Title

A California Statewide Exploratory Analysis Correlating Land Use Density, Infrastructure Supply and Travel Behavior

Authors
Yoon, Seo Youn
Golob, Thomas F.
Goulias, Konstadinos G.

Publication Date
2008-10-01

A California Statewide Exploratory Analysis Correlating Land Use Density, Infrastructure Supply and Travel Behavior

Seo Youn Yoon
Department of Geography & GeoTrans Laboratory
University of California, Santa Barbara

Thomas F. Golob
Institute of Transportation Studies
University Of California Irvine

And

Konstadinos G. Goulias
Department of Geography & GeoTrans Laboratory
University of California Santa Barbara
Santa Barbara CA 93106
805-284-1597
Goulias@geog.ucsb.edu

Abstract: In this paper land use densities by type of employment and infrastructure supply are used together with social and demographic characteristics to explain non-motorized travel, transit use, and solo driving in California. The land use database, the highway network database, and the travel survey used for the analysis here covers the entire state of Califonia. Land use and infrastructure have a significant, substantial, and very different role for each behavior indicator used here. They alternate in significance and importance depending on the specific behavior analyzed. We also performed experiments to identify the appropriate geographical aggregation by comparing US Census tract vs US Census block group based land use densities and infrastructure densities. Regression models gave us mixed results leading us to suggest the use of a combination between the two geographies. Next steps are also outlined in the paper.

Version July 8, 2008

Paper submitted for presentation at the 2009 TRB 88th Annual Meeting and Publication in the Transportation Research Record

1. Introduction

Indicative of the green house gases (GHG) policy momentum is a comment by California Senator Darrell Steinberg to the Planning Report (www.planningreport.com) in July 2008 about legislation aiming at reducing green house gas emissions:

"...I think everyone recognizes that in order to meet the requirements of this very aggressive law (AB32), we're going to have to employ market-based mechanisms, we're going to have to regulate, and we're going to have to focus on mobile sources, stationary sources, and land use, which is the subject of SB 375..."

Independently of the type of mechanisms employed to control GHGs, transportation policy analysis is based on relationships among land use, infrastructure available, and travel behavior while at the same time accounting for a variety of other relationships including demographic, social and economic circumstances of households and their members. Examples include earlier systemwide models reported in (http://ntl.bts.gov/DOCS/ornl.html) and later urban simulation models as in Miller (2003). These relationships are increasingly analyzed using microdata to correlate neighborhood characteristics to travel (Kitamura et al., 1997, Khattak and Rodriguez, 2005), identify more complex relationships the impact of accessibility on commuters behavior (Abreu et al., 2006), examine behavioral aspects of special groups such as the baby boomers (Goulias et al., 2007). Many strong significant relationships are found and a review is reported in (http://www.fhwa.dot.gov/tcsp/selrefev.html) but sometimes under specific circumstances research finds weak relationships between land use and travel (Boarnet, 1996). These micromodels are already used in regional modeling and simulation and there are many urban and regional studies that can provide a good starting point in the US. However, detailed statewide analyses are only emerging recently. Some strategies/mechanisms advocated in the political realm aim at individuals and their households. Policy support systems that operate at the level of individual and household decision making and are able to make informed assessments of policy action are currently assessed (see http://www.dot.ca.gov/hq/tpp/offices/ocp/projects.html) and they are developed mostly for Metropolitan Planning Organizations in the context of their long range transportation plans (see http://trb-forecasting.org/innovationsConference2008.html). Statewide models that are able to relate land use, travel behavior, and environmental impacts are also developed along the path created by the Oregon travel model (see the 1998 TRB statewide modeling conference http://onlinepubs.trb.org/onlinepubs/circulars/ec011/donnelly.pdf).

In this paper taking advantage of the 2000-2001 California Statewide household travel survey and a variety of land use and infrastructure indicators we provide an exploratory analysis that examines the relationship among strategically selected travel behavior indicators, land use indicators, and supply of highways. We aim at three objectives:

(1) Offer a feasibility test and insights about these relationships using a statewide sample of households;
(2) Illustrate possible ways to correlate land use indicators and supply of highways with travel behavior indicators; and
(3) Show if measurement at different geographic aggregation units leads to different conclusions about these relationships.

These findings will guide future efforts in California to develop models that will be used in planning simulation software to assess the impacts of the planned legislative initiatives. In the next section we provide a short description of the data used. This is followed by a sample of model estimation and findings. The paper concludes with a summary.

2. Data Used

In this analysis the state of California is divided into two different sets of geographical units. The first set contains 7,049 zones using the US Census 2000 tracts and the second contains the smaller geographical areas that are 22,133 zones using the US Census 2000 block groups. The Census tract (unit of analysis here) is selected as a first order geographical subdivision to make the analysis tractable at the state level and to provide sufficient detail to be meaningful. Table 1 contains a selection of tract analysis characteristics. Treating each tract as a traffic analysis zone we attach the tracts to a network dataset that contains the entire spectrum of roadways in California from local roads to interstate freeways. As indicators of available opportunities in a tract, numbers of workers classified according to the North American Industry Classification System (NAICS) were used. The original NAICS classification of fourteen types of industries was aggregated into five types: retail, health, services, manufacturing, and all other considering the types of activity in which people can participate related to the industries. Using Geographic Information Systems we compute for each tract the amount of activity opportunities reachable within 5 km, 5 to 10 km, and 10 to 50 km. We repeat the same for 20 minutes and 20 to 40 minutes travel time computed using information about speed limits on the roadway network at hand. For each geographical unit we create sums of the length of roadway segments by type in a series of concentric areas that are accessible in 20 minutes, 40 minutes, and 60 minutes of travel time to quantify roadways that are available from an origin (the centroid of a census tract polygon). We name these areas the *buffers*. The types of roadways we count for each buffer are: primary highways with limited access (*primary lim* herein), primary roadways without limited access (*primary nolim* herein), secondary and connecting roadways (*secondary* herein), local and rural roads (*local* herein), roads with special characteristics (*special* herein), all other roadways (*other* herein). Highways are counted based on lane kilometers by type (e.g., limited access freeways/motorways, secondary roads connecting limited access roadways, local roads). This same process is repeated for the block groups. In the second quadrant of Table 1 we provide a statewide summary of the block groups characteristics. A selection of these characteristics is used in regression models that are reported in later sections of this paper.

When considered separately, sociodemographic characteristics, spatial accessibility, and road infrastructure all influencing travel behavior. Dense urban areas make walking trips more feasible; extensive network density of freeways and arterials encourages vehicular trips; large households make more trips per day than small households, and so on. However, in the real world, all of these variables interact simultaneously: Households consider the costs and benefits of different locations and feasible travel modes in the light of their circumstances, and choose residential locations accordingly. Indeed, one could argue that households are not merely reacting to their circumstances, but rather are actively trying to improve their lot in any way they can. Adjustment strategies include moving residence, changing jobs, choosing different travel destinations, bundling individual single-occupancy vehicle (SOV) trips into high-occupancy vehicle (HOV) trips, and so on. One cannot merely consider the influence of spatial infrastructure characteristics in isolation. A household travel survey can provide the base information explaining how they impact household decisions on trip making along with other variables such as sociodemographic characteristics.

The California Statewide Travel Survey, conducted over several months in the years 2000 and 2001, provides an excellent starting point for disentangling the relationships between space, infrastructure, and sociodemographics. The survey sample, consisting of more than 17,000 households, is a quota sample by county and planning region, rather than a representative sample of California proportional to the population of each county. Each trip destination has been geocoded, usually to the nearest intersection, but sometimes to the approximate census tract centroid or block group centroid. The location of almost every household can also be determined from the survey data. A small selection of household characteristics which is available and used here is presented in Table 2. To these data have been added spatial accessibility variables and roadway infrastructure variables by census tract and block group described above. In this way, a sample of 17,040 households with a relatively even distribution across all California counties is created and contains a variety of travel behavior variables from a wide variety of spatial environments.

3. Models

In each of the models that follow, three blocks of variables are tested: (1) the same set of sociodemographic variables, (2) residential and activity site density variables, and (3) any road infrastructure variables found to be significant in explaining the dependent travel behavior variable after controlling for the first two sets of variables. For example, one variable is the retail employees within 10 km of a census tract. From the several measures developed using both time and distance to define the boundary of the buffer, only the 10 km and 50 km buffer variables are used here. Shorter time buffers could have been used and would have produced similar results, but these distances were found to be more effective in capturing the influence of infrastructure provision and access to activity opportunities. The shortest distance buffer zones are tested both in direct and difference (ring) format.

Modeling the contribution of spatial accessibility and infrastructure density was further complicated by the presence of spikes at zero and long positive tails. For example, some rural census tracts in California are extremely large with a very small population concentrated in a small portion of the tract. These need to be modeled together with census tracts that have some of the highest densities of roadway infrastructure in the nation. To overcome this distributional heterogeneity, spatial variables were converted to a scale in which the population was ranked into ten groups of equal frequency (deciles). This relieves the estimation bias caused by outlying observations and restrictions to the positive domain with spikes at zero value. It also facilitates estimation in which the spatial variables can contribute nonlinear and even non-ordinal effects.

We present omnibus tests of each set of variables, but the variable coefficients are shown only for one model for illustration purposes. These coefficients are displayed as odds ratios; the raw coefficient can be computed as the natural logarithm of the odds ratio. To aid in interpretation, only statistically significant (p = .05) coefficients are listed. All variables are categorical, and the continuous spatial variables are discretized into ten equal categories (deciles).

In the following sections we present the results of three sets of models (for nonmotorized travel, high occupancy vehicle usage, and solo driving) aimed at assessing the influence of the spatial environment on travel demand in California. The first set of models deals with nonmotorized travel. We estimate separate models for nonmotorized travel by any household member and by adult drivers only. In the second part, similar models are estimated for transit travel The final set of models is for solo driving, with one model for household solo driving demand, and one model for solo driving distance.

The same procedure of variable computation was conducted using block groups, which are smaller than census tracts, and the same models were built using the block group variables. We also analyze the impact of spatial aggregation level (e.g., census tract level measurement vs block group level measurement) on the explanatory power of the models. Artificial boundaries imposed on continuous geographical phenomena, such as accessibility, result in the generation of artificial spatial patterns, and the spatial patterns generated in different levels of spatial aggregation differ from each other. This is called the modifiable areal unit problem (MAUP; Openshaw and Alvanides, 1999). We analyze the existence and the impact of MAUP in the three sets of travel behavior models and show how spatial variables at different aggregation levels can be used in the models to mitigate this artificial spatial resolution considering the impact of unit area sizes. In the models that follow we show estimation results using census tract and block group accessibility variables and household sociodemographics and their implication on travel behavior.

3.1 Nonmotorized Travel

Of our 16,750 households with complete data (98.3% of the sample), 14.2% had a household member that made at least one trip walking or by bicycle. The highest concentration of these households was in the San Francisco Bay Area, where 25.9% of the households in this survey recorded a nonmotorized trip segment, followed by Santa Barbara County, with 19.2% of households. Spatial factors are important in nonmotorized travel demand (Table 3) under any combination of spatial variables specification. The second quadrant of Table 3 shows the chi-square contribution of spatial variables measured at the block group level. In the third quadrant we show the impact of combining in the model specification variables measured at the census tract with variables measured at the block group level. The combination provides a slightly better fit using the same amount of degrees of freedom. It is also important to note the virtually unaltered chi-square contribution of sociodemographics between the two specifications.

The sociodemographic predictors of household nonmotorized travel include the variables listed in Table 2. As expected, the presence of children older than 6 increases the likelihood of a household making a nonmotorized trip, while the presence of very young children decreases that likelihood. Lower income and the youngest households are more likely to make nonmotorized trips, but so are the most highly educated households. With regard to influences of the built environment on nonmotorized travel (Tables 4 and 5), the “rural” effect is very interesting for nonmotorized trips than for public transport trips. Here low housing density produces a lower propensity for nonmotorized trips, confirming that extreme distances among activities inhibit the use of slower modes. It is possible that for some households rural transit trips are taking the place of rural nonmotorized trips. In terms of road infrastructure, Table 5 shows that the lower percentiles of local roads within 10 to 50 km have much higher propensity for nonmotorized trips, as is the case for transit shown later. Higher levels of road infrastructure (local roads within 10 to 50 km) correspond to lower levels of nonmotorized trips. Both of these effects are perhaps related to using nonmotorized trips as a form of recreation, as it is pleasant to walk or bike in less developed, low traffic areas, while it is both unpleasant and dangerous to walk or bike in highly developed, high traffic areas.

It should also be noted the contribution of household density is larger in the census tract model, and the contributions of the other variable sets are larger in the block group model for household nonmotorized travel demand. Moreover, the block group model also shows that low household and retail employee density produce a lower propensity for nonmotorized trips, but the impact of retail employee density of 10 to 50 km distance is clearer in the block group model. The influence of primary roads without limited access within 10 to 50 km, which was not significant in the census tract model, is found to be significant in the block group model. Instead of local roads within 10 km which were significant in the census tract model, local roads within 10 to 50 km were found to be more significant in the block group model. This type of reasoning and cross-comparisons lead us to believe that a regression models that uses densities measured at the tract level and infrastructure measured at the block group may yield better fit and more meaningful findings.

The second model we examine in nonmotorized travel is nonmotorized travel by adult drivers that were used as a specialized subset to control for the various inevitable nonmotorized travel by other household members such as walking and biking by people under driving age. Of the 14,160 households with adult drivers, 10.4% had an adult driver that recorded at least one nonmotorized trip segment. For Bay Area households the split is 23.6% and it is 6.3% for Santa Barbara County. As in the models for nonmotorized travel by any household member, spatial density variables are very important in this case and they are more important than sociodemographic factors, when compared by the goodness-of-fit contributions (Table 6).

Road infrastructure also plays a relatively important role. The sociodemographic predictors of nonmotorized travel demand by adult drivers in the household are similar to the predictors of nonmotorized travel demand by any household member, with the notable exception of the influence of older children. As expected, such children travel by bicycle and walking, but their presence actually decreases the likelihood that adult drivers in the household engage in such trips. Also, there is no statistically significant difference in the propensity for nonmotorized trips among adult drivers in older households, indicating that the lower propensity observed earlier for these households is likely due to the immobility of non-drivers in such households.

The spatial density factors influencing nonmotorized travel by adult drivers indicate a similar pattern as for all nonmotorized travel, with a considerably enhanced positive effect of regional retail accessibility on nonmotorized travel demand by adult drivers. Likewise, the influences of road infrastructure are accentuated in the case of nonmotorized travel by adult drivers. This shows that adult drivers are largely responsible for the effects of the built environment on demand for nonmotorized travel discussed in the previous section. It is not surprising that nonmotorized travel by non-drivers is less sensitive to spatial factors. Less likelihood of nonmotorized travel by adult driver in higher levels of road infrastructure is also picked up in the block group models. But instead of primary roads without limited access within 50 km, the influence of primary roads without limited access within 10 to 50 km is found to be significant, and the impact is reversed when it was measured using block groups. The impact of local roads within 10 to 50 km is shown more clearly in the block group model. It implies that block group model can be better for discerning different impact of smaller segment of the space. The impact of combining spatial variables from the two levels of spatial aggregation is similar to the previous non-motorized model of Tables 3, 4, and 5 but this time the degrees of freedom are the same between the model that uses only block group level independent variables and the model that uses a combination of block group level with census tract level variables due to the inclusion of more variables.

3.2 Transit Travel

In a similar way as for nonmotorized, the analysis here is done first with a model of transit usage by any household member and then with a model of transit usage by any adult driver in the household. Transit use is defined as taking any local transit mode, including bus, rail, and light rail, but not including long distance bus trips. School bus trips are also included as household public transport trips. Of the 16,750 households with complete data (98.3% of the sample), 8.1% had a household member who made at least one trip by public transport (transit); the highest concentration of these households being in the San Francisco Bay Area, where 14.4% of households in this sample had transit users.

Socioeconomic factors are effective in explaining which households are transit users, but some spatial density factors and one set of road infrastructure variables are also important. Transit usage is a decreasing function of income, and an increasing function of household size. Transit usage is generally a decreasing function of age of the household head(s), but usage is greatest for the second youngest group, and lowest for the second oldest group. Transit services for the elderly probably increase the likelihood of transit usage for households with the oldest household heads. Education is not an effective predictor of transit usage, and only one ethnicity category is important: black households are 1.6 times more likely to use transit. Regarding children, households with only young children are less likely to use transit, while those with older children are more likely to use transit. Spatially, as expected, transit-using households are concentrated in the densest 10% of residential areas, and also in the least dense 20% of areas. But excluding areas in the highest 10% of housing density, households located in areas above median density are less likely to use transit. Census tracts with low density housing tend to be located in rural counties. While the presence of school age children in the household coupled with the inclusion of school bus trips as public transit trips may account for some of this effect, this result underscores the importance of rural public transport. Accessibility to retail services, particularly accessibility at the regional level (10 to 50 km), indicates lower transit usage for households located in low accessibility areas, and high transit usage for households located in the highest 10% of retail accessibility. This effect undoubtedly captures the urban core phenomenon. The influence of road infrastructure is complex. Controlling for sociodemographic factors and spatial density, households that live in areas in the lower quartile of regional primary surface road coverage (primary roads without limited access within 10 to 50 km of network distance) exhibit the highest transit usage, together with households in the 80th percentile. However, households above the 90th percentile have very low transit usage. Once again, the importance of rural public transport is picked up by the road infrastructure variable, even when controlling for housing and retail density. In tracts with both low housing density and lower levels of road infrastructure, the likelihood of transit usage is unusually high.

Using the block group level variables as explanatory variables in the model for transit usage, household density contributes slightly more to the model when it is measured at the census tract level, and the other spatial variable sets – retail employee density and road infrastructure - contribute more to the model when they are measured based on block groups (Table 7). Especially, the road infrastructure in the block group model contributed almost twice as much as in the census tract model in terms of chi-square. The spatial density variables show similar impact pattern on household transit usage in the block group analysis, too. However, in the block group model, the concentration of transit usage in the highest density area is stronger and the concentration in 10^{th} percentile of household density is not captured. The highest percentile of the block group retail employee density had higher impact in both buffers (0 to 10 km and 10 to 50 km). This can be a typical influence of MAUP. First, different sizes of unit area produce different statistics, household density in this case, and they reveal different patterns of influences. The patterns can have different impact in the models as the variable sets do in the Logit model of household transit use. Second, different levels of spatial aggregation lead different levels of approximation of the explanatory variables. From the comparison between the two models of household transit use, it appears that a better approximation of an explanatory variable by going one level of disaggregation down (from census tract to block group) improves the contribution of the independent variables by explaining variation in the dependent variable.

The influence pattern of road infrastructure of the block group model is similar to that of the census tract model, but in addition to primary roads without limited access within 10 to 50km, which was the only road infrastructure variable set significant in the census tract model of household transit usage, local roads variables were found to be significant in the block group model. In the block group model, the importance of rural public transportation is also picked up, and the likelihood of transit usage is low in the households which belong to the highest 10% road network areas. The first quadrant of Table 7 shows the fit comparison. Combining tract and block group level variables was not found a wise strategy in this case.

Analyzing transit trips made by *any* household member can be difficult to interpret, as children and non-driving adults may be skewing the results for some households but not others. The next model describes transit usage by adult drivers, being those adults who were either recorded as having a driver's license, or else were observed to have driven at least once. Only 2.7% of the 14,160 households with adult drivers and complete data have an adult driver that makes at least one transit trip.

As expected, it is much more difficult to predict which households these are, based on sociodemographic factors, as seen by comparing the goodness-of-fit log-likelihood-ratio model Chi-square statistics and the pseudo-R^2 indices in Table 7. However, spatial density is relatively more important in the case of adult drivers, and the same set of road infrastructure variables is also significant. The sociodemographic predictors of transit usage by adult drivers shows that such usage is concentrated in low income households, larger households, households in the middle age groups (35 to 55), black households, and more highly educated households. This effect probably captures central business district employment. Households less likely to have adult driver transit usage are high and middle income households, small households, households with heads in the 65-75 year range, lower educated households, and households with children. The effects of rural public transport (tracts with low density housing and road infrastructure) disappear when the focus is restricted to adult drivers. Moreover, households that live in areas with the highest residential and retail density are the heaviest transit users. The phenomenon of low relative transit usage households in the 90th percentile of regional primary surface road coverage still prevails (compared to the previous transit model). Households in the 90th percentile of regional primary arterial coverage are concentrated in Orange, Los Angeles, and San Mateo, and Alameda Counties, but there are also such households located in San Bernardino, Santa Clara, Riverside and Ventura Counties. An abundance of primary arterials appears to correlate with fewer household transit trips in these areas.

Shifting our attention to the models with block group variables, we can see the influence of using smaller unit areas in this comparison, too. Household density contributes more to the model than when it is measured in census tract level, and the other spatial variables contribute more to the model when they are measured in block group level. For example, the likelihood of transit usage by adult drivers is relatively low in the households in the 90th percentile of primary and local roads coverage as shown in the census tract model. However, the local roads variable set in the block group model still show the effect of rural public transport usage by adults drivers and the 70th percentile of primary road infrastructure had positive impact in the block group model, which couldn't be seen in the census tract model. The likelihood of transit usage by adult drivers was found to be the highest in the households in the 90th percentile of spatial density as it was in the census tract model. High transit usage in the 40th percentile of household density was marginally significant, which was not found in the census tract model. The impact of the highest deciles of retail employee density was higher and also clearer in the block group model.

3.3 Solo Driving Demand

Here we estimate a binomial choice model of whether or not a household engages in any amount of solo driving (i.e., whether the household generates single occupant vehicle, or SOV, travel). Of the 16,750 households in our sample, 74.2% recorded an SOV trip segment. The model using spatial opportunities computed at the tract level contains variables from all three sets, but the explanatory power was largely concentrated in the sociodemographic factors (Table 8).

A more interesting model is about the adult solo driving time. For this model an ordered logit model is used for the total time that household adults spend solo driving, where that time was divided into deciles. The model results are listed in Tables 9, 10, and 11 using census tract level spatial explanatory variables and block group level spatial explanatory variables.

In terms of Sociodemographics (Table 10), total household solo driving time is a monotonically increasing function of household income, and of household size. The presence of children also has a positive effect, which is greater for younger children. The only influence of age is that the oldest two categories of average age of heads generate less solo driving time, and there are three ethnic groups that also generate less than average solo driving time, controlling for all other variables: Asian households, mixed White and Asian households, and Hispanic households. The spatial density effects (Table 11) are particularly revealing. Controlling for sociodemographic differences, households located in the lowest quintile of residential density spend more time solo driving, while those in the highest quintile of residential density spend less time. This implies that policies aimed at densification of residential areas will likely reduce solo driving time, *ceteris paribus*. Similarly, households located in areas with lower local (within 10 km) retail accessibility spend more time solo driving than households located in the highest level of retail accessibility. But, the opposite is true for regional retail accessibility: Households located in areas with above median regional retail accessibility travel more by solo driving, while households located in areas with below-median regional retail accessibility travel less by solo driving. We can surmise that the availability of local retail services reduces solo driving time, while the availability of services further from home increases such time, and conversely. For adult solo driving time, the block group model worked slightly better and the pattern of the impact of the spatial variables did not show substantial difference in the two models (Tables 11). In addition to the sample of regression models illustrated in this paper we also examined the probability of no travel by adults in households and high occupancy vehicle travel (occurrence and travel time) leading to similar conclusions that for some travel behavior variables land use density plays a major role and for others it does not. Similarly, supply of roads sometimes influences travel behavior and other times it is unrelated to behavior.

4. Summary and Conclusions

The wealth of the spatial indicators developed using information from census tracts, census block groups, and an extensive roadway network in California was used as a major group of explanatory variables in regression models of travel behavior. Each set of these accessibility capturing variables affects different travel behaviors in different ways. Household density, retail employee density and road infrastructure provided meaningful explanation of the variety in travel behaviors with description of different dimensions of accessibility such as characteristics of residential area, availability of activity opportunity, and connectivity through road infrastructure. This analysis also gave us the opportunity to analyze the impact of modifiable areal unit problem (MAUP) on the travel behavior models. From the model estimation experiments a variety of findings emerge. First, from the comparisons between the census tract models and the block group models, we see that the variables describing a behavioral aspect can show different levels and patterns of impact on travel behaviors when they are measured using different areal unit sizes. To be more specific, household density measured in census tracts explained better the indicators used here than household density measured using block groups. From the comparisons, we can say that census tracts covering a larger area around a residence and therefore capture the density impact in more informative ways. However, this cannot be the golden rule for every travel behavior indicator. We need to think about the implications that a specific areal unit has about a type of travel behavior indicator, test its ability to explain behavior, and decide to use the one that is the most informative.

Second, spatial variables involving shortest paths in computation showed better ability of discerning the impacts of each spatial segment and also clearer impact patterns of each variable set when they are computed using smaller unit areas than when they are computed using larger unit areas. Smaller unit areas provide closer approximation of the variables and those variables seem to be less susceptible to measurement error than variables computed using larger geographical units. However, the trade-off between obtaining closely approximated explanatory variables and the computing demand of using smaller areal units has to be considered when we decide which areal unit we want to use. In fact, the improvement in the goodness of fit of some regression models was marginal or even totally absent. Moreover, the two aggregation levels used here have their own inherent advantages and disadvantages. Consequently, we also demonstrate building models using spatial variables with both levels with some clear benefits.

Overall, however, land use density and supply of roadways are strong and significant explanatory sets of variables and they provide a good candidate for linking land use to travel behavior in policy impact assessments.

The entire analysis was done using data from the year 2000. The data are from products such as the Census Transportation Planning Package and a roadway network vintage 2000. The household behavior data span a set of months in 2000 and 2001. As a result all the analytical findings are for that period and may not be extendable to other times. This analysis should be expanded to include other years as opportunities for new data are multiplying due to the American Community Survey, which in 2010 will release its 5-year estimates for areas with a population of less than 20,000, including census tracts and block groups. This may provide an unprecedented opportunity to study the evolution of accessibility in our state and identify the places and their sociodemographic groups that benefitted the most by pinpointing geographic areas that increased or decreased residents' accessibility.

In the third major area of next steps we can expand the microanalysis to a more comprehensive treatment of travel behavior. This includes activity participation and interactions among household members, trip consolidation in the form of tours, and also the more traditional analysis of trip making. In addition to offering a more detailed picture of the impact that infrastructure and density of opportunities cause on travel behavior, this next step has also the potential to improve the statewide transportation model maintained by Caltrans. This last area of analysis is also a fruitful research direction in developing a next generation of land use transportation integrated models.

Acknowledgments

Funding for this project was provided by California PATH Task Orders 5110 & 6110 and support for the University of California Transportation Center. The authors are also thankful to James Marca, Val Noronha and Bryan Krause who worked in earlier stages of the project. The paper does not constitute official policy of any government agency or organization.

References

Abreu J. Da Silva, T. Golob, and K.G. Goulias (2006) The effects of land use characteristics on residence location and travel behavior of urban adult workers. Paper accepted for presentation at the Transportation Research Board 2006 Annual Meeting and publication in the *Transportation Research Record*.

Boarnet M. And S. Sarmiento (1996) Can Land Use Policy Really Affect Travel Behavior? A Study of the Link between Non-Work Travel and Land Use Characteristics (http://repositories.cdlib.org/itsirvine/casa/UCI-ITS-AS-WP-96-5/).

Goulias K. G., L. Blain, N. Kilgren, T. Michalowski, and E. Murakami (2007) Catching the Next Big Wave: Are the Observed Behavioral Dynamics of the Baby Boomers Forcing us to Rethink Regional Travel Demand Models? Paper presented at the 86th Transportation Research Board Annual Meeting, January 21-25, 2007, Washington, D.C. and included in the CD ROM proceedings.

Khattak A. And D. Rodriguez (2005) Travel behavior in neo-traditional neighborhood developments: A case study in USA. Transportation Research Part A. Volume 39, Issue 6, July 2005, Pages 481-500

Kitamura R., P.L. Mokhtarian, and L. Laidet (1997) A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. Transportation, vol. 24, nº2, pp. 125-158.

Miller E. (2002) Land Use: Transportation Modeling. In Transportation Systems Planning: Methods and Applications (K. G. Goulias – Editor) CRC Press.

Openshaw S, Alvandies S (1999) Applying geocomputation to the analysis of spatial distributions. In Longley P, Goodchild M, Maguire D, Rhind, D (eds) *Geograhpic Information Systems: Principles and Technical Issues*. Vol 1, 2nd ed. New York: John Wiley and Sons Inc.

Table 1 US year 2000 Census tract and block group characteristics in California

Census tract characteristics	Mean	Std.Dev.	Maximum*
Tract Square Km	59.0	453.7	20486.8
Tract Population	4805.2	2143.1	36146.0
Tract Households	1631.8	763.0	8528.0
Within a 5 Km Buffer from Tract Centroid			
Workers in Retail (retail)	5031.1	6937.8	54745.0
Workers in Health (health)	2644.0	3524.4	26478.0
Workers in Services but not in Health or Retail (services)	28024.4	44497.0	373127.0
Workers in Manufacturing (manufacturing)	3391.0	5547.7	59059.0
Workers in All Other Occupations (other)	5753.4	6805.7	50287.0
Primary limited access roadways (primary lim)	284.1	448.6	3244.8
Primary without limited access roadways (primary nolim)	77.9	140.6	958.5
Secondary and connecting roadways (secondary)	1867.8	2711.3	17711.4
Rural, local and neighborhood roadways (local)	8549.4	11256.1	71318.1
Special roadways (special)	342.1	591.3	4612.7
All Other types of roadways (other)	778.6	1618.7	10511.1
Block group characteristics	**Mean**	**Std.Dev.**	**Maximum***
Block Group Square Km	18.51	179.59	12219.12
Block group Population	1530.3	1008.48	36146
Block group Households			
Within a 20 min travel time buffer from block group Centroid			
Workers in Retail (retail)	56324.49	48926.91	202513
Workers in Health (health)	96664.34	89718.16	389816
Workers in Services but not in Health or Retail (services)	23812.89	23757.93	87798
Workers in Manufacturing (manufacturing)	80640.04	88937.65	339848
Workers in All Other Occupations (other)	75843.44	68947.56	270979
Primary limited access roadways (primary lim)	266.53	206.05	885.86
Primary without limited access roadways (primary nolim)	78.4	82.01	552.42
Secondary and connecting roadways (secondary)	650.52	425.51	2333.31
Rural, local and neighborhood roadways (local)	2561.13	1782.39	12545.59
Special roadways (special)	23.2	39.44	483.4
All Other types of roadways (other)	223.78	275.34	1984.31

*The minimum is zero for all variables and tracts

Table 2 Sociodemographic variables used in the regression models

Variable	%	Variable	%	Variable	%
Annual Household income		Average age of heads		Highest education of head	
<$10,000	4.3	18-25	5.8	not high school	9.1
$10,000-$24,999	14.2	25.5-35	14.1	high school graduate	24.5
$25,000-$34,999	13.2	35.5-45	20.1	Some college	23.7
$35,000-$49,999	13.9	45.5-55	22.7	associates degree	7.4
$50,000-$74,999	19.9	55.5-65	15.5	bachelors degree	20.9
$75,000-$99,999	10.9	65.5-75	11.8	graduate degree	13.4
$100,000-$149,999	7.4	75.5+	7.5	Unknown	1.1
$150,000+	3.4	Unknown	2.5	Whether any children < 6	
Unknown	12.8	Ethnicity of heads		Yes	7.5
Household size		White	75.5	No	89.4
1	26.4	Hispanic	10.2	Whether any children 6-12	
2	40.8	Black	2.3	Yes	9.3
3	14.4	Asian/Pacific Islander	1.9	No	85.6
4	11.2	White & Hispanic	3.1	Whether any children 13-17	
5	4.7	White & Asian	1.3	Yes	9.0
6 or more	2.5	Other or unknown	5.8	No	2.9

Table 3 Logit Models of Any Household Nonmotorized Travel

Model	Variable set	Contribution of set		Cumulative model		
		Chi-square	Degrees of freedom	Chi-square	Degrees of freedom	Nagelkerke R^2
Census Tract	**Sociodemographic**	**1065.65**	**35**	**1065.65**	**35**	**.116**
	Spatial density	**373.08**	**27**	**1438.73**	**62**	**.147**
	Household density	*277.78*	*9*			
	Retail employee	*95.30*	*18*			
	Road infrastructure	**104.31**	**18**	**1543.04**	**80**	**.158**
Block Group	**Sociodemographic**	**1065.94**	**35**	**1065.94**	**35**	**.110**
	Spatial density	**325.60**	**27**	**1391.54**	**62**	**.143**
	Household density	*191.22*	*9*			
	Retail employee	*134.38*	*18*			
	Road infrastructure	**141.40**	**18**	**1532.94**	**80**	**.157**
Block G. variables with CT HH density	Sociodemographic	1065.54	35	1065.54	35	.110
	Spatial density	382.99	27	1488.53	62	.148
	Road infrastructure	123.29	18	1571.82	80	.160

Table 4 Logit Models of Household Nonmotorized Travel – Spatial Density

Ind. Variable (all bases = 50th %tile)	Block group		BG variables with CT HH density	
	Significance	Odds ratio	Significance	Odds ratio
household density	0.00		0.00	
<10 %tile			0.00	0.683
10th %tile	0.00	0.713	0.00	0.541
20th %tile	0.00	0.634	0.00	0.621
30th %tile			0.05	0.858
40th %tile				
60th %tile			0.05	1.160
70th %tile	0.02	1.186	0.01	1.227
80th %tile	0.03	1.168	0.00	1.432
90th %tile	0.00	1.475	0.00	2.019
retail employees within 10 km	0.00		0.00	
<10 %tile	0.00	0.668		
10th %tile				
20th %tile			0.01	1.235
30th %tile				
40th %tile				
60th %tile				
70th %tile			0.02	0.834
80th %tile			0.00	0.757
90th %tile	0.00	1.798	0.00	1.426
retail employees within 10 to 50km	0.00		0.00	
<10 %tile	0.00	0.610	0.00	0.660
10th %tile	0.00	0.675	0.00	0.736
20th %tile	0.01	0.788		
30th %tile	0.00	0.745	0.00	0.777
40th %tile	0.00	0.554	0.00	0.570
60th %tile				
70th %tile	0.02	1.254	(0.06)	(1.199)
80th %tile	0.00	2.082	0.00	1.916
90th %tile	0.00	3.401	0.00	2.910

Table 5 Logit Models of Household Nonmotorized Travel – Infrastructure

Ind. Variable (all bases = 50th %tile)	Block group		BG variables with CT HH density	
	Significance	Odds ratio	Significance	Odds ratio
primary roads w/o limited access within 10 to 50 km	0.00		0.00	
<10 %tile				
10th %tile	0.00	0.735	0.00	0.744
20th %tile	0.02	0.839	0.02	0.827
30th %tile	0.04	1.161	(0.06)	(1.149)
40th %tile				
60th %tile				
70th %tile	0.02	1.180	0.03	1.165
80th %tile				
90th %tile				
Local roads within 10 to 50 km	0.00		0.00	
<10 %tile	0.00	2.325	0.00	2.128
10th %tile	0.00	1.739	0.00	1.647
20th %tile	0.01	1.385	0.04	1.316
30th %tile	0.00	1.430	0.00	1.391
40th %tile				
60th %tile	0.04	0.842		
70th %tile	0.00	0.662	0.00	0.678
80th %tile	0.00	0.630	0.00	0.678
90th %tile	0.00	0.295	0.00	0.330

Table 6 Logit Models of Household Nonmotorized Travel by Adult Drivers

Model	Variable set	Contribution of set		Cumulative model		
		Chi-square	Degrees of freedom	Chi-square	Degrees of freedom	Nagelkerke R^2
Census Tract	**Sociodemographic**	**251.86**	**35**	**251.86**	**35**	**.036**
	Spatial density	**306.80**	**27**	**558.66**	**62**	**.079**
	Household density	*204.38*	*9*			
	Retail employee	*102.42*	*18*			
	Road infrastructure	**158.58**	**27**	**717.24**	**71**	**.101**
Block Group	**Sociodemographic**	**251.86**	**35**	**251.86**	**35**	**.036**
	Spatial density	**272.49**	**27**	**524.34**	**62**	**075**
	Household density	*131.98*	*9*			
	Retail employee	*140.51*	*18*			
	Road infrastructure	**182.18**	**27**	**706.52**	**89**	**.100**
BG variables with CT HH density	Sociodemographic	251.97	35	251.97	35	.036
	Spatial density	312.70	27	564.67	62	.080
	Road infrastructure	164.57	27	729.24	89	.103

Table 7 Transit Usage Models Comparison

Logit Models of Any Household Transit Use

Model	Variable set	Contribution of set		Cumulative model		
		Chi-square	Degrees of freedom	Chi-square	Degrees of freedom	Nagelkerke R^2
Census Tract	**Sociodemographic**	**1633.28**	**35**	**1633.28**	**35**	**.216**
	Spatial density	**177.90**	**27**	**1811.19**	**62**	**.238**
	Household density	*125.45*	*9*			
	Retail employee	*52.45*	*18*			
	Road infrastructure	**81.57**	**9**	**1892.76**	**71**	**.248**
Block Group	**Sociodemographic**	**1633.58**	**35**	**1633.58**	**35**	**.216**
	Spatial density	**180.37**	**27**	**1813.95**	**62**	**.238**
	Household density	*106.50*	*9*			
	Retail employee	*73.87*	*18*			
	Road infrastructure	**158.66**	**9**	**1972.60**	**71**	**.258**

Logit Models of Household Transit Use by Adult Drivers

Model	Variable set	Contribution of set		Cumulative model		
		Chi-square	Degrees of freedom	Chi-square	Degrees of freedom	Nagelkerke R^2
Census Tract	**Sociodemographic**	**216.32**	**35**	**216.32**	**35**	**.068**
	Spatial density	**282.90**	**27**	**499.22**	**62**	**.155**
	Household density	*205.39*	*9*			
	Retail employee	*77.51*	*18*			
	Road infrastructure	**64.76**	**9**	**563.98**	**71**	**.175**
Block Group	**Sociodemographic**	**216.34**	**35**	**216.34**	**35**	**.068**
	Spatial density	**297.52**	**27**	**513.86**	**62**	**.159**
	Household density	*180.12*	*9*			
	Retail employee	*117.40*	*18*			
	Road infrastructure	**116.93**	**18**	**630.76**	**80**	**.195**

Table 8 Logit Models of Household Solo Driving (SOV use)

Model	Variable set	Contribution of set		Cumulative model		
		Chi-square	Degrees of freedom	Chi-square	Degrees of freedom	Nagelkerke R^2
Census Tract	**Sociodemographic**	**2083.85**	**35**	**2083.85**	**35**	**.172**
	Spatial density	**53.91**	**18**	**2137.76**	**53**	**.176**
	Household density	-	-			
	Retail employee	*53.91*	*18*			
	Road infrastructure	**46.74**	**9**	**2184.50**	**62**	**.180**
Block Group	**Sociodemographic**	**2084.57**	**35**	**2084.57**	**35**	**.172**
	Spatial density	**39.05**	**9**	**2123.62**	**44**	**.175**
	Household density	-	-			
	Retail employee	*39.05*	*9*			
	Road infrastructure	**43.10**	**16**	**2166.73**	**60**	**.178**

Table 9 Ordered Logit Models of Total Household Solo Driving Time by Adults

Model	Variable set	Contribution of set		Cumulative model		
		Chi-square	Degrees of freedom	Chi-square	Degrees of freedom	Nagelkerke R^2
Census Tract	**Sociodemographic**	**1811.24**	**35**	**1811.24**	**35**	**.121**
	Spatial density	**237.64**	**27**	**2048.88**	**62**	**.136**
	Household density	*11.89*	*9*			
	Retail employee	*225.75*	*18*			
	Road infrastructure	(Not significant)				
Block Group	**Sociodemographic**	**1811.80**	**35**	**1811.80**	**35**	**.121**
	Spatial density	**267.61**	**27**	**2079.42**	**62**	**.138**
	Household density	*26.83*	*9*			
	Retail employee	*240.78*	*18*			
	Road infrastructure	(Not significant)				

Table 10 Ordered Logit Model of Household Solo Driving Time – Sociodemographic

Independent variable	Significance	Coefficient
Income (base = unknown)		
<$10,000	0.00	-0.422
$10,000-$24,999	0.00	-0.386
$25,000-$34,999	0.04	-0.124
$35,000-$49,999		
$50,000-$74,999	0.00	0.200
$75,000-$99,999	0.00	0.431
$100,000-$149,999	0.00	0.484
$150,000+	0.00	0.456
household size (base = 6 or more)		
1	0.00	-1.458
2	0.00	-1.121
3	0.00	-0.570
4	0.00	-0.310
5		
Average age of heads (base = unknown)		
18-25		
25.5-35		
35.5-45		
45.5-55		
55.5-65		
65.5-75	0.00	-0.442
75.5+	0.00	-0.653
Ethnicity (base = unknown)		
White		
Hispanic	0.03	-0.178
Black		
Asian/Pacific Islander	0.03	-0.272
White & Hispanic		
White & Asian		
Education (base = unknown)		
not high school graduate		
high school graduate		
some college		
associates degree		
bachelors degree		
graduate degree		
presence of children 0-5 yrs. Old	0.00	0.269
presence of children 6-12 yrs. Old	0.00	0.192
presence of children 13-17 yrs. Old	0.00	0.140

Table 11 Ordered Logit Models of Household Solo Driving Time – Spatial Density

Ind. Variable (all bases = 50th %tile)	Census tract		Block group	
	Significance	Coefficient	Significance	Coefficient
household density				
<10 %tile	0.05	0.185	(0.11)	(0.140)
10th %tile	0.05	0.154	0.05	0.148
20th %tile			0.00	0.196
30th %tile				
40th %tile				
60th %tile				
70th %tile				
80th %tile	(0.06)	(-0.125)		
90th %tile	0.01	-0.208	(0.09)	(-0.123)
retail employees within 10 km				
<10 %tile	0.05	0.191	0.00	0.351
10th %tile	0.00	0.234	0.00	0.262
20th %tile	0.01	0.183	0.19	0.096
30th %tile	0.02	0.158	0.01	0.186
40th %tile				
60th %tile				
70th %tile				
80th %tile				
90th %tile	0.01	-0.212	0.00	-0.244
retail employees within 10 to 50km				
<10 %tile	0.00	-0.487	0.00	-0.495
10th %tile	0.00	-0.273	0.00	-0.296
20th %tile	0.03	-0.145	0.01	-0.166
30th %tile	0.01	-0.168	0.01	-0.184
40th %tile				
60th %tile	0.01	0.179	0.01	0.166
70th %tile	0.00	0.391	0.00	0.362
80th %tile	0.00	0.432	0.00	0.403
90th %tile	0.00	0.512	0.00	0.503