects of all socioeconomic variables that explain residential density are translated to mileage and fuel usage by the direct effects between the endogenous variables. The total effects of the urban/rural dimension on mileage and fuel usage are also presented as the coefficients.

Table 5: Structural Model Coefficients with Bootstrapped t-statistics in parentheses

Explanatory variables	Household fuel usage per year in gallons	Total Mileage per year on all household vehicles	Dwelling units per sq. mile in units of 1,000 - census block group
Dwelling units per sq. mile in units of 1,000 - census block group	-4.8 (-1.32)	-1499 (-5.76)	
Total Mileage per year on all household vehicles	0.0439 (35.17)		
Annual household income in units of $10,000	8.4 (5.29)	660 (8.24)	-0.021 (-5.07)
Number of children	8.0 (1.69)	-98 (-0.28)	-0.040 (-2.28)
Number of workers	-53.9 (-0.62)		0.052 (1.88)
1-worker household	58.0 (0.65)	4504 (6.09)	
2-worker household	133.8 (0.75)	6867 (6.00)	
3-worker household	181.7 (0.64)	13792 (6.92)	
Number of drivers	21.1 (2.38)	9323 (3.97)	-0.047 (-0.77)
1-driver household		-3221 (-1.36)	-0.369 (-4.34)
2-driver household		-4737 (-1.00)	-0.610 (-4.58)
3-driver household		-7908 (-1.04)	-0.807 (-3.86)
Respondent has college degree	-28.3 (-2.54)		
Respondent has postgraduate degree	-50.2 (-3.34)		
Respondent is retired		-3647 (-3.33)	-0.148 (-2.68)
Single-person household		-2834 (-2.66)	-0.085 (-1.17)
Race is Asian	-48.0 (-3.29)	-2703 (-2.28)	0.461 (3.19)
Race is Hispanic		-1590 (-1.14)	0.402 (3.84)
Race is Black			0.118 (1.74)
Race is mixed White & Hispanic			0.207 (1.90)
Household resides big MSAs[1)]	-24.9 (-2.56)	-2137 (-2.93)	0.554 (10.64)
Household resides in Second City[2)]	-95.8 (-5.85)	-2727 (-2.43)	1.819 (38.59)
Household resides in Suburban	-93.1 (-6.42)	-2699 (-2.19)	1.576 (37.61)
Household resides in Town	-53.4 (-3.49)	-1882 (-2.15)	0.485 (18.61)
Household resides in Urban	-99.3 (-3.50)	-2525 (-1.41)	3.870 (53.78)

Notes: 1) Big MSAs include New York, Chicago, Boston, Washington and Philadelphia. 2) The urban/rural dimension variable divides urban form into 5 categories: urban, second city, suburban, town, and rural. The classification is based on the contextual density measure (density in the context of its surrounding area).

Table 6: Restricted Reduce-Form Coefficients with Bootstrapped t-statistics in parentheses[1)]

Explanatory variables	Household fuel usage per year in gallons	Total Mileage per year on all household vehicles	Dwelling units per sq. mile in units of 1,000 - census block group
Annual household income in units of $10,000	38.9 (10.35)	691 (8.63)	-0.021 (-5.07)
Number of children	6.6 (0.41)	-37 (-0.11)	-0.040 (-2.28)
Number of workers	-57.6 (-0.66)	-78 (-1.64)	0.052 (1.88)
1-worker household	255.8 (2.82)	4504 (6.09)	
2-worker household	435.4 (2.42)	6867 (6.00)	
3-worker household	787.6 (2.60)	13792 (6.92)	
Number of drivers	434.0 (4.30)	9394 (4.02)	-0.047 (-0.77)
1-driver household	-115.4 (-1.15)	-2668 (-1.15)	-0.369 (-4.34)
2-driver household	-165.0 (-0.82)	-3823 (-0.83)	-0.610 (-4.58)
3-driver household	-290.3 (-0.89)	-6697 (-0.90)	-0.807 (-3.86)
Respondent has college degree	-28.3 (-2.54)		
Respondent has postgraduate degree	-50.2 (-3.34)		
Respondent is retired	-149.8 (-3.04)	-3425 (-3.12)	-0.148 (-2.68)
Single-person household	-118.5 (-2.47)	-2707 (-2.53)	-0.085 (-1.17)
Race is Asian	-199.4 (-3.78)	-3395 (-2.77)	0.461 (3.19)
Race is Hispanic	-98.3 (-1.63)	-2193 (-1.59)	0.402 (3.84)
Race is Black	-8.3 (-1.66)	-177 (-1.69)	0.118 (1.74)
Race is mixed White & Hispanic	-14.6 (-1.73)	-310 (-1.69)	0.207 (1.90)
Household resides in big MSAs[2)]	-157.9 (-4.77)	-2968 (-4.01)	0.554 (10.64)
Household resides in Second City[3)]	-344.1 (-7.69)	-5454 (-5.42)	1.819 (38.59)
Household resides in Suburban	-323.0 (-6.42)	-5061 (-4.57)	1.576 (37.61)
Household resides in Town	-170.3 (-3.96)	-2609 (-3.00)	0.485 (18.61)
Household resides in Urban	-483.7 (-8.50)	-8327 (-6.19)	3.870 (53.78)

Notes: 1) The restricted reduced form coefficients are calculated from the equation (6), where A and B are the structural model coefficients. 2) Big MSAs include New York, Chicago, Boston, Washington and Philadelphia 3) The urban/rural dimension variable divides urban form into 5 categories: urban, second city, suburban, town, and rural. The classification is based on the contextual density measure (density in the context of its surrounding area).

2. Interpretation of the results

Relationship among three endogenous variables

The model implies that, if two households are identical in all aspects measured by the socio-demographics variables and residing in the same category of urban/rural dimension, but one household is located in a residential area that is 1000 housing units per square mile denser (roughly 50% of the sample average), the household in the denser area will drive 1500 (7.8%) miles per year less than the household in the less dense area. This is the net effect of vehicle ownership level and trip patterns. I also note that this is the partial effect of residential density within a category defined by the urban/rural dimension variable.

We can compare this result with that in California subsample analysis, which uses the same specification including the urban/rural dimension and imposing the same restrictions on the coefficient matrices as in this study. The California subsample is the data containing 2,079 observations, which was used in Brownstone and Golob (2009). The above coefficient of residential density on household annual mileage is slightly greater than the California subsample result of -1107 (4.4%) in response to a 1000 housing units per square mile density change (roughly 40% of the weighted sample average). The elasticity is almost same between the two cases.

The household in a 1000 housing units per square mile denser area will consume 71 (7.5%) fewer gallons of fuel, and this effect of residential density on fuel usage is decomposed into two paths of influence. The mileage difference of 1500 miles leads to a difference of 66 gallons, which is calculated by multiplying 0.0439 gallons per mile (from the coefficient of mileage on fuel consumption in Table 5) by mileage difference of 1500 miles. However, there is an additional direct effect of density on fuel consumption of 5 gallons per 1,000 housing units per square mile. This is due to the relationship between residential density and fleet fuel economy, a result of vehicle type choice; people in denser area choose more fuel efficient vehicles.

The total impact of density on fuel consumption in national data is slightly greater than the California subsample result of 60 (5.1%) fewer gallons of fuel in denser area. However, the direct effect of density on fuel consumption is not statistically significant at any usual confidence level and it accounts for only 7% (5/71) of the total impact of residential density on fuel consumption. It is much smaller than the California subsample result of 26% (16/60).

Urban/rural dimension and Rail transit dummy

The results above suggest that residential density has only a modest effect on mileage and fuel

consumption (a 50% increase in housing density gives rise to only about 7% decrease in mileage and fuel consumption). However, this is not necessarily the case if both residential density and the urban/rural dimension change simultaneously. This is total impact of urban/rural dimension on mileage and fuel consumption, and it is given in Table 6 as the reduced-form coefficients.

Since the urban/rural dimension is highly correlated with residential density, the total effect of the urban/rural dimension on mileage is due to both a direct effect and an effect channeled through residential density. From Table 5, because urban area is 3870 (71.96)[8] housing units per square mile denser than rural area, and it leads to 5802 (269.9) fewer annual mileages traveled per household. There is an additional direct effect of 2525 (1792) fewer mileages for household living in urban area compared to household in rural area (see Table 5). In total, the households residing in urban area tend to drive 8327 (1345) miles fewer than the household residing in rural area. These total effects are calculated in Table 6. The household living in suburban area drives 3266 (1137) mileages more than household in urban area, and most of them are from the indirect effect channeled through residential density. The direct effect is only -4 (1114) miles which is not statistically significant.

Note that the urban/rural dimension is defined from the contextual density measure (density in the context of surrounding area). For example, even when a certain region is dense, if its surrounding area is not dense, then the contextual density measure may indicate that the region is less dense area. From this fact, we can have more interpretations about the coefficients of the urban/rural dimension in the mileage equation, which is the direct effect of the urban/rural dimension. These coefficients capture the influences of moving a household from rural to other dimensions holding residential density fixed. In the context of this variable, it implies the change in density of its surrounding area holding density of the region fixed. We can also interpret it as if the development around the region has occurred. From the Table 5, we can see that these effects are statistically significant except for the change from the rural area to the urban area.

The effect of the urban/rural dimension variable on fuel consumption is the sum of the direct effect, the effect channeled through mileage, and the effect channeled through residential density. The effect channeled through mileage is obtained by multiplying the total effect of the urban/rural dimension on mileage and the direct effect of mileage on fuel consumption. The effect channeled through residential density is from multiplying the effect of the urban/rural dimension on density and the direct effect of density on fuel consumption. From Table 6, the household living in urban area consumes 483.7 fewer

[8] Standard errors are reported in parentheses in this paragraph.

gallons of fuel compared to household living in rural area. This total effect is decomposed of 99.3 gallons of direct effect, 365.8 gallons of effect channeled through mileage, and 18.6 gallons of effect channeled through residential density.

The effect of rail transit on mileage and fuel consumption can also be decomposed into direct and indirect effects. Surprisingly, the effect of rail transit on mileage is not statistically significant. The indirect effect channeled through residential density does exist, but the magnitude is not large and statistically insignificant. Among the MSAs which have rail transit, several big MSAs including New York, Chicago, Boston, Washington, and Philadelphia, were separated into a new dummy variable and tested. As a result, the big MSA dummy has significant influences on both mileage and fuel consumption, unlike the rail transit dummy. From Table 6, the household living in these big MSAs is predicted to drive 2968 (15% of the sample average) miles less and consume 158 (17%) fewer gallons of fuel.

Simulation using the urban/rural dimension

Total effect of the urban/rural dimension can be interpreted as the impact of a program that alters urban form from a category of the urban/rural dimension into another category. As described above, this program affects more than the sole residential density measure. The result of this experiment is summarized in Table 7. Several results stand out. When we move a household from the rural to urban area, the effect is to reduce annual household mileage by 35%. The household living in urban area is predicted to drive 15% fewer miles than the household living in suburban area. When we move the household from the rural to suburban area, the household will drive less by 22%. All these effects of category changes are statistically significant except for the change from suburban to second city category.

Percentage change in fuel consumption is slightly larger than percentage change in mileage in every dimension reflecting the tendency toward more fuel efficient vehicle choices in denser area. All these percentage changes increase if the households move from denser areas to less dense areas because of low base effect when we calculate percentage. The effects of this program is very similar to the results of Bento et al. (2005), where the authors implemented a similar simulation that moves a household from a city which has the same characteristics as Georgia to a city which has the same characteristics as Boston.

To compare the effect of this program with the California subsample result, I included the urban/rural dimension and imposed same restriction into the model using the California subsample. The effects of this program using California data are much smaller than the results using national data. The effects of category changes are only about half of those of national data. In California, the effects of category

changes are frequently statistically insignificant, unlike the results from national data. This finding supports the statement that there is no need to control for geographic heterogeneity in California analysis while it is necessary in analyzing national data. The effects of this program in California data is given in Table 8.

Table 7: Change in mileage and fuel consumption when the category of urban/rural dimension changes (national level data)

Change in mileage when urban/rural dimension changes

From	Urban		Second City		Suburban		Town		Rural	
To	Mileage change	%change	Mileage change	%change	Mileage change	%change	Mileage change	%change	Mileage change	%change
Urban			-2873 (1114)	-17.83%	-3266 (1137)	-14.74%	-5718 (1115)	-24.76%	-8327 (1345)	-35.17%
Second City	2873 (1114)	28.45%			-393 (1007)	-1.77%	-2845 (1007)	-12.32%	-5454 (1006)	-23.03%
Suburban	3266 (1137)	32.34%	393 (1007)	2.44%			-2453 (917)	-10.62%	-5061 (1108)	-21.38%
Town	5718 (1115)	56.62%	2845 (1007)	17.65%	2453 (917)	11.07%			-2609 (870)	-11.02%
Rural	8327 (1345)	82.45%	5454 (1006)	33.84%	5061 (1108)	22.85%	2609 (870)	11.30%		

Change in fuel consumption when urban/rural dimension changes

From	Urban		Second City		Suburban		Town		Rural	
To	Fuel usage change	%change	Fuel usage change	%change	Fuel usage change	%change	Fuel usage change	%change	Fuel usage change	%change
Urban			-140 (45.24)	-18.20%	-161 (45.53)	-15.24%	-313 (46.13)	-27.79%	-484 (56.88)	-40.49%
Second City	140 (45.24)	29.57%			-21 (42.35)	-2.00%	-174 (43.43)	-15.41%	-344 (44.75)	-28.81%
Suburban	161 (45.53)	34.04%	21 (42.35)	2.75%			-153 (41.71)	-13.54%	-323 (50.33)	-27.04%
Town	313 (46.13)	66.39%	174 (43.43)	22.66%	153 (41.71)	14.48%			-170 (43.03)	-14.26%
Rural	484 (56.88)	102.46%	344 (44.75)	44.86%	323 (50.33)	30.63%	170 (43.03)	15.10%		

Notes: 1) Standard errors are reported in parenthesis

2) Percentage changes are calculated based on the mean of each category of the urban/rural dimension

Table 8: Change in mileage and fuel consumption when the category of urban/rural dimension changes (California subsample)

Change in mileage when urban/rural dimension changes

From	Urban		Second City		Suburban		Town		Rural	
To	Mileage change	%change	Mileage change	%change	Mileage change	%change	Mileage change	%change	Mileage change	%change
Urban			-2089 (1410)	-9.90%	-2284 (1945)	-9.49%	-2476 (1897)	-9.85%	-2886 (2044)	-12.49%
Second City	2089 (1410)	11.09%			-194 (1756)	-0.81%	-387 (1763)	-1.54%	-796 (1909)	-3.45%
Suburban	2284 (1945)	12.12%	194 (1756)	0.92%			-193 (2136)	-0.77%	-602 (2511)	-2.61%
Town	2476 (1897)	13.14%	387 (1763)	1.83%	193 (2136)	0.80%			-409 (2120)	-1.77%
Rural	2886 (2044)	15.32%	796 (1909)	3.77%	602 (2511)	2.50%	409 (2120)	1.63%		

Change in fuel consumption when urban/rural dimension changes

From	Urban		Second City		Suburban		Town		Rural	
To	Fuel usage change	%change	Fuel usage change	%change	Fuel usage change	%change	Fuel usage change	%change	Fuel usage change	%change
Urban			-149 (64.09)	-14.68%	-114 (71.91)	-10.02%	-180 (86.12)	-14.56%	-210 (87.19)	-18.34%
Second City	149 (64.09)	17.09%			36 (73.02)	3.12%	-30 (86.39)	-2.46%	-61 (89.80)	-5.29%
Suburban	114 (71.91)	13.03%	-36 (73.02)	-3.49%			-66 (87.57)	-5.34%	-96 (100.10)	-8.39%
Town	180 (86.12)	20.57%	30 (86.39)	2.99%	66 (87.57)	5.80%			-30 (97.15)	-2.64%
Rural	210 (87.19)	24.03%	61 (89.80)	5.96%	96 (100.10)	8.45%	30 (97.15)	2.45%		

Notes: 1) Standard errors are reported in parenthesis

2) Percentage changes are calculated based on the mean of each category of the urban/rural dimension

Socio-demographic variables

Number of Drivers

The number of drivers has a strong influence on household travel behavior. Considering the effect of number of drivers on residential density, the total effect can be decomposed into direct and indirect effects. The total effects on each of the three endogenous variables are nonlinear, as captured by up to four variables: a continuous "number of drivers" variable, and dummy variables for one-driver, two-driver and three-or-more-driver households.

Drivers per household have a negative diminishing marginal effect on choice of residential density. For example, all else held constant, the model predicts that a household with one driver will locate in a residential area that is less dense by 416 (22% of the sample average) dwelling units per square mile, when compared with a household with no drivers; a household with two drivers will locate in a residential area that is less dense by about 288 (15%) dwelling units per square mile, when compared with a household with one driver; and the difference in density between two- and three-driver households declines to about 244 (13%) dwelling units per square mile.

The influence of drivers per household on annual vehicle usage also exhibits such diminishing marginal effect till 3-driver household, but from three to four it does not hold. Based on the restricted reduced form results in Table 6, adding the first driver in the household increases annual mileage by 6725 (35% of the sample average), and adding an additional driver leads to an additional 5571 (29%) miles per year. From two to three drivers per household the added mileage per year is 2697 (14%) miles, and from three to four it is 9394 (48%) miles. The effects of the number of drivers on fuel usage follow the same trend. These patterns of diminishing marginal effects are similar to the California subsample result, but the magnitudes of the effects are slightly lower than the California subsample result.

Number of workers

As in the previous case of household drivers, the total effects of number of workers on annual mileage and fuel usage are both nonlinear, each being captured by three variables: a continuous variable and dummy variables for one-worker, two-worker and three-or-more-worker households. Based on the restricted reduced form results in Table 6, adding the first worker in the household increases annual mileage by 4425 (23% of the sample average), and adding an additional worker leads to an additional 2285 (12%) miles per year. From two to three workers per household the added mileage per year is 6847 (35%) miles. In contrast to the number of drivers, the greatest marginal effect for number of workers is

the difference in mileage and fuel consumption attributed to the difference between two to three workers. This pattern can be applied to the California subsample result from the model using same specification and magnitudes of the effects are also similar.

Income

The model predicts that fuel usage increases linearly with income, and this is caused by all three factors. Higher income translates into: (1) choice of lower density residential location, (2) greater total driving distances, which is independent of the greater distances caused by lower densities, and (3) lower overall fuel economy of the household fleet.

All these effects are statistically significant. Other things equal, if one household earns $10,000 higher income annually, the household of higher income will consumes about 40 (4%) more gallons of fuel per year. It is almost twice greater than the California subsample result of 26 (2.2%) more gallons for a household of higher income per year. The effect of (2) is the largest among the three factors mentioned above and it is over twice larger than the California subsample result. The effect of (1) is almost similar to the California result. The direct effect of income on fuel usage (3) is still significant, but it is smaller than the California subsample result, which suggests that the influence of income on vehicle choice is larger in California than for the average U.S. household.

Number of children

Fuel usage is expected to increase with the number of children due to three factors. (1) Larger families tend to choose lower residential density, which in turn increases total mileage. (2) All other things equal, large family and more children tend to drive more. (3) Fuel economy decreases as a function of the number of children, due to increased likelihood of a least one van or SUV in the household fleet. The effect of (1) is significant, but is smaller than the California result. The effect of (2) is insignificant, and the effect of (3) is also insignificant, in contrast to the California result, where the effect is significant. Overall, the influence of the number of children on annual mileage and fuel consumption is statistically insignificant, unlike the California subsample result using the model using same specification.

Education

The influences of education level on residential density choice and mileage traveled were dropped

because of insignificance, so we can capture only the effect of education on vehicle type choice. Only two education dummy variables are found to be significant. Households headed by a respondent with a college degree tend to have a vehicle fleet with lower fuel economy than their less educated counterparts. This effect is accentuated if the household is headed by a respondent with a postgraduate degree. Although magnitudes are slightly lower than the California subsample result, the pattern is similar.

Life Cycle Effects

Retired households tend to live in lower-density residential areas, which imply that retired households tend to drive more. However, there is a negative direct effect of retired household on mileage traveled, which is greater than the indirect effect channeled through residential choice. Summing up these two effects, the retired household tends to drive less.

The influence of a single-person household on residential density is not statistically significant. However, the influences on household annual mileage and fuel consumption are significant, implying that single-person households tend to drive less than non-single-person households.

Race and Ethnicity

Four race and ethnicity variables are included in the residential density equation. Households which are solely Black, solely Asian, solely Hispanic, or mixed White and Hispanic, all tend to reside in higher-density areas, compared to other households, predominant solely White households. This leads to lower vehicle usage and fuel consumption for all of these groups. There are statistically significant direct effects on mileage and fuel consumption decisions only for solely Asian. In contrast to the California result, solely Hispanic does not have any direct effects on mileage and fuel consumption.

V. Conclusion and Further Research

This paper measures the impact of residential density on vehicle usage and fuel consumption following the same methodology as Brownstone and Golob (2009), but using national level data from the 2001 NHTS. To account for residential self-selection effects, a simultaneous equation model is specified and OLS (3SLS with the restriction of a diagonal covariance matrix) is used for estimation. A weighting

approach is used to correct for the bias caused by non-random missing data. After controlling for various socio-demographic and geographic locations, the model cannot reject the null hypothesis that there is no significant self-selection effect, which is consistent with Bhat and Guo (2007) and Brownstone and Golob (2009).

I find that residential density has a statistically significant but economically modest influence on household travel behavior. However, the contextual density measure (also defined by the urban/rural dimension) has a considerable impact on household annual mileage and fuel consumption. A program that alters urban form across the urban/rural dimension affects household vehicle usage significantly.

Compared to the California subsample analysis, which uses the same specification as in this study, the impact of residential density on mileage and fuel consumption is slightly bigger. The direct impact of residential density on fuel consumption is statistically insignificant unlike, the California subsample result. I also find that the urban/rural dimension does not have an explanatory power in explaining household vehicle usage in California data, which is different from the result using national data.

This research can be usefully extended in several directions. First, there might be some ways to overcome the limitations from using residential density as the measure of urban sprawl. Considering that the phenomenon of urban sprawl has a dynamic aspect, new measures can be devised using the information on urban development of the U.S since World War II. Although the urban/rural dimension was devised to consider the density surrounding the area, as it is pointed out by the devisers, the definition of each category is ad-hoc. To overcome this problem, adjunct geographic location information can be merged into the NHTS dataset to provide more information about the households' neighborhood characteristics.

Second, considering that the travel demand is also multidimensional in nature, we can test the model with different selection nodes other than household mileage traveled and fuel consumption. Especially, this paper assumes that household mileage is just the summation of all cars a household owns. However, the decision of mileage traveled can be a different choice node from purchasing and owning cars. In that case, the number of cars can be also regarded as another endogenous variable. The daily travel data set, which is included in the NHTS set, can be also used for further analysis of travel behavior. If it is possible to obtain information on accessibility to public transportation for those households in major metropolitan areas, a model can be developed to jointly determine public transit accessibility along with residential density and transportation energy use. Finally, we can investigate historical patterns of travel behavior by utilizing the 2008 NHTS, which is supposed to be released in the near future.

REFERENCES

Badoe, D. and E. Miller. 2000. Transportation-land use interaction: empirical findings in North America, and their implications for modeling. Transportation Research D 5, 235-263.

Bento, A.M., M.L. Cropper, A.M. Mobarak and K. Vinha, 2005. The impact of urban spatial structure on travel demand in the United States. Review of Economics and Statistics 87, 466-478.

Bhat, C.R. and J.Y. Guo. 2007. A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. Transportation Research B 41, 506-526.

Boarnet, M. and R. Crane. 2001. The influence of land use on travel behavior: specification and estimation strategies. Transportation Research Part A 35, 823-845.

Brownstone, D and Golob, T.F. 2009. The impact of residential density on vehicle usage and energy consumption. Journal of Urban Economics 65, pp. 91-98.

Brueckner, J. K. 2001. Urban sprawl: lessons from urban economics. Brookings-Wharton papers on urban affairs.

Ewing, R. 1997. Is Los Angeles Style Sprawl Desirable? Journal of the American Planning Association, 63, 107-126.

Gordon, P. and H.W. Richardson. 1997. Are Compact Cities a Desirable Planning Goal? Journal of the American Planning Association, 63 95-106.

Handy, S. 1996. Methodologies for exploring the link between urban form and travel behavior. Transportation Research Part D 35, 823-845

Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 46, 1251-1271.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47, 153-161.

Miller, D. R. and Hodge, K. 1994. A Population Density Approach to Incorporating an Urban-Rural Dimension into Small Area Lifestyle Clusters. Paper presented at the Annual Meeting of the Population Association of America, Miami, Florida

Newman, P. and J. Kenworthy. 1999. Costs of automobile dependence: global survey of cities. Transportation Research Record 1670, 17-26.

Horowitz, J. L., 2002. The bootstrap, in E. E. Leamer and J.J. Heckman (Eds.), Handbook of Econometrics, Vol. 5, Elsevier, Amsterdam, Chapter 52.

Kahn, M. E., 2000. The Environmental Impact of Suburbanization. Journal of Policy Analysis and Management 19, 569-586.

Wooldridge, J.A. 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge, MA.