Land Use and Vehicle Miles of Travel in the Climate Change Debate: Getting Smarter than Your Average Bear

Marlon G. Boarnet, Douglas Houston, Gavin Ferguson, and Steven Spears

Department of Planning, Policy, and Design,
Department of Economics (Boarnet),
Institute of Transportation Studies,
and Metropolitan Futures Initiative of the School of Social Ecology

University of California, Irvine

Revised Version: ~~August 13~~October 24, 2010

Presented at the 5th annual Land Policy Conference of the
Lincoln Institute of Land Policy
Cambridge, Massachusetts
May 24-25, 2010

Acknowledgements: We thank the Lincoln Institute of Land Policy, the University of California Transportation Center, and the University of California Multi-Campus Research Program on Sustainable Transportation for financial support. The Southern California Association of Governments (SCAG) generously provided data and advice about the data sets. Hsin-Ping Hsu and Dongwoo Yang provided valuable research assistance. We thank Kenneth Small for helpful comments on an earlier draft, and we thank the volume editors, Greg Ingram and Yu-Hung Hong, for their suggestions throughout the process. The authors alone are responsible for any shortcomings or interpretations in this paper, and the opinions expressed here do not necessarily represent those of SCAG or any of the persons or entities mentioned in these acknowledgements.

Abstract

Almost everything we know about land use and travel behavior is average effects for metropolitan areas or larger. While averages are an important part of the policy puzzle, planning requirements like California's Senate Bill 375 can also benefit from an assessment of where land use policies might be associated with travel behavior effects that are smaller or larger than the metropolitan average. We use regression analysis of a detailed travel diary survey in the greater Los Angeles metropolitan area to show how the effect of regional employment accessibility on vehicle miles of travel (VMT) varies in different places. For households in the third and fourth quintiles of employment accessibility, the estimated elasticity of VMT with respect to employment accessibility is as much as three to four times larger than common estimates in the literature. This suggests a more important role for land use in transportation and climate change policy, and also suggests a focus on employment accessibility as a key variable. The results also imply that the interaction of land use and transportation at the metropolitan scale, as opposed to focusing only on neighborhoods, is fundamental if the goal is to reduce VMT and associated greenhouse gas emissions.

A common planning response to climate change has been to focus on the relationship between land use and travel behavior. The transportation sector was responsible for 27 percent of the greenhouse gas (GHG) emissions in the U.S. in 2008 (US EPA, 2010).[1] If cities were built at higher densities, with mixed land uses and alternatives to car travel, would that help reduce GHG emissions? The answer is hotly debated (see, e.g., Moore, Staley, and Poole, 2010, Winkelman and Bishins, 2010, and Boarnet, 2010), yet amidst that debate, a key issue has been overlooked.

Most everything that we know about land use and travel behavior is regional averages, typically from studies that analyze travel diary data for a metropolitan area or larger geographies. We know almost nothing about departures from metropolitan area averages. Are there places where the impact of land use on vehicle miles of travel (VMT) might be larger than the metropolitan average, or smaller than the average? Is the relationship between particular land use variables and VMT characterized by non-linearities and thresholds? Both logic and the limited evidence that we have suggest "yes", but the question of non-linearities, or thresholds, in the land use – VMT relationship has rarely been examined. Policy application necessarily requires an understanding not only of an average effect, but also where limited resources and attention might be applied. This paper is an attempt to move beyond broad averages, to search for thresholds in land use – VMT links, and to do so in a way that begins to illuminate the role for land use planning in the climate change debate.

We use an exceptionally detailed travel diary data for the five-county Southern California Association of Governments region, covering the greater Los Angeles metropolitan area. Our primary methodological tool is a standard land use – travel behavior regression, regressing household VMT on a set of household sociodemographic variables and land use measures at the household's place of residence, but we estimate the regression for threshold values of key land use variables, to test for non-linearities and interactions among key land use variables.[2] Overall, we find that access to regional employment has a non-linear effect on VMT. Converting our regression results to elasticities, we find elasticities of VMT with respect to employment accessibility that are, in some cases, three to four times as large as corresponding elasticities in the literature, suggesting that the influence of land use can vary in ways largely overlooked by previous research.

---

[1] A more common estimate is that the transportation sector is responsible for about 1/3 of all U.S. greenhouse gas emissions. That estimate is from years before 2008 and compares the transport sector to net emissions, after carbon sinks, rather than to gross emissions without sinks, which for 2008 puts the transport sector at 27 percent of the total (U.S. EPA 2010).

[2] The data only provide information on VMT, not GHG emissions. We note that conditional on vehicle fleet composition, reductions in VMT will be associated with reductions in GHG emissions. There is a broader debate about whether GHG reduction should come from increases in vehicle fuel efficiency or from VMT reduction. We note that policy will likely need both levers – high fleet fuel efficiency and measures that reduce the growth of VMT – partly because experts estimate that available increases in fuel efficiency will not be sufficient to meet GHG reduction goals and more importantly because planning efforts that slow the growth of VMT have many co-benefits, including reductions in criteria pollutants and improvements in community quality of life.

I. Background: The Literature on Land Use and Travel Behavior

There are scores of studies of land use and travel behavior. For reviews of the literature, see Badoe and Miller (2000), Boarnet and Crane (2001, chapter 3), Brownstone (2008), Crane (2000), Ewing and Cervero (2001), and Handy (2005). A recent National Research Council (NRC, 2009) report concluded that the elasticity of VMT with respect to population density is in the range from -0.05 to -0.12, and that accounting for the effect of changing multiple land use variables together, the impact of compact development on VMT might imply an elasticity on the order of -0.25. The NRC (2009) report, while stating that the data sources and hence the scientific evidence were at times thinner than would be preferred, concluded that both logic and available evidence suggest that the relationship is at least in part causal, implying that increases in density would reduce VMT. Translating the density – VMT relationship into greenhouse gas emissions reductions, the NRC (2009) report estimated that more compact development could reduce GHG emissions, below a baseline trend, by an amount in the range of less than 1 percent to approximately 11 percent by 2050.

Viewed more broadly, the question of land use and travel behavior is at once obvious and an exceptionally slippery topic of empirical study. In 1954, Mitchell and Rapkin wrote one of the earliest tomes on travel demand modeling, titling that work *Urban Traffic: A Function of Land Use*, reflecting both the obvious nature of that link and a viewpoint that underpins travel demand modeling to this day. Yet viewing travel as a behavioral response complicates matters substantially. Persons can change when and how they travel, and by choosing their place of residence, job location, and activity locations, persons and economic actors choose their trip origins and destinations. That behavioral complication, spanning several markets all characterized by imperfect data, combined with few opportunities to observe natural experiments, have all conspired to complicate the empirical literature on land use and travel behavior.

In the past two decades, a standard approach has developed. Measures of individual or household travel, typically from a travel diary, are regressed on sociodemographic variables (e.g. income, age, number of children, employment status, and the like) and land use variables typically measured near the person's or household's residence. The use of data for individuals or households avoids aggregation problems and potentially allows better causal influence. Recent high quality applications of this approach include Bhat and Guo (2007), Bento et al. (2005), Brownstone and Golob (2009), and Fang (2008). Yet amidst that standard approach, several methodological issues have been debated. The most important methodological issues are discussed below.

Measuring Land Use

The NRC (2009) report's elasticity estimate focused on density, because that is the most common variable used in the literature. Yet density is certainly a proxy – and likely often a weak proxy – for a broad range of land use variables. Generally, land use is measured according to what are now called the D variables – density, diversity (i.e. land use mix), design (i.e. the character of the street network and for some studies the quality of the pedestrian environment), access to destinations (often measured as access to regional job centers or employment), and distance from transit or, more completely, access to and characteristics of the regional transportation network. Some of these variables are very local, often measured for neighborhoods that approximate walking distance, typically ¼ or ½ mile areas. Population density, land use mix, and the character of the street network are typically measured for localized neighborhoods, and are intended to measure some of the neighborhood-scale ideas associated with Smart Growth. Other variables – access to regional employment and characteristics of and distance from the transportation network – describe geographies at the scale of the metropolitan area. For a discussion of local and regional access, and these different geographies, see Handy (1993).

Residential Selection

Economic actors might choose their locations in part based on how they wish to travel or, for firms, how persons can travel to them. Households who wish to walk might choose to live in a walking oriented neighborhood. Firms will locate based on transportation accessibility. The literature on land use and travel, being almost exclusively focused on household travel due to data constraints, has focused on the former question, called "residential selection."

The idea that households might locate based on how they wish to travel was first suggested by Boarnet and Sarmiento (1998), who modeled this as an endogeneity problem, using instrumental variables as a correction. Since then, several studies have modeled residential selection, using a variety of techniques. Cao et al. (2009) reviewed the literature on residential selection, finding that virtually every study that attempted to control for the possible endogeneity of household location found that econometric controls did not change the sign or significance of the land use – travel behavior association. In some cases the magnitude of the land use – travel effect was attenuated after controlling econometrically for residential selection. The association between land use and travel appears to be partly causal, and partly persons sorting (or choosing) residential locations that match their travel preference.

The Cao et al. (2009) review treats residential selection as an econometric problem, as is typical of this literature. But whether selection occurs or not and how to correct for residential selection in a regression ignores the question of whether selection should or should not be considered part of the policy impact.

The selection question is motivated in part by econometric studies of labor markets. A good analogy for the labor economics perspective is job training programs. Suppose a voluntary job training and job search program is offered to persons at a mythical factory which recently closed. Presumably the voluntary program will attract the most motivated, and possibly the more highly skilled, former employees. Graduates of the training program might find jobs more quickly, but one would want to adjust for the selection of persons into training. If, in the extreme case, the training program imparts no skills, through selection participants could still do better in their job search than non-participants. In that extreme case the entire "program effect" is selection, rather than the effect of the training program itself, and presumably the training participants would have found the same jobs had there never been a training program.

Suppose, instead, that there is a distribution of persons, the most motivated of which will benefit from job training (possibly because their motivation will cause them to better leverage the training into better jobs), but the least motivated of which will not benefit from job training. All participants select into the training program (maybe the program was so effectively marketed that even the non-motivated chose to participate). If there are other, pre-existing and perfectly suitable training or education options, the new training program would have no impact on labor market outcomes – the motivated would have obtained training elsewhere and all the new program did was enroll less motivated persons who will not benefit from the training. If, on the other hand, no other training programs exist, our mythical factory's job training program improves labor market outcomes. In a case where a good (the job training program in this example) produces an impact for at least some who select in and the good is in short supply such that all who could benefit cannot consume (i.e. there are no other training programs) the supply of the good itself can be considered part of the policy impact.

Which analogy better fits compact development and Smart Growth? Is compact development an adequately supplied service, providing no intrinsic change in travel behavior, such that selection is not part of the policy effect? Or is compact development a desired (and desirable) but under-supplied good, such that selection could legitimately be the whole of the policy effect? More importantly, if there is a distribution of travel preferences, how does the supply of neighborhoods match the demand for travel and hence neighborhood type? Levine (2006) argues that compact development is under-supplied, and that selection should be viewed as a legitimate part of any land use – travel impact.

A full analysis would be complex and, to our knowledge, has never been attempted. One would have to formally model developer behavior and possibly government regulations that govern the supply of neighborhoods, and also travel behavior, and the interactions between those via land use change.[3] This would go somewhat beyond trying to econometrically control for residential selection to uncover an effect absent selection, as those methods do not address the question of whether selection is or is not part of the policy impact. So while residential selection has become possibly the central

[3] An alternative possibility would be to examine land use – travel behavior relationships in a metropolitan area with an abundant supply of compact development, hypothesizing that in such locations the marginal effect of land use on travel is not due to persons who desire compact neighborhoods sorting into such neighborhoods but instead is due to a direct effect of land use on travel. To the best of our knowledge, such a method for illuminating the residential selection question has not been tried. Note, though, that "abundant supply" would have to be measured relative to consumer preferences, complicating the research design as simply looking for locations with a large amount of compact development would not necessarily imply "abundance" if those places also have high demand for compact development. Relatedly, it is possible that some persons might migrate across metropolitan areas based in part on characteristics that include the supply of compact development. While there is no good formal evidence for migration based on a metropolitan areas compactness, commentators have speculated on such possibilities (e.g. *The Economist*, 2010). All of these issues illuminate the need for a more structural approach to questions of land use – travel – residential selection, but such efforts are beyond the scope of the study presented here.

methodological issue in this literature, the approach has been narrow in ways not broadly appreciated. Yet our concern here is not broadening the argument beyond econometric corrections for residential selection, but putting on the table an equally important methodological concern that has received almost no attention.

Regional Average Impacts and Departures from Regional Averages

Regression analyses of land use and travel give coefficients that represent an average effect for the data set. There are many such studies – enough that the literature has settled on a common range for various land use variables. Studies of individual travel diary data that correct for residential selection suggest that the elasticity of VMT with respect to population density is in a range from -0.05 to -0.12 (Bento et al., 2005, Brownstone and Golob, 2009, Fang, 2008, Del Valle and Niemeier, 2010, NRC, 2009).[4] Studies give an elasticity of VMT with respect to land use mix in a range from -0.01 to -0.06 (Bento et al., 2005; Chapman and Frank, 2004; Frank et al., 2005; Kockelman, 1997; Pushkar et al., 2000). Regional accessibility to jobs has a larger VMT elasticity – typically in the range of -0.15 to -0.31 (Ewing and Cervero, 2010, Table A-4). Meta-analyses give results that are typical of these individual studies; two meta-analysis studies by Ewing and Cervero (2001 and 2010) found elasticities of VMT with respect to population density and land use mix in the range of -0.04 to -0.09 and an elasticity of VMT with respect to regional access to jobs in the range of -0.20 to -0.22, leading to the general conclusion that the regional (or metropolitan-wide) distribution of employment has a larger magnitude effect on VMT than do neighborhood-scale land use variables.[5]

While an average effect for a metropolitan area is one piece of the policy puzzle, knowing whether and how the effect sizes vary across different land use contexts is just as important. The literature gives little information about departures from regional (or metropolitan area) average effects, but some small-area studies suggest that such departures can be important. For example, Boarnet et al. (2011) found substantial travel behavior variation within their relatively small study neighborhoods. For one neighborhood, controlling for household characteristics, residents within ¼ mile of a commercial concentration averaged five times more walking trips and 25 percent fewer driving trips than persons in the balance of the same neighborhood, all of whom were within a mile of the same commercial concentration.[6] Moving from that specific

[4] Del Valle and Niemeier (2010) find an elasticity of VMT with respect to residential density equal to -0.19. That paper was not available at the time of the NRC (2009) study and so the range typically identified by NRC (2009) and similar efforts is -0.05 to -0.12.

[5] Note that this conclusion overlooks some complexity. Ewing and Cervero (2010), in their meta-analysis, find that the VMT elasticity of the street network – measured by the density of street intersections and the percent of intersections that are four-way – is -0.12 and the elasticity of walking with respect to intersection density is 0.39. But looking across a range of variables and impacts, when focusing on VMT, the regional employment elasticity is typically larger than elasticticies for neighborhood level land use variables.

[6] Using data in Boarnet et al. (2011), the results from the neighborhood in question, Artesia Boulevard, imply an approximate 25 percent decrease in car trip generation rates across two locations where housing unit density differs by about 50 percent. While density was not included

example to the general, where might land use policy achieve more, or less, bang for the "compact development" buck, so to speak?

California's SB 375 moves this question to the fore of metropolitan planning. SB 375 requires that the state's 18 MPOs develop sustainable community strategies (SCS's) which will demonstrate how the MPO's regional transportation plan and affordable housing strategy combine to meet GHG emission reduction targets established by the California Air Resources Board. In the Southern California Association of Governments (SCAG) region – the largest MPO in the state and the nation – this requires an SCS that is, in part, an amalgamation of the transportation investment and general plan decisions of 189 cities and six counties. Development patterns in the SCAG region range from the high-rise canyons of the highly urban Wilshire corridor to exurban commuter suburbs separated from job centers by commutes of an hour or more. Should attention be focused on inner locations with urban character, or would policies targeted toward the exurban fringe have a larger impact? While it is obvious that land use – transportation plans should be sensitive to local context, beyond good planning intuition there is little in the way of analytics that can help refine and apply that insight. SCAG has conducted a visioning process that identifies areas for targeted infill development (called "Compass opportunity areas"), but that identification was largely based on access to bus or rail transit, rather than the larger menu of land use variables used in this study.

II. Data

VMT data are derived from the SCAG 2001 regional travel survey. The survey was conducted in Spring and Fall 2001 and Spring 2002. The survey included a travel diary, completed by 16,939 households in the SCAG six county region (Los Angeles, Orange, Riverside, San Bernardino, Ventura, and Imperial Counties.) The diary was one day for most households, with 2,422 households completing a two-day travel diary that included at least one weekend day. Start days were staggered throughout the week. For a more complete description of the survey and results, see NuStats (2003).

The travel survey asked respondents to record all trip origin and destination locations for all household members, plus travel mode, trip purpose, and time of day. Trip origins and destinations were geocoded by SCAG and their survey contractor, NuStats. We used the geocoded trip origins and destinations to obtain trip distances and total household VMT for the diary period.

All trips were routed over a street and highway network using Mapquest in February of 2010 with options that choose minimum travel times. The routing method is internal to Mapquest, but the research team experimented with other routing tools, namely CloudMade and Google maps, and found little difference in overall VMT calculations.

---

directly in the regression, an approximate 50 percent increase in density associated with an approximate 25 percent reduction in car trip generation rates implies a larger elasticity than is typical in the literature. As a comparison, trip generation elasticities with respect to population density in the more regional-level studies reviewed by Ewing and Cervero (2001) are never larger than 0.15 in magnitude.

The street/highway network is complete, but is for early 2010 rather than the survey year of 2001-2002. In concept, that could create errors due to road and street construction in the intervening years, but in practice we doubt that is much of an issue. The overwhelming majority of locations within the SCAG region had no new center-line miles of street or highway construction during the past nine years; in most locations, the current street/highway network is what existed in 2001. We examined several routed trips visually and found that the routing method was reasonable and gave the expected routes.

As a further check, we compared our household VMT to an estimate of household VMT calculated by SCAG using a 2003 road network.[7] SCAG's VMT calculation was only available for approximately 70 percent of the households in the data set, leaving 10,630 households with VMT from both SCAG's 2003 calculation and our method of routing over a 2010 network. For those 10,630 households, the mean and median difference in VMT were -1.45 and 0.67 miles, respectively, compared to a mean household VMT of approximately 48 miles. Differences are distributed roughly evenly across both negative and positive values, suggesting no obvious bias from either source. The middle 80 percent of the differences are approximately ± 7 miles. After visual inspection of households with large difference of both signs, we concluded that any differences appear unrelated to the year of the road network and are more likely related to differing decisions about whether to include out-of-region travel. We prefer the inclusion criteria for trips that we developed, described below.

Trips were discarded if the survey variable spdflag indicated respondent error (spdflag=1) or unresolved speed violation (spdflag=5). The spdflag variable was developed by SCAG to indicated trips that could not be routed over the network. Those trips often indicated errors in the respondent's identification of trip origin or destination location or cases where the implied travel speed was much faster than allowed travel speeds, as identified in SCAG's quality check of the data. Some of these erroneous trips are very long – in some cases several hundred miles – but correspond to short travel times, implying impossible travel speeds and indicating an error in respondent response. We included trips that were outside of the five-county metropolitan area, as our interest is in total VMT, subject to the condition that the spdflag variables did not indicate respondent errors. Lastly, we took care to avoid double counting the same trip more than once when multiple household members travelled in the same vehicle.

The sociodemographic variables are from the travel diary. The land use variables were obtained from GIS, using the geocoded residential location of each travel diary household. Data included the 2000 census, SCAG data on land use categories for census block groups and also number of jobs by block group for 2000, and rail and bus routes.

The land use variables are in three groups:

[7] SCAG's VMT data were only available to us at the household level, not by individual trip, allowing no ability to recalculate VMT with different assumptions or decompose travel by different sets of trips. For that reason, we believe it is preferable to use SCAG as a quality check for our routing method since, if our method gives similar results – and it does – our VMT data are disaggregated to the trip level and so allow more detailed future analysis.

- variables that measure neighborhood level characteristics: population density within ¼ mile radius around the household's residence, fraction of land that is in commercial use within ¼ mile from the household, fraction of land that is in medium or high density residential use within ¼ mile from the household, total number of street intersections within ½ mile from the household (a measure of block size), and fraction of street intersections within ½ mile from the household that are four-way (a measure of grid-orientedness of the street network),
- variables that measure access to jobs throughout the metropolitan area (regional access to jobs): distance from the central business district (Los Angeles City Hall), and a gravity variable that sums census block group employment damped by straight-line distance (in meters) from household residence to census block group centroid, and
- variables that measure access to the transportation network: distance from the nearest freeway on-ramp, dummy variables indicating if the household is within ¼ and ½ miles from a rail transit station, dummy variable indicating if the household is within ½ mile of a bus station, dummy variable indicating if the household is within ½ mile of an express bus station, dummy variable indicating if a household is within ½ mile of a rapid bus station.

In addition, a dummy variable indicating whether the household was within a Compass opportunity area was used in the regression. The Compass opportunity areas were developed as part of a planning exercise that culminated in 2004 in the SCAG Compass Blueprint plan.[8] Opportunity areas are judged by SCAG to be good targets for infill development, and were chosen based on access to employment or activity centers, access to rail or bus transit, and infill development opportunities. See http://www.compassblueprint.org/opportunityareas for more information. Collectively, these land use variables provide a comprehensive treatment of the typical 5D approach to land use measurement. Descriptive statistics for the land use variables are shown in Appendix A, for all households which have a full set of sociodemographic variables.

III. VMT in the SCAG Region: Descriptive Results

Table 1 gives descriptive statistics for VMT in the full sample for the SCAG travel diary. There are 118 households for which diary VMT could not be derived, and another 3,220 households (19% of the sample) with zero vehicle miles traveled during the diary period. That proportion of "zero vehicle miles-traveled" households is not uncommon in one-day travel diary surveys.

[8] Compass is not an acronym. Apparently the phrase "Compass" was used to indicate a direction or a charting of a course, presumably to communicate the future-oriented nature of the growth vision plan. See Southern California Association of Governments (2004).

Table 1: Household VMT for the Travel Diary Period,
SCAG 2001 Survey

	vehicle miles traveled		
	total	weekday	weekend
mean	47.81	42.89	34.51
standard deviation	77.52	67.37	69.37
10^{th} percentile	0.00	0.00	0.00
25^{th} percentile	4.75	4.09	0.00
75^{th} percentile	62.11	56.81	41.78
90^{th} percentile	116.32	104.49	91.26
n=	16939	16939	2414

Source: SCAG Travel Diary, authors' calculations. Diary is one-day for 14,517 households, and two-day for 2,422 households. Eight of the two-day households had missing information that did not allow VMT to be estimated. "Total" is for the full diary period, either one day or two day. Households that completed a two-day diary did so for one weekday and one weekend day, allowing VMT descriptive statistics for weekdays (from the full sample) and weekend days (from the households who completed two-day diaries), as shown in the table.

Table 2 shows the distribution of total VMT, for all trips in the diary, by trip length. For comparison, Table 2 shows the distribution of total VMT traveled by trip length from the SCAG diary, the 2001 National Household Travel Survey (NHTS) Los Angeles consolidated metropolitan area sub-sample, and the national NHTS. Note first that the cumulative distributions are similar for all three data sources, providing some assurance that any inaccuracy in network routing is, in the aggregate, likely small. More importantly, Table 2 illustrates the importance of long trips in total VMT. Both the NHTS and the SCAG diary show that approximately 40 percent of all VMT in the L.A. region is from trips of 30 miles or longer. This suggests that land use innovations that influence short trips may be less effective than land use policies that influence long trips if the goal is reducing greenhouse gas emissions. We note, though, that California's SB 375 is primarily a regional planning tool, targeting transportation infrastructure and affordable housing, although the specific implementation has yet to unfold. We also note that Table 2 suggests that attention to regional accessibility may be more important than attention to neighborhood-level land use measures, although for completeness we include both in our analysis.

Table 2: VMT by Trip Length

Trip Length (Miles)	SCAG L.A. Percent of VMT	SCAG L.A. Cumulative Percent	NHTS L.A. CMSA Percent of VMT	NHTS L.A. CMSA Cumulative Percent	NHTS National Percent of VMT	NHTS National Cumulative Percent
0 - 2	2.96	2.96	1.72	1.72	1.57	1.57
2 - 4	5.75	8.71	5.12	6.84	5.48	7.06
4 - 6	5.62	14.33	7.35	14.20	6.48	13.54
6 - 8	5.14	19.46	5.92	20.12	5.70	19.24
8 - 10	4.73	24.20	4.20	24.32	4.70	23.94
10 - 20	20.87	45.07	21.91	46.23	21.80	45.75
20 - 30	14.20	59.27	16.92	63.16	13.78	59.53
30 - 50	16.04	75.30	14.86	78.02	14.16	73.68
50 - 100	15.19	90.49	15.02	93.04	11.05	84.73
100 - 200	6.56	97.05	4.10	97.14	7.03	91.76
> 200	2.95	100.00	2.86	100.00	8.24	100.00

Figure 1 shows a VMT surface for the 11,218 survey households that reside in the two largest urbanized areas (Los Angeles – Long Beach – Santa Ana and Riverside – San Bernardino) in the SCAG region. The map shows survey household VMT, smoothed by interpolating between household locations. A spatial pattern is clearly evident. The largest area of dark blue (the lowest VMT quintile) is the central and south-central part of the City of Los Angeles, with another large dark blue area near Santa Ana and Irvine in central Orange County. Parts of both places have exceptional job accessibility. Studies of sub-centers (e.g. Funderburg and Boarnet, 2008; Redfearn, 2007) suggest that downtown Los Angeles and central Orange County are among the two largest and most economically complex job centers in the metropolitan area. Income also is an important explanatory factor, as south-central Los Angeles and Santa Ana are lower income areas, but the locations of low VMT extend into high income, job accessible locations such as Irvine and Santa Monica. The red areas on the fringe of the region suggest that, as expected, households in more exurban locations have higher VMT.

Figure 1: Interpolated VMT Quintiles, Households in the Los Angeles - Long Beach – Santa Ana and Riverside – San Bernardino Census Urbanized Areas

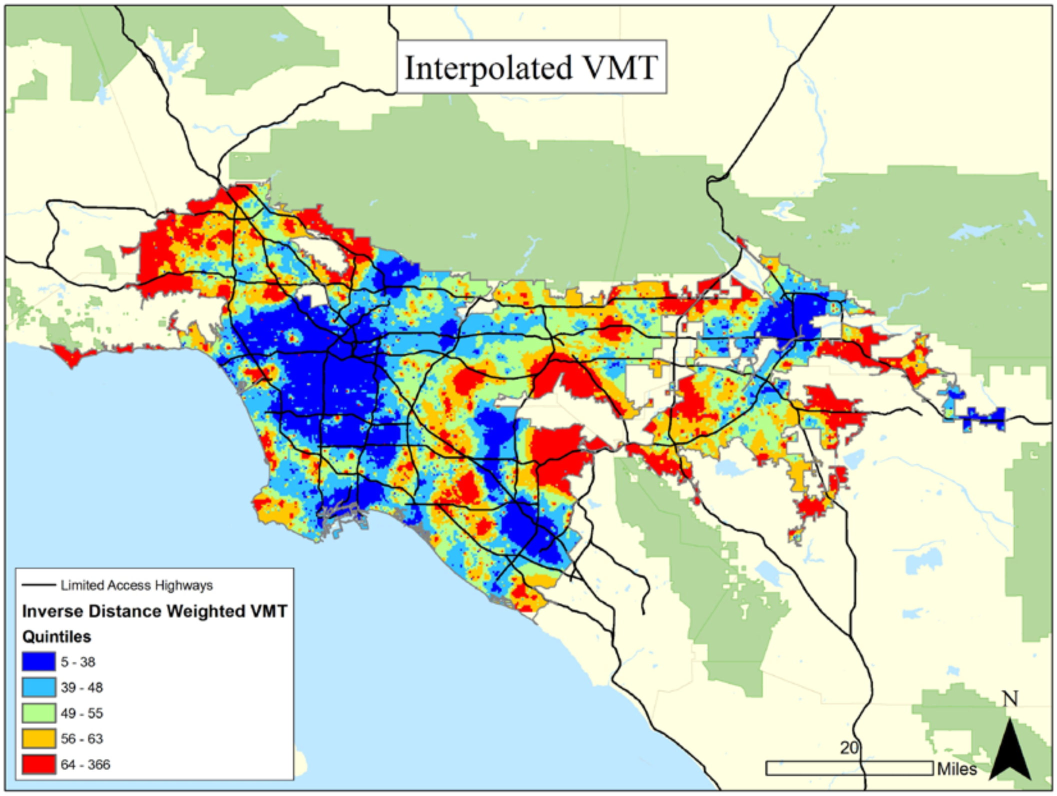


Table 3 shows VMT by county. The inner counties of Los Angeles and Orange have lower household VMT. Table 4 shows VMT split according to whether the household lives inside a Compass 2 percent opportunity area. As Table 4 shows, households that lived in places that would later be designated as Compass opportunity areas had lower household VMT than other households. Table 5 splits the sample by ¼ mile distances from rail or bus stops, and Table 6 splits the sample by ½ mile distances from rail, express bus, and rapid bus stops. All the VMT differences in Tables 4-6 are statistically significant using two-sample t-tests with unequal variances across groups. (A small number of households could not be geocoded to locations, so the sample in Tables 4-6 includes 16,821 households.)

Table 3: Household VMT by County, diary period

County	*Obs*	*Mean*	*Std. Dev.*
Imperial	897	35.36	66.5
Los Angeles	7222	44.1	74.14
Orange	2304	48.32	77.53
Riverside	2324	54.47	74.63
San Bernardino	2155	57.57	91.95
Ventura	1919	51.13	80.34

Table 4: Household VMT, inside and outside of Compass Opportunity Areas

	in Compass opportunity area	*Outside Compass opportunity area*
Mean	36.79	52.99
st dev	61.54	83.17
Obs	5005	11816
t-stat		13.98

Table 5: Household VMT, split at ¼ mile from rail or bus station

	< 1/4 mi from rail	*> 1/4 mi from rail*	*< 1/4 mi from bus*	*> 1/4 mi from bus*
Mean	21.05	48.39	44.58	55.18
st dev	32.63	77.94	77.31	78.04
Obs	137	16684	11122	5699
t-stat		9.59		8.36

Table 6: Household VMT: split at ½ mile from rail, express bus, or rapid bus station

	< 1/2 mi from rail	*> 1/2 mi from rail*	*< 1/2 mi from express bus*	*> 1/2 mi from express bus*	*< 1/2 mi from rapid bus*	*> 1/2 mi from rapid bus*
Mean	25.49	48.98	42.23	50.18	32.9	50.51
st dev	46.93	78.48	64.86	81.51	63.17	79.45
Obs	583	16238	4250	12571	2233	14588
t-stat		11.53		6.46		11.82

IV. Methodology

We use a standard land use – travel behavior regression to examine how VMT is related to land use. The basic form of the regression is shown below.

$$VMT_i = \mathbf{SD}_i\beta_1 + \mathbf{LU}_i\beta_2 + u_i \quad (1)$$

Where VMT = household vehicle miles of travel during the travel diary period (one or two days),

SD = a row vector of sociodemographic variables for the household or the primary respondent in the household,

LU = a row vector of land use variables measured at the household's place of residence,
u = regression error term,
i indexes households,
and β_1 and β_2 are column vectors of parameters, with one including a constant.

A fairly extensive set of sociodemographic variables are available, and are shown in Appendix A with descriptive statistics for observations for which all sociodemographic variables are available. The sociodemographic variables compare favorably to the most extensive sets of variables used in the literature (e.g. Bhat and Guo, 2007). Note that while we make no formal claims, Brownstone (2008) has suggested that including an extensive set of sociodemographic variables in the travel behavior regression might be sufficient to control for residential selection.

Past literature has argued that land use variables should be divided into two geographic scales – localized, neighborhood-scale measures of development patterns and measures of the overall metropolitan settlement pattern (e.g. Handy, 1993). Much of the literature has been silent on this distinction, but the evidence that is available suggests that this local/regional distinction is important (e.g. Boarnet and Sarmiento, 1998; Ewing and Cervero, 2001 and 2010). The two levels of geography correspond to different policy agendas. The neighborhood scale corresponds to discussions of Smart Growth or compact development, while the regional scale links more closely to metropolitan growth patterns, including recent research on the pattern of employment centering and centralization within urban areas (e.g. Giuliano et al., 2008).

We believe the local – regional distinction is vital, and we choose one variable as a key measure of each of the local and regional land use patterns. Population density within ¼ mile of the place of residence is the local land use measure, and employment accessibility (the gravity variable with linear straight-line distance damping) is the regional accessibility measure. We examine thresholds and interactions for both of those two variables, to examine thresholds in local accessibility, thresholds in regional accessibility, and interactions between local and regional accessibility. Population density and employment accessibility were chosen because those two variables are, respectively, the most common measures of local and regional accessibility in the literature on land use and travel, so a focus on thresholds in those two variables more easily allows comparisons to the literature. Other choices are possible, and in a fully generalized approach one could run a second-order approximation of a trans-log functional form for the land use variables, using linear and quadratic terms and interactions for all land use variables. We did not pursue that approach as such specifications become difficult to interpret for cases like ours with a large number of variables, and at this initial exploratory stage the simplification from examining thresholds and interactions in two variables has benefits in suggesting patterns and possible directions for future research.

The question of thresholds or non-linearities was examined three decades ago by Pushkarev and Zupan (1977) who studied population density thresholds and transit ridership. Yet since then, the question of thresholds and non-linearities has been overlooked, leaving the modern land use – travel literature to statements of average effects for data sets that typically cover metropolitan areas or larger geographies.

Most of the analysis in the next section examines questions of thresholds in population density, employment accessibility, or interactions between the two. Yet we also examine thresholds for transit access and highway access. We examined VMT across threshold distances of a mile or less to transit, and threshold distances of ten miles from freeways as exploratory analyses suggested that those were key breakpoints in the data, and results for transit and highway distance are presented at the end of the next section.

V. Regression Analysis and Results

Regression Specifications

We used two regression approaches to examine the question of non-linearities. In the first (and simpler) approach, we stratified the full sample into quintiles by population density and employment accessibility, and then estimated the full regression shown in equation (1) on the households in each quintile.[9] This entails two sets of regressions, first running the regression in equation (1) on the sub-samples of households in the five different population density quintiles, and then running the regression in equation (1) on the sub-samples of households in the five different employment accessibility quintiles. We call this the "stratified sample" approach.

As a second specification, we developed spline variables for population density and employment accessibility to implement a piecewise regression approach, as described in, e.g., Pindyck and Rubinfeld (1981, pp. 126-127). Specifically, we estimate the regression shown below, which we call the "spline regression" approach.

$$VMT = X\gamma + \sum_{q=1}^{5} \alpha_q PD^q + \sum_{q=1}^{5} \beta_q EA^q + u \qquad (2)$$

Where X = the SD + LU matrices from equation (1), without the variables for population density (denoted PD) and employment accessibility (denoted EA).

The five spline variables EA^q are defined for each quintile $q \in [1,5]$,

$EA^q = 0$ if EA_i is below quintile "q"
$EA^q = (EA_i - EA_{q-1})$ if EA_i is within quintile "q"

[9] The quintiles are defined relative to the SCAG household travel survey, and so the population density and employment accessibility quintiles represent the distribution of SCAG survey households, which will depart from the underlying geography of the SCAG region to the extent that the 2001 travel survey sampled households non-uniformly across density or employment access. As a sensitivity test, it would be sensible to examine other groupings, such as deciles, both to better illuminate non-linearities in the impact of land use variables and to test whether the results obtained here are sensitive to the number of groups. Due to time constraints, we do not examine other groupings in this paper, but we believe such sensitivity tests are an important question for future research.

$EA^q = (EA_q – EA_{q-1})$ if EA_i is above quintile "q"

Where EA_i is the EA value for each household, EA_q is the highest value for EA^i within each quintile "q", EA_{q-1} is the highest value for EA_i in the previous quintile, and $EA_{q-1} = 0$ for the first quintile, q = 1. Hence, EA^1 = EA for each household in quintile 1 and the maximum value of EA in quintile 1 for all households in higher quintiles, EA^2 is zero for households in quintile 1, the difference between EA_i and the upper bound of EA in quintile 1 for households in quintile 2, and the difference between the maximum and minimum values of EA in quintile 2 for households in quintiles higher than the 2nd, and so on. The spline variables for PD are defined in the same way.

The regression in equation (2) is a special case of a spline function, allowing the estimated coefficients to vary across the quintiles while maintaining the continuity of the relationship between VMT and both PD and EA at the quintile breakpoints.[10] Testing the statistical significance of each of the spline variables, PD^q or EA^q, tests the hypothesis that the slope for the quintile "q" is different from the slope in the previous quintile-segment, "q-1". The regression in equation (2) thus allows simple tests for the hypothesis that the relationship between PD and the latent variable for VMT or the relationship between EA and the latent variable for VMT varies across quintiles.[11] Those formal tests are not possible in the stratified sample approach to estimating equation (1), as the stratified samples are not nested regressions. The regression in equation (2) also has more degrees of freedom, using all of the households rather than running the regression separately on each quintile as in the stratified sample approach. Yet the regression in equation (2) does not allow interaction effects between PD and EA, which can be examined in the stratified sample approach.[12] We report results from both approaches here.

Marginal Effects and Elasticities

It has become common to use elasticities to quantify the effect of land use variables on travel (e.g. Ewing and Cervero, 2010; Tal, Handy and Boarnet, 2010). The elasticity of VMT with respect to PD is $\frac{\partial VMT}{\partial PD} \cdot \frac{PD}{VMT}$ and the elasticity of VMT with respect to EA is $\frac{\partial VMT}{\partial EA} \cdot \frac{EA}{VMT}$. For linear regression specifications, if the slope (either $\frac{\partial VMT}{\partial PD}$ or $\frac{\partial VMT}{\partial EA}$) is constant, the elasticity will change as PD/VMT or EA/VMT changes. Stated

10 In general, spline functions are piecewise relationships of continuous functions that may include higher order terms. Equation (2) represents the special case where the piecewise relationships are linear within the quintiles.

11 Because we use Tobit regression, the coefficients estimated in equations (1) and (2) show the relationship between the independent variables and the latent variable for VMT rather than observed VMT. This is discussed more later in this section.

12 It would be possible to interact the spline variables with each other, but that creates 25 possible pairs of interactions, which did not reveal insights beyond what can be obtained from the stratified sample approach and so are not reported here.

more generally, elasticties vary for linear relationships as one moves along the line. For that reason, we examine both the marginal effects, ∂VMT/∂PD or ∂VMT/∂EA, and the elasticities, to assess whether changes in elasticities across the quintiles are more due to changes in marginal effects or changes in the values for PD, EA, and VMT.

Because VMT for our sample of households is left-censored, we use Tobit regression for both the stratified sample and spline regression approaches. Because Tobit regression is non-linear, estimating elasticties at sample means (either for the full sample or at means within quintiles) can be misleading (e.g. Brownstone, 2008). Both terms in the elasticity, the marginal effects, ∂VMT/∂PD or ∂VMT/∂EA, and the ratios PD/VMT and EA/VMT, vary for each household. We calculate the marginal effect and the elasticity for each household, and then average those marginal effects and elasticities for each quintile to obtain average marginal effects and elasticities within quintiles. The marginal effect for the stratified sample approach is shown below.[13]

$$me_i = \beta_k \Phi\left(\frac{X_i \gamma}{\sigma}\right) = \frac{\partial VMT_i}{\partial X_k} \qquad (3)$$

Where Φ = the cumulative normal probability distribution

X = the full vector of independent variables in equation (1)

σ = standard error of the regression

X_k = the land use variable, either PD or EA

β_k = coefficient on the land use variable

i indexes households.

Note that the expression for me_i is listed as the partial derivative of VMT with respect to the land use variable. The expression for me_i in equation (3), and the analogous expression for me_i for spline regressions given in equation (4) below, give marginal effects of the land use variables on VMT including the effect on those households with zero VMT – in other words, the marginal effect for the full sample used in the regressions.

Within each quintile, the average marginal effect and elasticity are then calculated as the average of the household values within that quintile.

$$me_q = \frac{1}{N_q}\sum_{j=1}^{N_q} me_i$$

$$elasticity_q = \frac{1}{N_q}\sum_{j=1}^{N_q} me_i \frac{lu_i}{VMT_i}$$

Where me = marginal effect

q indexes quintiles

N_q = number of households in quintile q

[13] This is equivalent to the regression coefficient on the land use variable, PD or EA, multiplied by the probability that household "i" has non-zero VMT. See, e.g., Johnston and Dinardo (1997, p. 437).

lu_i = land use variable (either PD or EA) for household "i"
VMT_i = VMT for household i.

For the spline regression, the marginal effect is shown below.

$$me_i = \beta_k \Phi\left(\frac{X_i^{-}\gamma + \sum_{j=1}^{5}\beta_j lu_j}{\sigma}\right) = \frac{\partial VMT_i}{\partial X_k} \tag{4}$$

Where X^{-}_i is the set of independent variables in equation (1) less the five spline variables for the land use variable in question, either PD or EA
lu = PD or EA
$\sum_{j=1}^{5}\beta_j lu_j$ = the sum of the five quintile variables for the household "i"
and, as before, Φ is the cumulative normal probability distribution, β_k is the coefficient on the land use variable for the quintile of PD or EA that the household is in, and σ is the standard error of the regression.

As before, within each quintile the marginal effects and the elasticities for households are averaged to obtain average marginal effects and elasticities for the quintile.

Results

Results from the two specifications – the stratified sample approach and the spline regression approach – are shown in Tables 7-9. Those tables show the coefficients (from the Tobit regression) for population density and employment accessibility only. All the regression coefficients from a version of equation (1) fit on the full sample are shown in Appendix B.[14] The marginal effects and elasticties are presented in Tables 10 and 11. In Table 7, each row shows results from fitting equation (1) on the five population density sub-samples. Table 7 shows the population density ranges for each quintile. Note that quintiles were split based on the full sample of 16,939 households, and given that some households have missing variables and were excluded from the regression analysis the number of households is not identical across the stratified samples used in the regressions. Table 8, similar to Table 7, shows results for the coefficients on the population density and employment accessibility variables when the sample is stratified by quintiles for the employment gravity variable.

[14] The results in Appendix B are consistent with the literature. Note that the coefficient on distance from downtown Los Angeles in Appendix B is negative, counter to simple expectations that persons living distant from the central business district drive more. The regression also controls for access to employment through the gravity variable, which likely better controls for accessibility to Los Angeles' decentralized employment.

Table 7: Stratified Sample Approach:
Neighborhood Population Density and Regional Employment Accessibility Coefficients, by Population Density Quintile

Population density quintile	Minimum density (persons/sq mile in surrounding ¼ mile)	Maxium density (persons/sq mile in surrounding ¼ mile)	Number of observations	Population density, ¼ mile area		Gravity variable for employment accessibility	
				Coefficient	t-statistic	Coefficient	t-statistic
1	0.00	2578.34	2139	-0.0004802	-0.18	**-0.1290879**	**-2.41**
2	2583.44	5034.40	2330	0.0003823	0.12	**-0.1259769**	**-2.24**
3	5039.49	7485.35	2435	0.0021975	0.77	*-0.1092746*	*-1.84*
4	7490.45	12122.29	2522	0.0002719	0.21	-0.034606	-0.69
5	12127.39	72952.87	2603	-0.0001099	-0.57	**-0.1140818**	**-2.62**
Full Sample	0.00	72952.87	12029	0.0000316	0.18	**-0.1142238**	**-5.55**

Note: Coefficients that are statistically significant at 5 percent level (two-tailed test) are in bold, and those significant at 10 percent level are in italics.

Table 8: Stratified Sample Approach:
Neighborhood Population Density and Regional Employment Accessibility Coefficients, by Employment Gravity Variable Quintile

Emp. gravity variable quintile	Minimum of employment gravity variable	Maximum of employment gravity variable	Number of observations	Population density, ¼ mile area		Gravity variable for employment accessibility	
				Coefficient	t-statistic	Coefficient	t-statistic
1	19.33	87.54	1809	0.0020428	1.54	-0.2382835	-0.55
2	87.56	135.97	2518	-0.0000936	-0.15	-0.0394491	-0.25
3	135.98	237.36	2551	-0.0004998	-1	**-0.1694827**	**-2.25**
4	237.38	280.50	2560	0.0000902	0.22	*-0.3620431*	*-1.94*
5	280.51	768.98	2591	-0.0000835	-0.4	0.0039499	0.08
Full Sample	19.33	768.98	12029	0.0000316	0.18	**-0.1142238**	**-5.55**

Note: Coefficients that are statistically significant at 5 percent level (two-tailed test) are in bold, and those significant at 10 percent level are in italics.

Table 9 shows results from the spline regression approach, fitting equation (2) on the full sample. The general pattern of results in Table 9 is similar to the pattern from Tables 7 and 8. Population density is never significant, while employment accessibility is significant in the full sample regression for equation (1). The elasticity of VMT with respect to employment accessibility implied by the full sample estimate from Table 7 is -0.29 (see Table 10). That "full sample" elasticity is similar to other estimates in the literature (see, e.g., Ewing and Cervero, 2010, or Tal, Handy, and Boarnet 2010.) Tables 7 and 8 give little evidence of interactions between population density and employment accessibility, so we turn our attention to threshold or non-linear effects of employment accessibility across different ranges of the employment gravity variable.[15]

Table 9: Spline Regression Approach: Coefficients on population density and employment gravity variables

		Coefficient	t-stat
Population density quintile	1	-0.00040	-0.21
	2	0.00006	0.04
	3	0.00163	1.12
	4	-0.00061	-0.81
	5	-0.00001	-0.04
Employment accessibility quintile	1	-0.10523	-0.48
	2	**-0.19481**	**-2.43**
	3	**-0.08462**	**-2.01**
	4	**-0.25612**	**-2.96**
	5	-0.03545	-0.77

Note: Coefficients that are statistically significant at 5 percent level (two-tailed test) are in bold.

From the spline regression results for employment accessibility in Table 9, the Tobit regression slope on the employment gravity variable changes across each employment gravity variable quintile boundary except when moving from the fourth to the highest quintile. Yet the pattern is not monotonic, as the magnitude of the coefficient becomes smaller when moving from the second to the third quintile and then grows larger in the fourth quintile. The coefficients in Table 9 show impacts on the latent variable for the Tobit regression, and for information on marginal effects and elasticities we turn our attention to Tables 10 and 11.[16]

[15] Note, though, that the coefficients in Tables 7 and 8 are not marginal effects, so while the coefficient on employment accessibility does not visually change much across different population density quintiles, the marginal effects could change across population density quintiles. Table 10, though, shows not much variation in the marginal effect of employment accessibility across population density quintiles.

[16] Note, though, that the spline regression results reject the hypothesis that the employment accessibility coefficient is the same for the first four quintiles, verifying a non-linear or piecewise relationship between the VMT latent variable and the employment accessibility gravity variable.

Table 10: Marginal Effects and Elasticities, by population density quintiles

	Effect of Population Density						Effect of Employment Accessibility					
	Basic Regression		Stratified Regressions		Spline Regression		Basic Regression		Stratified Regressions		Spline Regression	
Pop. Density quintile	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity
1	0.00002	0.00043	-0.00034	-0.00671	-0.00028	-0.00545	**-0.08024**	**-0.14048**	**-0.09015**	**-0.16218**	-0.09387	-0.16782
2	0.00002	0.00142	0.00026	0.01570	0.00004	0.00271	**-0.07977**	**-0.20197**	**-0.08408**	**-0.20257**	**-0.09843**	**-0.24788**
3	0.00002	0.00232	0.00146	0.15125	0.00112	0.11895	**-0.07830**	**-0.25662**	*-0.07266*	*-0.23147*	**-0.10274**	**-0.32668**
4	0.00002	0.00376	0.00018	0.03352	-0.00039	-0.07195	**-0.07358**	**-0.33296**	-0.02296	-0.10473	**-0.09926**	**-0.41089**
5	0.00002	0.00893	-0.00007	-0.04144	-0.00001	-0.00278	**-0.06404**	**-0.46175**	**-0.07012**	**-0.61506**	-0.06403	-0.38447
Full Sample	0.00002	0.00354	0.00031	0.03053	0.00010	0.00795	**-0.07485**	**-0.28578**	-0.06701	-0.26998	-0.09122	-0.31333

Note: Marginal effects and elasticities shown in bold (italics) are based on Tobit regression coefficients that are significant at the 5 percent (10 percent) two-tailed level. The "full sample" marginal effects and elasticities in the stratified regression and spline regression columns are averages for all households across the five quintiles, and so are based on five regression coefficients, some of which are significant for employment accessibility and some not. The full sample marginal effect and elasticity for the basic regression is based on the coefficient estimate from using Tobit regression on equation (1) for the full household sample.

Table 11: Marginal Effects and Elasticities, by employment accessibility (employment gravity variable) quintiles

	Effect of Population Density						Effect of Employment Accessibility					
	Basic Regression		Stratified Regressions		Spline Regression		Basic Regression		Stratified Regressions		Spline Regression	
Emp. access quintile	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity	Mean marginal effect	Mean elasticity
1	0.00002	0.00104	0.00004	0.05623	-0.00004	0.00864	**-0.07689**	**-0.09035**	-0.08127	-0.15909	-0.07096	-0.08302
2	0.00002	0.00195	0.00040	-0.00563	0.00017	0.01834	**-0.07997**	**-0.14418**	-0.07482	-0.04850	**-0.13654**	**-0.24593**
3	0.00002	0.00245	0.00049	-0.04363	0.00021	0.01712	**-0.07989**	**-0.24962**	**-0.06974**	**-0.41355**	**-0.05934**	**-0.18428**
4	0.00002	0.00431	0.00037	0.01218	0.00010	0.00135	**-0.07376**	**-0.36713**	*-0.05698*	*-1.15600*	**-0.16482**	**-0.82639**
5	0.00002	0.00715	0.00016	-0.02291	0.00003	-0.00514	**-0.06457**	**-0.51504**	-0.05671	0.02152	-0.02000	-0.15975
Full Sample	0.00002	0.00354	0.00031	-0.00432	0.00010	0.00795	**-0.07485**	**-0.28578**	-0.06701	-0.36316	-0.09122	-0.31333

Note: Marginal effects and elasticities shown in bold (italics) are based on Tobit regression coefficients that are significant at the 5 percent (10 percent) two-tailed level. The "full sample" marginal effects and elasticities in the stratified regression and spline regression columns are averages for all households across the five quintiles, and so are based on five regression coefficients, some of which are significant for employment accessibility and some not. The full sample marginal effect and elasticity for the basic regression is based on the coefficient estimate from using Tobit regression on equation (1) for the full household sample.

In Tables 10 and 11, we show the average marginal effects and elasticities for households within quintiles for both the stratified sample and spline regression approaches. For information, the quintile averages for population density, employment accessibility, and VMT are shown in Table 12. Given that no coefficient on population density was statistically significant in any regression, we discuss only the marginal effects and elasticities for the employment accessibility gravity variable. Three sets of marginal effects and elasticities are shown, labeled "basic regression", "stratified regressions", and "spline regression." The marginal effects and elasticities for the basic regression use the full sample regression coefficients from the bottom row of Table 7 or 8, and then evaluate the marginal effects and elasticties for each observation (or household) and average within the quintiles. The stratified regression and spline regression columns use the coefficients for each quintile from the respective approach.

Visual inspection shows that the marginal effects in the "basic regression" column vary little across quintiles. The marginal effects in the "stratified regression" column also vary little across quintiles, but the corresponding elasticities show substantial change. For example, the elasticity of VMT with respect to employment accessibility is -0.29 for the full sample, -0.41 for households in the third quintile of employment accessibility, and -1.16 in the fourth quintile using comparisons to the stratified regression approach in Table 11. Looking at the "spline regression" column of Table 11, the marginal effect of employment accessibility is -0.16 for households in the fourth quintile of employment accessibility, compared to -0.07 using equation (1) fit on the full sample, and the elasticity of VMT with respect to employment accessibility is -0.83 in the fourth quintile.

Table 12: Means, by population density quintile and by employment accessibility quintile

	Descriptive Statistics		
Population density quintile	Mean population density	Mean employment gravity variable	Mean household VMT
1	1177.74	108.59	58.53
2	3898.63	153.30	57.52
3	6230.60	188.90	53.90
4	9518.46	231.17	47.62
5	19897.63	286.98	32.69
Full Sample	8527.19	197.81	49.52

	Descriptive Statistics		
Employment accessibility quintile	Mean population density	Mean employment gravity variable	Mean household VMT
1	2717.01	66.56	54.93
2	5208.96	109.82	57.72
3	6555.44	189.20	55.77
4	10556.06	258.31	46.94
5	15745.24	323.66	34.16
Full Sample	8527.19	197.81	49.52

Note: Mean values are for the subset of households for which all data are available and hence which were used to estimate the regression. Mean values in Section III are for the full sample of 16,939 households.

Overall, two conclusions are evident. Employment accessibility is much more important than population density as a determinant of VMT in the SCAG data studied here, and the impact of employment accessibility on VMT is largest for households in the third and fourth quintile of employment accessibility. While the evidence suggests that the larger part of the changes in elasticity across quintiles is from changes in the household values for the employment accessibility and VMT variables, in the fourth quintile the spline regression marginal effect of the employment gravity variable is over twice as large as the full sample marginal effect from the basic regression. The elasticities in those quintiles, as large as -0.83 (spline regression) and -1.16 (stratified sample) are approximately three to four times as large as both common elasticities in the literature and our full sample average elasticity from the basic regression. On net, we conclude that examining departures from regression averages is important for land use – transportation research, and the third and fourth quintiles of employment accessibility appear to be fruitful places for policy attention.

Figure 2 shows employment accessibility quintiles, by household location. The third and fourth quintiles are in green and yellow, respectively. The regression results suggest a focus on increasing employment accessibility in those locations, or increasing the number of persons who live in those places. This would direct policy attention primarily to north and central Orange County, South Central Los Angeles, and some locations in the San Fernando Valley and near Ontario and Riverside in what is called the Inland Empire (Riverside and San Bernardino Counties.)

Figure 2: Employment Accessibility Quintiles, by Household Location

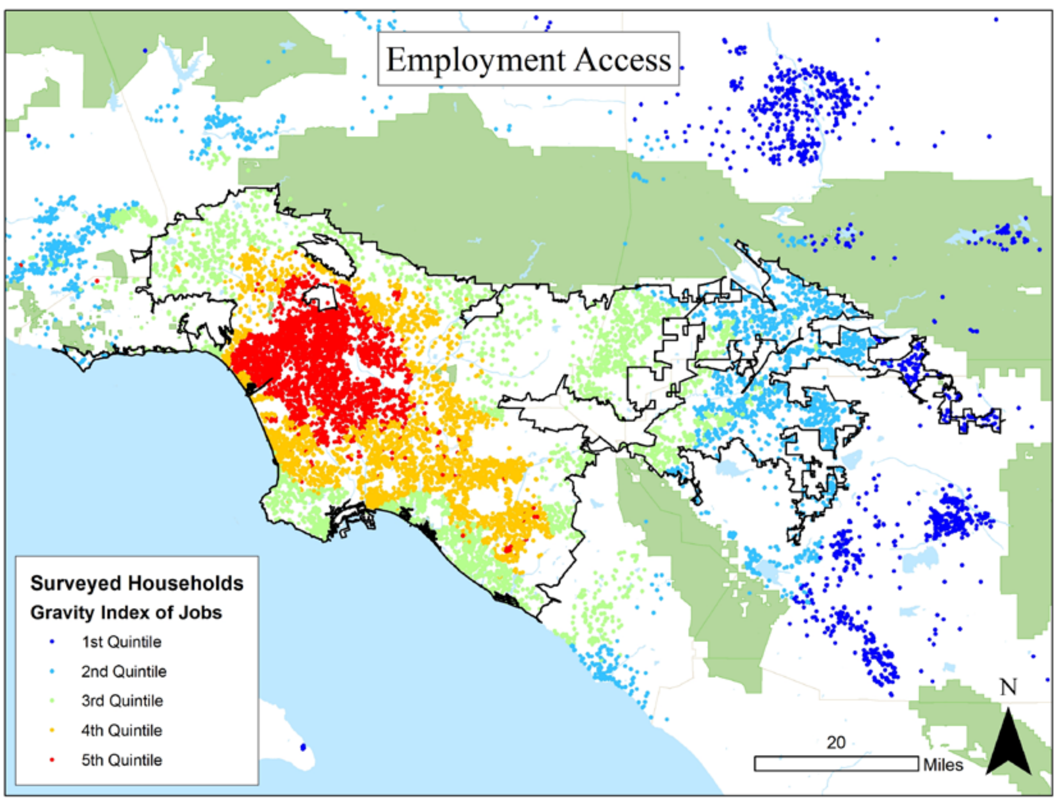
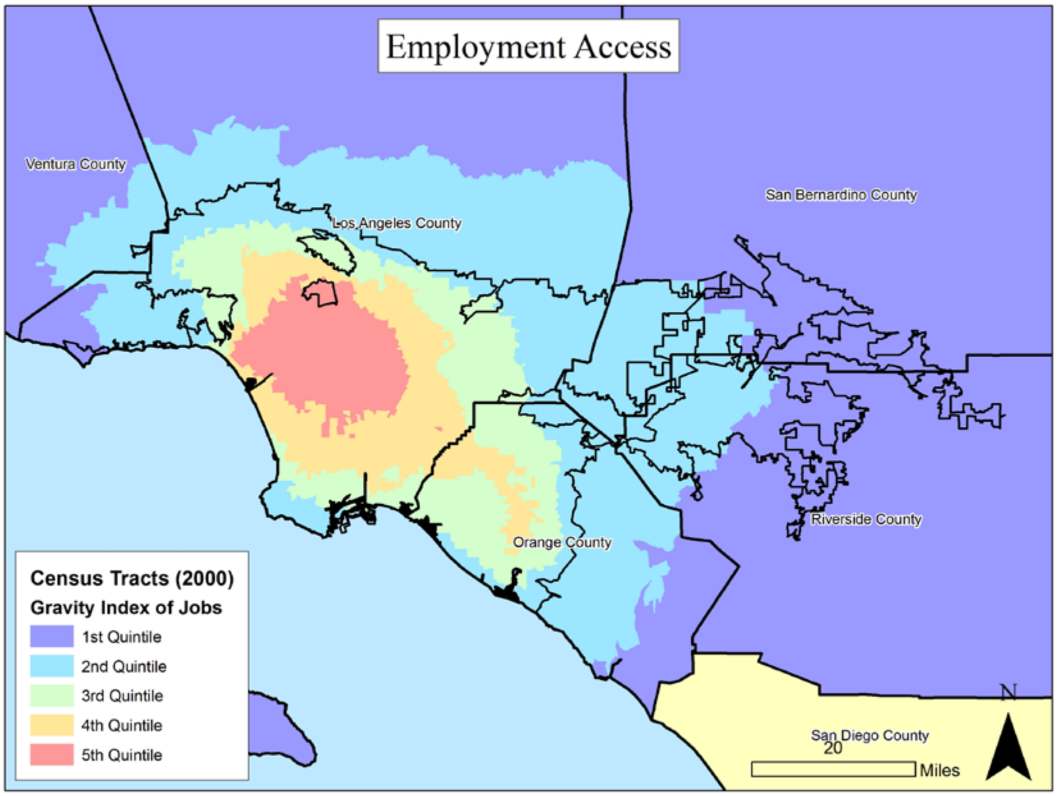


Figure 3 shows employment accessibility based on census tracts, not surveyed households. That view would tighten the policy focus to the central spine of Orange

County and areas extending from downtown Los Angeles. In general, we believe that a focus on employment accessibility would lead to prominent attention on the region's job sub-centers.

Figure 3: Employment Accessibility Quintiles, by Census Tracts



Lastly, we briefly present results for thresholds by distance from rail and bus and by distance from freeway. For distance from rail transit, we ran the full regression specification from equation (1), including dummy variables that indicate the concentric distance categories less than 1/8 mile from a rail station, 1/8 to 1/4 mile, 1/4 to 1/2 mile, 1/2 to 3/4 mile, and 3/4 to 1 mile from a station. The dummy variables stopped at 1 mile; we did not test for effects at farther distances. We then ran the full regression specification with the same concentric circle dummy variables for bus station distance. In both regressions, the dummy variable indicating whether the household lived within a Compass opportunity area was dropped, as the Compass areas overlap with transit access, by construction, and so colinearity might mask any effects of transit access. Results are shown, for the concentric circle dummy variables only, in Table 13.

Table 13: Rail and Bus Concentric Circle Dummy Variables (full regression without Compass dummy, once each for rail and bus)

	rail concentric circle variables		*bus (any) concentric circle variables*	
	coeff	t-statistic	coeff	t-statistic
< 1/8 mile	-11.1903	-0.48	**-7.7698**	**-2.10**
1/8 - 1/4 mile	-12.0599	-1.12	-4.0911	-1.12
1/4 - 1/2 mile	-9.0098	-1.63	**-8.2585**	**-2.20**
1/2 - 3/4 mile	-1.5748	-0.33	*-8.0430*	*-1.69*
3/4 - 1 mile	2.7003	0.64	-0.9640	-0.16

Note: Coefficients shown in bold are significant at the 5% level (two-tailed test). Coefficients shown in italics are significant at the 10% level (two-tailed test).

The results show no significant association between rail access and VMT. This might be due to the small number of households surveyed near existing rail stations in 2001-2002. For example, there were 139 surveyed households within ¼ mile of a rail station, of which 25 were within 1/8 of a mile. Living near a bus station is associated with lower VMT. The 1/8 to 1/4 mile ring is not significant, which is odd, but ignoring that the pattern appears to be a relatively unchanged slope coefficient out to 3/4 mile from the bus station, beyond which the relationship is statistically insignificant.

We found that distance from nearest freeway was highly non-linear. Both quadratic and cubic terms were significant in a regression with the full set of sociodemographic and land use variables. The non-linearity appears to be driven by outliers, based on our initial analysis. When the sample is split into two groups, depending on whether the household is within 10 miles from the nearest freeway on-ramp, only the linear term for distance from the freeway is significant, and then only for households that are within 10 miles of a freeway on-ramp. The results are shown in Table 14 below, which suggests the intuitive finding that being nearer a freeway is associated with reduced VMT, but only for households that are within 10 miles of a freeway.

Table 14: Coefficient on Distance from Freeway, which is measured in meters

	coeff	*t-statistic*	*N*
< 10 miles from nearest freeway	**0.0011**	**3.06**	12396
> 10 miles from nearest freeway	0.0001	0.17	558
full sample	**0.0004**	**2.80**	12594

Note: Coefficients shown in bold are significant at the 5% level (two-tailed test).

VI. Interpretation

Consistent with previous research, our results suggest that land use policies aimed at reducing VMT should focus on employment accessibility as opposed to neighborhood population density.[17] Largely untested in previous work, our results also suggest that the link between employment accessibility and VMT varies in a non-linear (threshold) fashion. The "mid-range" locations – in the third and fourth quintile of employment accessibility – are where the association between VMT and employment accessibility is strongest.

One narrative that has developed in the land use – travel debate is that the magnitude of any associations, even if they are causal, are too small to be important for policy purposes (see, e.g. Brownstone, 2008 and Brownstone and Golob, 2009). If the results in this paper are confirmed by further analysis, such a view would have to be tempered. The point is not that land use is a weak tool for VMT, but that land use may be either a weak or a meaningful tool depending on where policy is focused. Such a relationship is intuitive, and was conjectured by the National Research Council's (2009) report "Driving and the Built Environment," but to the best of our knowledge the evidence on thresholds in this paper gives the first quantified illustration of such an effect. The results of this paper give a threefold message: (1) regional access is a more appropriate focus than neighborhood population density if the objective is VMT or GHG reduction, (2) some places will yield a stronger association with VMT than others, (3) and, in the Los Angeles region, a focus on employment sub-centers is likely a fruitful path for future research and policy.

What, then, of California's ambitious land use – transportation planning required by SB 375? SB 375 is a work in progress, and the question of what implementation will ultimately entail has yet to be worked through. At the time this chapter was written, in summer of 2010, the California Air Resources Board had just released draft transport-sector GHG emission targets for the states eighteen Metropolitan Planning Organizations (MPOs.) MPOs are to develop, by the fall of 2011, sustainable community strategies that document how the MPO will comply with the transport-sector GHG targets for the years 2020 and 2035. While there has been much speculation about the impact of SB 375, we note that all of that speculation has occurred in advance of the MPO plans, which are being drafted and will not be delivered until, most likely, the last quarter of 2011.

While SB 375 is a work in progress, we suggest two possible paths for SB 375. One path would focus directly on employment accessibility, linking residences to job and activity centers through a combination of transportation infrastructure and land use planning. Such an approach would be consistent with the evidence in this paper and, more importantly, is consistent with the legislative language of SB 375. Briefly, SB 375 requires MPOs to develop plans (sustainable community strategies) that demonstrate that the regional transportation plan (RTP) and the regional housing needs assessment (RHNA) are both consistent with GHG emission reduction targets. The RTP is an

[17] For the 16,939 households in the SCAG travel survey, the correlation coefficient for population density and employment accessibility is 0.60. That illustrates how many land use variables are correlated, and population density has often been used in the literature to proxy a range of land use characteristics. Our results illustrate that, notwithstanding those correlations, VMT was only statistically related to employment accessibility, echoing the finding of the larger literature that the magnitude of the VMT – accessibility relationship is larger than the magnitude of the VMT – neighborhood population density relationship.

MPO's program of infrastructure investment, possibly including travel demand management techniques. The RHNA is a plan to meet fair-share affordable housing requirements established by the State. In short, SB 375 is a requirement that transportation planning be coordinated with affordable housing strategies in ways that reduce GHG emissions – almost the policy translation of the analytical statement that employment accessibility, measured by a gravity variable, is key. Our conclusion is that there is nothing in our analysis that is at odds with SB 375, and we believe that future refinements of this analysis should be useful for SB 375 implementation.

Yet there is another possible SB 375 path. The popular imagery of land use – transportation links is informed by ideas from Smart Growth, Transit-Oriented Development, and the New Urbanism. Those ideas are typically focused on neighborhood scale design and land use, with an often explicit link to rail transit. The discussion of SB 375 among lay persons in the region often morphs into a transit-oriented, Smart Growth view of the legislation. The evidence in this paper suggests that a more mundane focus on employment accessibility will deliver larger VMT reduction. Certainly neighborhood-level Smart Growth and transit-oriented development can be consistent with regional efforts to improve access to job centers. The difficulty lies not in any inherent tension between local and regional planning, but in the risk that a powerful neighborhood narrative will obscure the need to plan more regionally. SB 375, as written, is a possibility to focus on metropolitan-scale links between transportation access and land use. The results in this paper suggest that maintaining a focus on the metropolitan scale, even while also fostering innovative local or neighborhood planning, will be vital.

References

Badoe, Daniel and Eric J. Miller. 2000. Transportation – Land Use Interaction: Empirical Findings in North America and Their Implications for Modeling. *Transportation Research Part D* 5,4: 235-263.

Bento, Antonio M., Maureen L. Cropper, Ahmed Mushfiq Mobarak, and Katja Vinha. 2005. The Effects of Urban Spatial Structure on Travel Demand in the United States. *The Review of Economics and Statistics* 87,3: 466-478.

Bhat, Chandra R. and J.Y. Guo. 2007. A Comprehensive Analysis of Built Environment Characteristics on Household Residential Choice and Auto Ownership Levels. *Transportation Research B* 41,5: 506-526.

Boarnet, Marlon G. and Sharon Sarmiento. 1998. Can Land Use Policy Really Affect Travel Behavior? *Urban Studies* 35,7: 1155-1169.

Boarnet, Marlon G. and Randall Crane. 2001. *Travel by Design: The Influence of Urban Form on Travel*. New York: Oxford University Press.

Boarnet, M.G. Planning, climate change, and transportation: Thoughts on policy analysis. *Transport. Res. Part A* (2010), doi:10.1016/j.tra.2010.03.001

Boarnet, Marlon G., Kenneth Joh, Wally Siembab, William Fulton, and Mai Thi Nguyen. 2011. Retrofitting the Suburbs to Increase Walking: Evidence from a Land Use – Travel Study**.** *Urban Studies* (forthcoming), doi number 10.1177/0042098010364859.

Brownstone, David. 2008. *Key Relationships Between the Built Environment and VMT*. Draft paper prepared for Transportation Research Board panel on "Relationships Among Development Patterns, Vehicle Miles Traveled, and Energy." October. Available at http://onlinepubs.trb.org/Onlinepubs/sr/sr298brownstone.pdf, accessed August 7, 2010.

Brownstone, David and Thomas Golob. 2009. The Impact of Residential Density on Vehicle Usage and Energy Consumption. *Journal of Urban Economics* 65: 91-98.

Cao, Xinyu, Patricia Mokhtarian, and Susan L. Handy. 2009. Examining the impacts of residential self-selection on travel behavior: A focus on empirical findings. *Transport Reviews* 29 (3), 359–395.

Cervero, R., and K. Kockelman. 1997. Travel demand and the 3Ds: Density, diversity and design. *Transportation Research D* 2,3: 199-219.

Chapman, James and Lawrence Frank. 2004. *Integrating travel behavior and urban form data to address transportation and air quality problems in Atlanta, Georgia* (Research Project No. 9819, Task Order 97-13). Washington, DC: U.S. Department of Transportation.

Crane, Randall. 2000. The Influence of Urban Form on Travel: An Interpretive Review. *Journal of Planning Literature* 15,1: 3-23.

Del Valle, David Heres and Deb Niemeier. 2010. CO_2 Emissions: Are Land-Use Changes Enough for California to Reduce VMT? Specification of a two-part model with instrumental variables. Working Paper, University of California, Davis.

Ewing, Reid and Robert Cervero. 2001. Travel and the Built Environment: A Synthesis. *Transportation Research Record* number 1780: 87-114.

Ewing, Reid and Robert Cervero. 2010. Travel and the Built Environment: A Meta-Analysis. *Journal of the American Planning Association* 76,3: 265 – 294.

Fang, Hao Audrey. 2008. A discrete–continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B* 42: 736–758.

Frank, Lawrence D., and Peter Engelke. 2005. Multiple impacts of the built environment on public health: Walkable places and the exposure to air pollution. *International Regional Science Review*, *28*(2), 193-216.

Funderburg, Richard and Marlon G. Boarnet. 2008. Agglomeration Potential: The Spatial Scale of Industry Linkages in the Southern California Economy. *Growth and Change* 39,1: 24-57.

Giuliano, Genevieve, Ajay Agarwal, and Christian Redfearn. 2008. Metropolitan Spatial Trends in Employment and Housing: Literature Review. Paper prepared for the Committee on the Relationships Among Development Patterns, Vehicle Miles Traveled, and Energy Consumption, Transportation Research Board and the Division on Engineering and Physical Sciences. Available at http://onlinepubs.trb.org/Onlinepubs/sr/sr298giuliano.pdf, accessed August 7, 2010.

Handy, Susan L. 1993. Regional versus Local Accessibility: Implications for Nonwork Travel. *Transportation Research Record*, No. 1400: 58-66.

Handy, Susan L. 2005. Smart Growth and the Transportation - Land use Connection: What Does the Research Tell Us? *International Regional Science Review* 28,2: 146-167.

Johnston, Jack and John DiNardo. 1997. *Econometric Methods*. New York: McGraw-Hill.

Kockelman, Kara M. 1997. Travel behavior as a function of accessibility, land use mixing, and land use balance: Evidence from the San Francisco Bay Area. *Transportation Research Record*, *1607*, 116-125.

Levine, Jonathan. 2006. *Zoned Out: Regulation, Markets, and Choices in Transportation and Metropolitan Land Use*. Washington, D.C.: Resources for the Future.

Mitchell, Robert B. and Chester Rapkin. 1954. Urban Traffic: A Function of Land Use. New York: Columbia University Press.

Moore, Adrian T., Samuel R. Staley, and Robert W. Poole. 2010. The Role of VMT Reduction in Meeting Climate Change Policy Goals. *Transportation Research A*, (forthcoming), doi:10.1016/j.tra.2010.03.012.

National Research Council (NRC), Committee on Relationships Among Development Patterns, Vehicle Miles Traveled, and Energy Consumption. 2009. *Driving and the Built Environment: The Effects of Compact Development on Motorized Travel, Energy Use, and CO2 Emissions*. Washington, D.C.: National Academies Press.

Nustats, 2003. Post Census Regional Household Travel Survey, Data User's Manual. Los Angeles, CA: Southern California Association of Governments.

Pindyck, Robert S. and Daniel L. Rubinfeld. 1981. *Econometric Models and Economic Forecasts*. New York: McGraw-Hill.

Pushkar, A. O., Bruce Hollingworth, and E. J. Miller. 2000. *A multivariate regression model for estimating greenhouse gas emissions from alternative neighborhood designs*. Paper presented at the 79th annual meeting of the Transportation Research Board, Washington, DC.

Pushkarev, Borris and Jeffrey M. Zupan. 1977. *Public Transportation and Land Use Policy*. Bloomington: Indiana University Press.

Redfearn, Christian L. 2007. The Topography of Metropolitan Employment: Identifying Centers of Employment in a Polycentric Urban Area. *Journal of Urban Economics* 61,3: 519-541.

Southern California Association of Governments. 2004. Southern California Compass: Growth Vision Report. June. Available at http://www.compassblueprint.org/files/scag-growthvision2004.pdf, access August 7, 2010.

Tal, Gil, Susan Handy, and Marlon G. Boarnet. Draft Policy Brief on the Impacts of Regional Accessibility Based on a Review of the Empirical Literature. Sacramento: California Air Resources Board. Available at http://arb.ca.gov/cc/sb375/policies/policies.htm.

The Economist. 2010. Portland and "Elite Cities: The New Model. April 15.

U.S. Environmental Protection Agency, 2010. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990 – 2008. Report EPA 430-R-10-006. April. Washington, D.C.: U.S. Environmental Protection Agency. Available at http://www.epa.gov/climatechange/emissions/usinventoryreport.html, accessed May, 2010.

Winkelman, Steve and Allison Bishins. 2010. Planning for Economic and Environmental Resilience. *Transportation Research A*, (forthcoming), doi:10.1016/j.tra.2010.03.011.

Appendix A: Sociodemographic and Land Use Variables, descriptive statistics

Sociodemographic Variable[a]	**Mean**	**Std. Dev.**	**Min**	**Max**
# persons in HH	2.267	1.296	1	9
# bicycles in HH	0.922	1.317	0	10
Homeowner	0.608	0.488	0	1
English primary language at home	0.917	0.276	0	1
Spanish primary language at home	0.072	0.258	0	1
HH income level				
$10,000 - $25,000	0.157	0.364	0	1
$25,000 - $35,000	0.129	0.335	0	1
$35,000 - $50,000	0.148	0.355	0	1
$50,000 - $75,000	0.223	0.416	0	1
$75,000 - $100,000	0.130	0.336	0	1
$100,000 - $150,000	0.098	0.297	0	1
Over $150,000	0.054	0.226	0	1
# workers in HH	1.158	0.842	0	6
# students in HH	0.632	0.977	0	7
# licensed drivers minus vehicles	-0.152	0.759	-7	4
# persons under 16 in the HH	0.455	0.902	0	7
# persons with disabilities in the HH	0.104	0.337	0	3
Age of main respondent	46.543	16.720	0	95
Age of main respondent squared	2445.753	1692.631	0	9025
Main respondent male	0.471	0.499	0	1
Main respondent employed	0.648	0.477	0	1
Main respondent's race/ethnicity				
White	0.656	0.475	0	1
Hispanic	0.197	0.398	0	1
African-American	0.072	0.258	0	1
Asian/Pacific Islander	0.046	0.209	0	1
Main respondent's education level				
High school graduate	0.254	0.435	0	1
2 years of college/Associates Degree	0.245	0.430	0	1
4 years of college/Bachelors degree	0.255	0.436	0	1
Post-graduate	0.154	0.361	0	1
Other	0.015	0.121	0	1
Travel diary included a Saturday	0.090	0.286	0	1
Travel diary included a Sunday	0.066	0.249	0	1
HH surveyed in Fall 2001	0.333	0.471	0	1
HH surveyed in Spring 2002	0.464	0.499	0	1

Land Use Variable[a]	**Mean**	**Std. Dev.**	**Min**	**Max**
Persons per sq. mi. within 1/4 mile	8527.193	7699.033	0	72953
Employment gravity variable (linear damping)	197.811	95.478	19	769
Commercial land use share within 1/4 mile	8.713	11.926	0	98
Med/high residential share within 1/4 mile	12.892	17.294	0	98
Distance to nearest freeway onramp (meters)	3055.287	6101.653	1	86014
# street intersections within 1/4 mile	103.292	41.842	1	433
Fraction of those intersections that are 4-way	0.246	0.179	0	1
Distance to L.A. City Hall (proxy for CBD)	55046.910	45176.290	276	357987
HH within 1/4 mile of rail station	0.009	0.096	0	1
HH within 1/2 mile of rail station	0.038	0.191	0	1
HH within 1/2 mile of a bus stop	0.868	0.339	0	1
HH within 1/2 mile of an express bus stop	0.272	0.445	0	1
HH within 1/2 mile of a rapid bus stop	0.143	0.350	0	1
HH in Compass 2% Area	0.311	0.463	0	1

[a] Statistics calculated for the 12,029 households used in the regressions.

Appendix B: Regression Results, Equation (1), Tobit regression fit on full sample

	Coefficient	Std. Error	t-stat
Persons per sq. mi. within 1/4 mile	0.000032	0.000174	0.18
Employment gravity variable (linear damping)	-0.114224	0.020572	-5.55
# persons in HH	14.345970	1.653724	8.67
# bicycles in HH	0.054772	0.673829	0.08
Homeowner	5.430469	2.102542	2.58
English primary language at home	0.652045	7.921144	0.08
Spanish primary language at home	-9.739624	8.803417	-1.11
HH income level			
$10,000 - $25,000	17.722770	4.237788	4.18
$25,000 - $35,000	28.006450	4.387237	6.38
$35,000 - $50,000	34.952750	4.398933	7.95
$50,000 - $75,000	42.640510	4.362773	9.77
$75,000 - $100,000	42.289310	4.749126	8.90
$100,000 - $150,000	43.366690	5.003719	8.67
Over $150,000	47.719730	5.573739	8.56
# workers in HH	7.222494	1.937953	3.73
# students in HH	2.644149	1.265575	2.09
# licensed drivers minus vehicles	-4.148310	1.113866	-3.72
# persons under 16 in the HH	-13.864560	1.988888	-6.97
# persons with disabilities in the HH	-14.307080	2.641084	-5.42
Age of main respondent	1.069729	0.296083	3.61
Age of main respondent squared	-0.013879	0.002978	-4.66
Main respondent male (= 1 if yes)	5.329279	1.667394	3.20
Main respondent employed (= 1 if yes)	4.757199	2.822731	1.69
Main respondent's race/ethnicity			
White	10.968380	4.844847	2.26
Hispanic	4.783170	5.252993	0.91
African-American	-1.111039	5.679793	-0.20
Asian/Pacific Islander	13.859130	6.108430	2.27
Main respondent's education level			
High school graduate	19.424300	3.770977	5.15
2 years of college/Associates Degree	26.261660	3.928971	6.68
4 years of college/Bachelors degree	28.757260	4.039997	7.12
Post-graduate	31.457440	4.297053	7.32
Other	33.863910	7.404212	4.57
Commercial land use share within 1/4 mile	-0.107810	0.082520	-1.31
Med/high residential share within 1/4 mile	-0.065333	0.060908	-1.07
Distance to nearest freeway onramp (meters)	0.000470	0.000156	3.02
# street intersections within 1/4 mile	-0.030230	0.021469	-1.41
Fraction of those intersections that are 4-way	-17.428490	5.738423	-3.04
Distance to L.A. City Hall (proxy for CBD)	-0.000176	0.000039	-4.52
HH within 1/4 mile of rail station	-3.094354	10.285560	-0.30

	Coefficient	Std. Error	t-stat
HH within 1/2 mile of rail station	-8.554579	5.377510	-1.59
HH within 1/2 mile of a bus stop	0.134049	2.755370	0.05
HH within 1/2 mile of an express bus stop	-2.131127	2.062802	-1.03
HH within 1/2 mile of a rapid bus stop	0.610553	2.981448	0.20
HH in Compass 2% Area	-4.087729	2.159362	-1.89
Travel diary included a Saturday	39.081790	2.876388	13.59
Travel diary included a Sunday	30.844950	3.295887	9.36
HH surveyed in Fall 2001	1.123705	2.296653	0.49
HH surveyed in Spring 2002	10.338650	2.235780	4.62
Constant	-58.073450	13.775350	-4.22
/sigma	84.8923	0.6164098	
Pseudo R2 = 0.0178			
n = 12029	2207 Left-censored observations		