Note to Readers: This paper's pagination is "off" (relative to the original and to the Table of Contents shown here), and its graphics are disfigured. These problems are due to the document's original creation in WordPerfect and later retrieval into Word 97. Please see *Transportation Research Record 1607* for a short version of this paper, and please see the original thesis (at the U.C. Berkeley Libraries) for correct pagination when quoting this document directly. Sorry for the inconvenience. **Thanks very much!**

TRAVEL BEHAVIOR AS A FUNCTION OF ACCESSIBILITY, LAND USE MIXING, AND LAND USE BALANCE: Evidence from the San Francisco Bay Area

by

Kara Maria Kockelman

B.S. (University of California, Berkeley) 1991

A thesis submitted in partial satisfaction of the
requirements for the degree of
Master of City Planning

in City and Regional Planning

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA, BERKELEY

Committee in charge:

Professor Robert Cervero, Chair
Professor Elizabeth Deakin
Professor William Garrison

1996

The thesis of Kara Maria Kockelman is approved:

Chair Date

Date

Date

University of California, Berkeley

1996

Abstract

TRAVEL BEHAVIOR AS A FUNCTION OF ACCESSIBILITY,
LAND USE MIXING, AND LAND USE BALANCE:
Evidence from the San Francisco Bay Area

by

Kara Maria Kockelman

Master of City Planning

in City and Regional Planning

University of California, Berkeley

Professor Robert Cervero, Chair

By incorporating characteristics of the built environment into models of travel behavior, much can be said about household travel distances, automobile ownership, and mode choice. This research investigates the relative significance of a variety of measures of urban form, both at trip-makers' home neighborhoods and at trip ends. The travel data come from the 1990 San Francisco Bay Area Travel Surveys, and the land-use data are largely constructed from hectare-level descriptions provided by the Association of Bay Area Governments. After controlling for demographic characteristics, the measures of accessibility, land use mixing, and land use balance proved to be highly statistically significant and influential in their impact on household vehicle miles traveled (VMT), automobile ownership, and mode choice. In contrast, under the majority of models (with the important exception of the vehicle-ownership models), density's impact was

negligible, after controlling for accessibility. In many cases, balance, mix, and accessibility were found to be more relevant (as measured by elasticities) than several household and traveler characteristics that often form a basis for travel behavior prediction. Moreover, the apparent influence that these variables, particularly accessibility, have on travel behavior is dramatic.

If a societal objective is reduced automobile use and dependence, while maintaining or improving general accessibility levels, these results lend empirical support to the promotion of a variety of land-use policies, such as regional growth containment, the raising and/or removal of density/intensity caps, and the establishment of mixed-use and flexible zoning standards throughout urban areas. These results also represent a step forward in the inclusion of measures of urban form in travel behavior forecasting models; thanks to the technology of geographical information systems and the increasing availability of detailed land-use data sets, such measures can be computed for a multitude of zones at relatively low cost.

TABLE OF CONTENTS

ACKNOWLEDGMENTS:

Many of the data sets used in the analyses are the product of work supported by a University of California Transportation Center research grant. Without these monies, the project and my results would have been greatly limited.

I wish to extend my sincere gratitude to Professor Robert Cervero for his guidance and his provision of UCTC research grant monies for much of the data acquisition. Thanks also go to Kangli (Tony) Wu, Alfred Round, Alexander Skabardonis, Charles (Chuck) Purvis, and Gordon Ye for their help in acquiring and compiling several of the data sets.

INTRODUCTION:

There is much interest in the connection between urban form and travel behavior. Neo-traditionalists and others assert the virtues of transit-oriented design (TOD) (*e.g.*, Kelbaugh 1989, Calthorpe Associates 1992, Bookout 1992), yet much work remains to be done in gauging the influence of TOD features. Kulash (1990) and McNally and Ryan (1993) have looked at the *theoretical* travel reductions associated with a grid network (and a mixed land-use pattern),[1] while others have attempted paired-sample analyses of neighborhoods (*e.g.*, Handy 1992 and Friedman *et al.* 1994). In investigations of the connection between travel behavior and the built environment, few have been able to move beyond simplistic measures of neighborhood structure, such as density (*e.g.*, Newman and Kenworthy 1989, Pushkarev and Zupan 1977, Holtzclaw 1990 & 1994, and Cheslow and Neels 1980), or subjective measures, such as the "Pedestrian Environment Factor" (PBQD 1993) and respondents' perceptions of neighborhood quality (Kitamura *et al.* 1995).

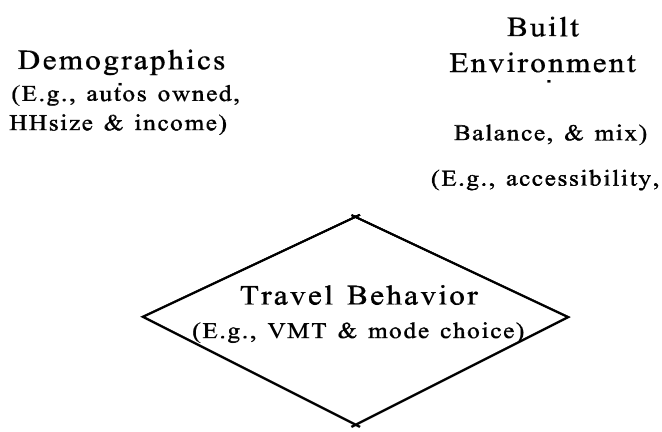


FIGURE 1. Generalized Diagram Of Travel Behavior Influences

The research pursued here intends to operationalize what are generally complex descriptors of land patterns such as "accessibility," land use "balance," "diversity," "mix," and "integration." It brings the power, speed, and precision of geographical information databases and software into formal urban analysis, instead of relying on subjective and/or simplistic descriptors of the built environment. Finally, this research attempts to elucidate and quantify the likely effects of land-use patterns, such as mixing and balance, as they relate to travel decisions. The models are intended to explore the degree of association between several dimensions of land use and travel behaviors, after controlling for socio-economic factors. Figure 1 provides a simplistic rendering of this structure, while Figure 3, preceding the thesis' conclusions, adds several levels of complexity to the relational diagram.

This thesis begins with a detailed discussion of the measures of the built environment that were used in the research, along with related work by other researchers. Then, the results of the analysis for vehicle miles traveled (VMT) per household[2], auto ownership, and mode choice are presented, along with conclusions and suggestions of directions for future research.

MEASURES OF THE BUILT ENVIRONMENT AND RELATED RESEARCH:

The association between aspects of the built environment at the *employment site* and workers' commute choices has been studied by others (*e.g.*, Cervero 1989, Pivo 1993, Cambridge Systematics 1994, and Hopper 1989). The general consensus is that features such as pedestrian-oriented design, high levels of transit service, and the presence of certain amenities (such as banks and shops) are associated with non-single-occupant-vehicle (SOV) mode choice.[3]

In looking at the *place of residence* (and/or trip origin and destination zones), the conclusions drawn are very similar (*e.g.*, Kitamura *et al.* 1995, Frank 1995). Yet, weaknesses in the neighborhood descriptors used (such as density alone) and/or technique (such as simple correlations and/or few neighborhood observations) are pervasive. For this and other reasons, questions regarding the connection between travel behavior and the local built environment have remained difficult to answer. The choice of measures of the built environment is critical in helping answer many, if not all, of these questions. The measures constructed for use in this research are intended to quantify the essence of the following characteristics of relatively local urban form: intensity, balance, and integration of land uses. These are modeled here under the following titles: density and accessibility, entropy, and the "dissimilarity index."

The expectation is that increases in land-use intensity, balance, and integration reduce automobile use via one or more causal channels. For example, an increase in opportunity sites (*i.e.*, places where one can satisfy typical needs, such as working and shopping) reduces the need for long trips (as well as the roadway and parking space available for the space-intensive automobile). Thus, increased *accessibility* should reduce travel distances while making alternatives to the automobile more viable. In a similar vein, local land-use *balance* brings different, yet vital, uses into relative proximity and is therefore expected to shorten trip distances. Finally, land-use *mix* details the integration of these distinct uses even further, for a finer grain of proximity -- and a further shortening of trip distances.

Thanks to the availability of regional data sets and the power of geographical information systems, these various land-use variables were computed for over 1,000 traffic analysis zones (in the case of accessibility and developed-area densities) and over 1,200 Census tracts (in the cases of mix and balance). Their construction is discussed here now.

Density:

The density of population or employment is a common measure of the built environment in literature exploring the connection between urban form and travel behavior (*e.g.*, Pushkarev and Zupan 1977, Cheslow and Neels 1980, Frank 1994, Cambridge Systematics 1994, Kockelman 1995). Its statistical significance in many models may be almost entirely due to its strengths as a *proxy* variable for many difficult-to-observe variables that affect travel behavior[4]. For example, high densities are often associated with high levels of access to opportunity sites, because the high land costs of accessible locations are an incentive for intense development. With a high degree of access to likely destinations, one expects shorter trips. Density is also associated with higher parking costs (via higher land costs), as well as limited parking and, in many cases, congestion. Finally, density is often thought a proxy for level of transit service (*e.g.*, Cambridge Systematics 1994), since economies of density in service provision are present and because the supply of transit tends to follow demand.

Largely due to its simplicity of measurement, density is a very popular measure of the built environment. However, too often it is the only measure used (*e.g.*, Holtzclaw

1990, Levinson and Wynn 1963), and researchers may then purport it to be a "silver bullet" for a myriad of transportation problems. In this study, density was tested as a dependent variable, while controlling for "accessibility." *Developed*-area population and jobs densities were used to avoid bias against ex-urban census tracts where development is often found only on one side or the edge of the tract.[5]

Accessibility:

As mentioned, density is very often a proxy for accessibility because higher land prices encourage more intensive development. As constructed here, accessibility is more descriptive of opportunity intensities than is density because it can be regional in scope and not confined to a local area (such as a census tract). The index used here is constructed from a very popular functional form for quantifying accessibility -- that of the

$$Accessibility_i = \sum_j \frac{A_j}{f(t_{ij})},$$

$$\text{where } A_j = \text{Attractiveness of Zone } j \text{ and } t_{ij} = \text{Travel Time from Zone } i \text{ to } j.$$

gravity model, as shown in Equation 1.[6]

There is an extensive body of literature devoted to the importance and measurement of accessibility; most suggested functional forms are directly comparable to that used here[7]. Discussions centering on the concept and measurement of accessibility can be found in Hägerstrand (1970), Burns (1979), Pirie (1981), Allen (1993), Wachs and Kumagai (1972), Mitchell and Town (1977), Wilbanks (1970), Doling (1979), Handy (1993), and Hansen (1959). While the literature is considerable, the author knows of few studies (*e.g.*, Handy 1992b & 1993 and Levinson 1995), where accessibility has been incorporated into exploratory models of travel behavior.

The basic functional form for accessibility used here assumes a direct proportionality with the number of opportunities (although one could certainly argue that diminishing marginal returns to opportunity exist) and an inversive relation relative to the cost of accessing those opportunities (with "cost" proxied by travel time). The specific index was constructed in a number of ways, with the majority based on an exponential time function incorporating coefficients estimated by Levinson and Kumar (1995). The form depended on trip type (work versus non-work, for example) and mode used (SOV versus walk trips, for example). Total jobs as well as sales and service jobs per traffic analysis zone (TAZ) were used as measures of attractiveness.[8]

While a location's regional accessibility can be expected to be highly relevant to travel behaviors (*e.g.*, VMT per household), there is not much that *local* planners can do to change this. Moreover, *local* accessibility is difficult to measure since census tracts and TAZs can be relatively large and data on jobs are rarely available at a more detailed geographic level. However, the jobs-density variable discussed and evaluated in this research can be expected to provide some measure of local accessibility. And, the indices of entropy and dissimilarity are useful for quantifying certain dimensions of local access to different land use types.

Entropy Index (Land Use Balance):

Most transportation researchers are familiar with the concept of "jobs-housing balance," which is generally measured as deviation from a larger region's average balance (*e.g.*, 1.5 jobs/household).[9] The concept of a balance of *land uses* is more difficult to define. The index used here is based on a measure of "entropy," which was originally defined for the energy state of a system in the Second Law of Thermodynamics, proven by Ludwig Boltzmann in the 1870s. Entropy is commonly used to quantify the uniformity of gaseous mixtures and has been extended to gauge the uncertainty represented by probability distribution functions (Shannon and Weaver 1949) and, more recently, in the regional-science literature, to index sectoral balance across distinct industries (*e.g.,* Garrison and Paulson 1973, Attaran 1986).

This measure is useful in appraising the uniformity in dispersion of a certain trait across *many* zones (as demonstrated by Miller and Quigley [1990] in gauging the degree of racial segregation across the San Francisco Bay Area's census tracts). In its most effective form for describing *spatial* distribution, the entropy index requires a double summation, over proportions of different characteristics as well as averaged (with zone-size weightings applied) across distinct zones. It is maximized when the characteristics are found in constant proportions across the many zones[10].

The measure of entropy used here and elsewhere in the land-use literature incorporates only a single summation (Equation 2), since this type of research is based on local or zonal (rather than multi-zonal) characteristics[11]. To the author's knowledge, the entropy measure was first used to quantify *land-use* balance by Cervero (1989), in looking at suburban employment centers. Following Cervero's example, Frank and Pivo (1995) have used this measure for Seattle-Tacoma area census tracts. Most recently, citing Frank and Pivo, Messenger and Ewing (1996) have incorporated an "adaptation" of

$$Entropy = \sum_j \frac{(P_j \, x \ln(P_j))}{\ln(J)},$$

where P_j *is the proportion of developed land in the jth use type.*

this variable[12].

As constructed here, the entropy measure is normalized with respect to the natural log of the number of distinct uses considered and thus varies between zero and one (with one signifying "perfect" balance of the uses considered). There is a weakness in the single-summation measure since maximum entropy (or "balance") requires that each use type be represented in the same proportion (1/J, where J is the number of different use types considered). In order to address this aspect of the definitional problem in future research, this measure of balance may be made more regionally realistic by incorporating weightings of $P_j \times \ln(P_j)$ for each "j" [13].

The six (J=6) land-use types considered distinct and used in the computation of this index are the following: residential, commercial, public, offices and research sites, industrial, and parks and recreation. A "non-work" entropy measure was also computed, where office/research and industrial uses were ignored (J=4); this measure was tested primarily in the models of non-work home-based trip-making.

As constructed here, the index considers only proportions of *developed* area and thus avoids any biases for or against zones/neighborhoods that are not fully developed (as well as the problem of incorrect computations when proportions do not sum to one).

Furthermore, to avoid bias against smaller tracts, in which there is relatively little area to allow for a variety of land-use types (*e.g.*, in densely developed San Francisco tracts where there may be as few as seven hectares), and to more adequately represent the concept of "neighborhood," a "mean entropy" was constructed (in addition to the tract-

$$Mean\ Entropy = \sum_k \frac{\sum_j \frac{(P_{jk} x \ln(P_{jk}))}{\ln(J)}}{K},$$

$$\begin{gathered} where\ K = Number\ of\ Actively\ Developed\ Hectares\ in\ Tract \\ and\ P_{jk} = Proportion\ of\ Use\ Type\ j\ within\ a\ Half\text{-}Mile\ Radius \\ of\ Developed\ Area\ Surrounding\ the\ kth\ Hectare. \end{gathered}$$

bounded entropy measure).

As shown in Equation 3, the mean entropy is the average of neighborhood entropies computed for all developed hectares within each census tract, where "neighborhood" is defined to include all developed area within one-half mile of each, relevant, active hectare. This measure proved to be far more useful for estimation of travel behavior than the tract-bounded entropy measure (as will be discussed in the section on results).

Dissimilarity Index (Mix):

While entropy helps quantify the degree of balance across distinct land uses, the degree to which these land uses come into contact with one another is also expected to be of importance (*e.g.*, Procos 1976, Lynch 1981). For example, if all shopping is housed in a major shopping center and all jobs are to be found in an industrial park, one may expect few walk trips to and from these uses by residents of relatively nearby homes -- particularly if large spaces surrounding the centers are consumed by parking. Everything else constant (*e.g.*, population and square footage of retail space), one expects the average distance between origins (such as homes) and destinations (such as work places and stores) to be longer when uses are segregated, even if they are in relative balance.[14] Thus, a measure of land-use "mix" based on the hectare-level land-use data set was sought. After many discussions and the consideration of several ideas, a relatively simple measure, offered by Robert Cervero, was accepted and computed. Cervero termed this the "dissimilarity index" since it is based on "points" awarded to each actively developed hectare based on the dissimilarity of its land use from those of the eight adjacent hectares.

$$\begin{gathered}
\textit{Dissimilarity Index} = \textit{Mix Index} = \sum_k \frac{1}{K} \sum_i^8 \frac{X_{ik}}{8}, \\
\textit{where } K = \textit{Number of Actively Developed Hectares in Tract and} \\
\textit{and } X_{ik} = 1 \textit{ if Central Active H'ectares Use Type differs from} \\
\textit{that of a Neighboring Hectare } (X_{ik} = 0 \textit{ otherwise}).
\end{gathered}$$

(Equation 4).

The average of these point accumulations across all active hectares in a tract is the dissimilarity or mix index for the tract. (Equation 4 & Figure 2) Since one-eighth of a point is awarded for each of the adjacent hectares whose active use differs from that of the center hectare, the final average varies between zero and one.

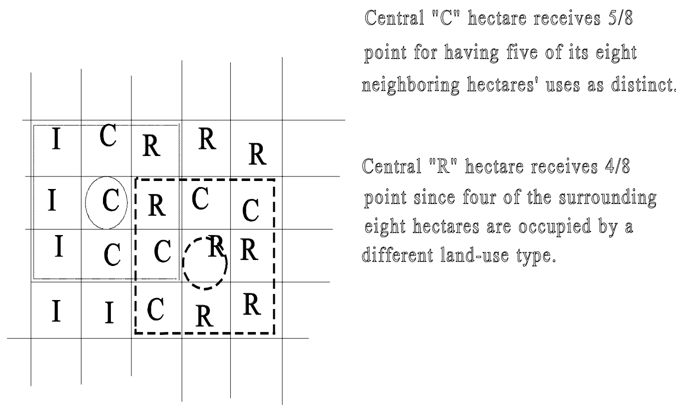


FIGURE 2. Example of Dissimilarity Index Computation Method

Two forms of this index were computed: the first, which relies on only four distinct land-use types (commercial [including industrial and office uses], residential, educational, and outdoor recreational), is termed the "general mix" index; the second index tested depends on eleven distinct land-use types[15] and is termed the "detailed mix" index.

The mix index can also be thought of as a local use dispersion index, since it values diversity over clusters of homogenous use. At the regional level, experimental model results indicate that dispersion of relatively intensively developed activity clusters (*i.e.*, polycentric form) reduces trip lengths; however, this lack of monocentricity in regional form does act to reduce economies of density in transit use and thus the level of service of transit systems and their ridership. This result was the consensus of many works reviewed by Handy (1992a), such as Hamilton (1982) and Schneider (1981). Moreover, experimental work by the author suggests that at a reasonably *local* level, activity dispersion reduces total travel distance of the population slightly while acting to rather dramatically increase the percentage of short trips which could then lend themselves more easily to walk trips; however, the effect on transit patronage may be negative. (A description of this research and its outcomes can be found in the Appendix.)

OTHER METHODS FOR MEASURING LOCAL LAND-USE PATTERNS:

While the built-environment variables constructed and incorporated into this research attempt to measure three distinct dimensions of local land-use patterns (*i.e.*, intensity, balance, and mixture of uses), other measures have been used elsewhere. For example, as a measure of accessibility, there is the "connectivity quotient" which is common in regional science (Bendavid-Val 1991). This index is essentially a normalized average travel distance to other sites/cities, and its summation can be weighted by destination size to better account for differences in destination "attractiveness." There also exists what regional scientists refer to as the "distribution quotient" (Bendavid-Val 1991) which is just a normalized density or a ratio of any two zonal traits (*e.g.*, jobs to housing units).

As a measure of land use mix, Tomanizis' (1979) calculated a "land use variability" index along several bus routes, wherein the number of frontage land-use changes was divided by the length of the route. The construction of such a variable across the over-1000 neighborhood zones modeled here would require a spatial wedding of the ABAG hectares and all streets in the neighborhood/census tract, but should be feasible.

As a possible measure of land-use balance, a "Gini coefficient," based on a Lorenz curve, is commonly used to describe the distribution of regional characteristics (such as income classes by percentage of population [Rodgers 1957] or jobs by industry type)[16].

For a "perfect" score of zero here, however, this measure would require that land uses be present in the same amount, which, as with the entropy index used here, is probably not a regionally practical requirement and depends greatly on the definitional division of use types.

Another measure of balance is the "index of concentration" (Equation 4)

$$\begin{gathered} \textit{Index of Concentration} = 100 - \sum_j \left| \%\,DA_j - \%\,Attribute_j \right|, \\ \textit{where \%DA and \%Attribute are the percentages of developed} \\ \textit{area and a single attribute type within each sub-zone } j. \end{gathered}$$

computed across many zones, typically for an entire region (Bendavid-Val 1991).

As indicated in Equation 4, the index of concentration requires a division of a single zone into several sub-zones, which, given the available data sets and for reasons previously discussed, is not very practical at the neighborhood level in the San Francisco Bay Area.

More recently, as a measure of the *walkability* of a neighborhood, Parsons Brinckerhoff Quade & Douglas (PBQD 1993) constructed a "pedestrian environment factor" which sums a set of rather subjective "scores" for a variety of local factors, such as ease of street crossings and the availability of sidewalks. Unfortunately, subjective measures are difficult to replicate in practice and require significant survey work in the field; for these reasons, this measure is unlikely to be feasible in many situations.

MODELS AND RESULTS:

The models explored here estimate household VMT, auto ownership, and mode choice as a function of demographic and land-use variables. The many variables used are summarized in Table 1. The hypotheses intended for examination are that land-use intensities, balance, and mix are useful for explaining travel behavior and that each of these dimensions of the built environment contributes to reduced automobile use (as evidenced via VMT, auto ownership, and mode choice).

TABLE 1. Summary of Variables Used in Travel-Behavior Models

Dependent Variables:

Total VMT per Household - Sum of Euclidean miles estimate of distance for all household trips made by a "personal vehicle" (*i.e.*, car, truck, van, taxi, or limo), after accounting for occupancy level of vehicle.

Total NWHB VMT per Household - Computed in the same fashion as Total VMT per household, but summed over only those trips that were non-work home-based by households that made at least one such trip.

Auto Ownership - Autos plus trucks owned by household, divided by the number of members in the household that are five years of age or older.

Personal-Vehicle Mode Choice - Binary variable with a level of 1 signifying that trip was made by a personal vehicle (*i.e.*, car, truck, van, taxi, or limo). Only trips by "adults" (*i.e.*, persons over 19 years of age) were considered.

Walk/Bike Mode Choice - Binary variable with a level of 1 signifying that trip was made by walking or by bicycle. Only trips by "adults" (*i.e.*, persons over 19 years of age) were considered.

Independent Variables:

Household Size - Number of household members five years of age or older.

(Household Size)$^{-1}$ - Inverse of total number of household members (all ages included).

Auto Ownership - As defined under the Dependent Variable sector (above).

Income per Member - Household income (as estimated from one of fifteen surveyed income brackets, in 1990$) divided by household size (as defined above).

Accessibility - Accessibility to jobs, based on the gravity model, as described in the body of this thesis. Several forms of this were constructed, such as an index that considered only jobs within 30 minutes by car and another which incorporated walk times within a 30-minute radius and relied only on sales and service jobs.

General Mix - Dissimilarity Index (as described previously) based on four distinct land-use types.

Tract-Confined Entropy - Balance across six distinct land-use categories using the proportions of each category found within each census tract's boundary.

Mean Entropy - Average of entropies computed for all one-half mile radius neighborhoods surrounding each of a census tract's actively developed hectares.

Mean Non-Work Entropy - Average of non-work entropies (computed for all one-half mile radius neighborhoods surrounding each of a census tract's actively developed hectares), where non-work entropy depends only on four distinct use types (*i.e.*, residential, commercial, public, and parks and recreation).

Population Density - Developed-area population density (per acre) of home traffic analysis zone (TAZ) (in the VMT and auto ownership models) and of trip origin and destination TAZs (in the mode-choice models).

Jobs Density - Developed-area employment density (per acre) of home TAZ (in the VMT and auto ownership models) and of trip origin and destination TAZs (in the mode-choice models).

Trip Distance - Euclidean-miles estimate of trip distance.

Sex - Gender of trip-maker; Male = 1.

Age - Age (in years) of trip-maker.

License - Dummy variable for whether trip-maker holds a driver's license (Yes = 1).

Employed - Dummy variable for whether trip-maker holds a full or part-time job (Yes = 1).

Race - Dummy variable for whether trip-maker is white/Caucasian (Yes = 1).

Professional - Dummy variable for whether trip-maker holds a professional job (Yes = 1).

The VMT models look at VMT from all trips made by a household as well as VMT due only to non-work home-based (NWHB) trip-making by a household on the survey day. The auto ownership models were studied in recognition of the endogeneity of this variable in models of VMT as well as those of mode choice. In other words, one can argue that the a household's level of auto ownership is simultaneously determined along with other forms of travel behavior, such as distance and mode choice (*e.g.*, if a household needs to travel long distances by car, it is likely to acquire and, thus, own more cars). The endogeneity of this variable might mean that its use in a simple ordinary least-squares model for VMT, for example, is picking up some of the influence of more fundamental variables (such as accessibility and density) which act to determine the household's choice of auto ownership[17]. From the significance of the auto-ownership models explored here and their cofactors, one may reject a hypothesis of auto-ownership exogeneity. Finally, the models of mode choice considered here are of personal-vehicle choice[18] and walk/bike choice; both are based on binary dependent variables along with logit model assumptions.

The results presented in this thesis for *all* model types have been selected following a process consisting largely of step-wise variable deletion and addition. While a strict application of this technique is essentially atheoretical, in an exploratory analysis this process allows one to focus on the most statistically significant of variables while minimizing the impacts of multi-collinearity. For example, fairly redundant and competing measures of accessibility (such as sales-and-service-jobs accessibility versus total-jobs accessibility) are dropped, leaving a single, relatively robust estimate of the impact of this built-environment dimension. The results of "base" models, which exclude all measures of the built environment but control for a variety of socio-economic factors[19], are shown here as a contrast to the final models:

TABLE 2. Model Results

VMT PER HOUSEHOLD MODELS:

TOTAL VMT PER HOUSEHOLD:

	Coefficient (SE) Base Model	Coefficient (SE) Final Model
Constant	-8.708 (1.243)**	+7.076 (1.61)**
Household Size	+8.842 (0.306)**	+7.890 (.307)**
Auto Ownership	+13.86 (0.674)**	+10.80 (.697)**
Income per Member	+1.574e-4 (1.96e-5)**	+1.646e-4 (1.96e-5)**
Accessibility (all jobs, 30 min. radius)	--	-7.299e-5 (7.96e-6)**
General Mix	--	-17.91 (5.79)**
Mean Entropy	--	-5.805 (3.19)*
R^2	0.1131	0.1452
n = 8,050		

NON-WORK HOME-BASED VMT PER HOUSEHOLD:

	Coefficient (SE) Base Model	Coefficient (SE) Final Model
Constant	+1.038 (0.755)	+8.172 (.979)**
Household Size	+2.450 (0.177)**	+2.020 (.178)**
Auto Ownership	+4.293 (0.425)**	+3.023 (.436)**
Income per Member	+3.087e-5 (1.23e-5)**	+3.160e-5 (1.23e-5)**
Accessibility (all jobs., 30 min. radius)	--	-2.538e-5 (4.00e-6)**
General Mix	--	-9.680 (3.477)**
Mean Non-Work Entropy	--	-4.911 (1.712)**
R^2	0.0366	0.0567
n=6,311		

*Statistically significant at the 10% level.
**Statistically significant at the 1% level.

AUTO OWNERSHIP MODELS:

AUTO OWNERSHIP

	Coefficient (SE)	
	Base Model	**Final Model**
Constant	+1.016 (0.017)**	+1.274 (.022)**
Household Size	-0.106 (0.005)**	-0.117 (.005)**
Income per Member	+6.003e-6 (3.15e-7)**	+5.490e-6 (3.07e-7)**
Accessibility (all jobs)	--	-9.142e-7 (1.22e-7)**
General Mix	--	-0.110 (0.085)
Tract-Confined Entropy	--	-9.160e-2 (3.77e-2)*
Population Density	--	-6.370e-3 (4.54e-4)**
R^2	0.1482	0.2185

n=8,050

*Statistically significant at the 10% level.

**Statistically significant at the 1% level.

MODE CHOICE MODELS (for adult trip-makers, all trip types):

PERSONAL VEHICLE CHOICE

	Coefficient (SE)	
	Base Model	**Final Model**
Constant	-0.900 (.0800)**	+0.0376 (.0900)
Trip Distance	+0.3553 (.0081)**	+0.32696 (.0081)**
Sex	-0.172 (.0303)**	-0.0851 (.0323)**
Age	+8.781e-3 (1.08e-3)**	+7.019e-3 (1.13e-3)**
License	+1.353 (.0571)**	+1.349 (.0589)**
Race	+0.1057 (.0365)**	-0.1209 (.0397)**
Employed	+0.1692 (.0400)**	+0.3453 (.0419)**
Professional	-0.2313 (.0360)**	-0.1963 (.0386)**
Auto Ownership	+1.294 (.0423)**	+0.9666 (.0441)**
Income per Member	-3.251e-6 (9.25e-7)**	+ 9.555e-7 (1.02e-6)
(Household Size)$^{-1}$	-1.413 (.0599)**	-1.158 (.0647)**
Accessibility (all jobs, 30 min. radius)		
Origin Zone	--	-1.977e-6 (6.68e-7)**
Destination Zone	--	-1.604e-6 (6.69e-7)**
Population Density		
Origin	--	-7.145e-3 (1.12e-3)**
Destination	--	-4.684e-3 (1.13e-3)**
Jobs Density		
Origin	--	-3.237e-3 (2.54e-4)**
Destination	--	-3.496e-3 (2.60e-4)**
Pseudo-R^2	0.2021	0.2686

n=52,650

Notes: 88.4% of adults trips were by personal vehicle. Pseudo-R^2 is one minus the ratio of the logs of the residual and null likelihoods (where the residual is fit over the specified model and the null allows only a constant term -- in other words the null assumes no information). The pseudo-R^2 using only HOME-zone accessibility, mix, entropy & density = 22.1%.

WALK/BIKE CHOICE:

	Coefficient (SE) Base Model	Final Model
Constant	+0.318 (.0861)**	-0.5662 (.109)**
Trip Distance	-0.5626 (.0122)**	-0.5247 (.0125)**
Sex	+0.1649 (.0323)**	+0.08926 (.0344)**
Age	-9.808e-3 (1.16e-3)**	-7.892e-3 (1.21e-3)**
License	-0.9542 (.0622)**	-0.9458 (.0647)**
Race	+0.0322 (.0403)	+0.1889 (.0436)**
Employed	-0.1906 (.0428)**	-0.3407 (.0450)**
Professional	+0.2541 (.0386)**	+0.2207 (.0412)**
Auto Ownership	-0.9488 (.04205)**	-0.6719 (.0433)**
Income per Member	+4.012e-6 (9.39e-7)	+1.158e-6 (1.01e-6)
(Household Size)$^{-1}$	+1.314 (.0628)**	+1.064 (.0676)**
Accessibility (walk mode, to sales & service jobs, 30 min. radius)		
Origin Zone	--	+1.834e-4 (1.30e-5)**
Destination Zone	--	+1.798e-4 (1.30e-5)**
Mean Non-Work Entropy		
Origin	--	+0.3027 (.136)*
Destination	--	+0.3378 (.136)*
Pseudo-R^2	0.2191	0.2663

n=52,650
*Statistically significant at the 10% level.
**Statistically significant at the 1% level.
Note: 9.90% of adult trips were by this mode.

Summary of Results:

Clearly the inclusion of built-environment variables has strengthened each of the models. Unexplained variation was reduced substantially, with the explained sum of squares (related proportionally to the R^2's) rising almost 30% in the all-trips VMT model, 55% in the NWHB VMT model, close to 50% in the auto-ownership model, 33% in the PV choice model, and 22% in the walk/bike choice model. However, day-to-day travel can vary dramatically within a single household, so the models of VMT showed high levels of unexplained variation -- particularly those of non-work home-based trip-making (as evidenced by the relatively low coefficients of determination, or R^2's), since many of these trip types do not represent regular/daily trip-making.

From the model results the elasticities of dependent variables with respect to many of the independent variables are estimated (at the medians) to be the following:

TABLE 3. Elasticity Estimates from Model Results

With respect to:	**VMT/HH**	**NWHB VMT/HH**	**AUTO-OWN.**	**PV CHOICE**	**WALK/BIKE CHOICE**
Household Size	+0.82	+0.68	-0.23	--	--
(Household Size)$^{-1}$	--	--	--	-0.067	+0.48
Income per Member	+0.16	+0.10	+0.10	-0.0021	+0.020
Auto Ownership	+0.56	+0.51	--	+0.11	-0.60
Accessibility[*]	-0.31	-0.35	-0.075	-0.036[**]	+0.22[***]
General Mix	-0.10	-0.17	-0.01	--	--
Mean Entropy	-0.10	--	--	--	--
Mean Non-Work Entropy	--	-0.30	--	--	+0.23
Tract-Confined Entropy	--	--	-0.03	--	--
Jobs Density	--	--	--	-0.0019	--
Population Density	--	--	-0.068	-0.013	--

Elasticities are computed at *median* variable values (to avoid upward bias from means, particularly on variables such as accessibility and income). The elasticities of probability estimates in the mode-choice binomial logit models are $\beta x(1-\pi)$; the estimate of probability "π" is taken to be the "no-information" estimate (*i.e.*, 88.4% in the PV models and 9.90% in the walk/bike models -- which represent the sample population's average choice probabilities). Much of the reason for the lower elasticities under the PV model versus the walk/bike model is that it is harder to induce a large percentage change in a number (88.4%) that is much higher to begin with (versus 9.90%).

[*] The accessibility measures used in each distinct model's elasticity results corresponds to that type used in the relevant final model.

[**] Note that without the jobs-density variable in the PV-choice model, accessibility's elasticity would have been estimated as -0.095. However, jobs density, while not influential, is highly statistically significant in this model.

[***] If the usual auto-mode, 30-minute accessibility index had been used in the walk/bike mode choice model, its elasticity estimate would have been +0.61! (But then the model's pseudo-R^2 would have been 0.254 -- instead of 0.266 [and the entropy measure would have lost its statistical significance].)

From the computed elasticities, one can judge the relative influence of the independent variables, and it is clear that accessibility is generally very powerful. However, even after controlling for this variable, balance and mix variables are almost always statistically significant at the 1% level -- in addition to being influential (as evidenced by their elasticities). Plots 1 and 2 illustrate the influence accessibility has on four of the response-variable estimates. Plot 2 also tracks the change in the probability that an adult will choose to walk or bike across three different non-work entropy levels.

All model signs are as expected *a priori* except, perhaps, for income's effect, which is estimated to be negative in the PV mode-choice models and positive (but *not* significant) in the walk/bike models[20]. The sign on the income variable is a consistent finding for these two model types in the San Francisco Bay Area (however, the author does estimate a positive income effect in models of non-work SOV mode choice). For other regions, this result may differ. For example, Phoenix and Los Angeles have developed later in history (when different transportation modes were available) and with fewer geographical constraints, so these may exhibit a different relationship between income and mode choice. Possible explanations for the result presented here are that a premium is placed on home prices in walkable neighborhoods, so that only the relatively wealthy can afford to live there, and/or the relatively wealthy have more leisure time to devote to this mode (which many people value as a form of exercise). Moreover, exclusionary zoning in the more central and transit-served Bay Area locations may be forcing low-income households to the region's periphery, where viable alternatives to personal vehicle use are often non-existent.

Elasticity results are not shown for many of the mode-choice demographic variables, but these can range from +0.10 for trip distance to +0.04 for employment status to -0.01 for sex in the PV mode-choice model; thus, they are often less than the elasticities estimated for the land-use variables.

A summary of other significant findings to be drawn from the various models shown, as well as the exploratory analysis used to reach them, is provided here:

- Accessibility to opportunities is very strongly associated with automobile use; the elasticities of VMT, auto ownership, and walk/bike mode choice with respect to this variable are high -- particularly in relation to other variables common to models of this type. It seems reasonable to expect that this association stems from a direct effect -- since closer opportunities diminish trip distances, thus reducing miles travelled as well as making the automobile less of a necessity -- as well as an indirect effect, since accessibility is generally associated with higher land prices, less convenient parking options, and probably more roadway congestion.
- Accessibility is a far better predictor of VMT and mode choice than is density. And, once regional accessibility is controlled for, local jobs density (which can be thought of as a measure of local accessibility) offers little if any explanatory power in all but the PV-choice models. However, population densities prove highly useful (and should not be neglected!) in the models of auto-ownership.
- Land use balance (entropy) and mix (dissimilarity) do appear to matter, affecting VMT and walk/bike probabilities substantially; but, generally, they are not as influential as accessibility.
- Mean entropy is much more useful in all models studied of VMT per household and mode choice than is the tract-confined entropy measure. The latter was of significant use in the models of auto-ownership (although is was of minor influence, as measured by its elasticity); but it did not come close to statistical significance in most other models, even in isolation. Thus, one may infer that

tract boundaries bias measures such as this and that a less restricted view of a household's "neighborhood" is often more appropriate for land-use balance.

- Auto ownership appears to be more significantly influenced by *local* attributes of the built environment, relying on population density and tract-confined entropy measures for the final model. Local measures are also important in the mode-choice models (*i.e.*, jobs and population density in the PV models and sales-and-service-jobs walk accessibility within 30 minutes for the walk/bike model). One might have expected this result, given that parking convenience and roadway congestion (which are probably associated with densities) and walk environments (which may depend substantially on very local use balance) are all very relevant at a local level, whereas vehicle trip *distances* (and VMT) are influenced at more of a community and regional level.
- The entropy index that considered all types of active land uses performed substantially better than the non-work entropy (which ignored office and industrial uses) in most all models -- with the important exceptions of the NWHB VMT-per-household and walk/bike mode-choice models. This result is not very surprising, given the models' different dependent variables.
- The "detailed" dissimilarity index (*i.e.*, detailed mix) performed poorly, in comparison with the more general index (which considered only four, rather than eleven, land use types as distinctly different). So, the level of *definitional* detail in a land-use dataset is probably not as critical for a measure of neighborhood mix as one might expect *a priori*.[21] However, additional spatial detail, *i.e.*, data on a

finer grid of areal cells, could be highly useful.

- Population density generally was more useful than jobs density, even when accessibility (a type of *regional* jobs density measure) was removed from the models. However, in the PV mode-choice models, jobs density was helpful, even after controlling for jobs accessibility.

Suggestions for Model Improvement:

There are several areas in which the model structure and data sets could be strengthened. Figure 3 illustrates many of the methods suggested here for model improvement, depicting many of the feedbacks and bi-directional causal linkages that make land use-transportation modelling so complex.

In terms of improving the models, it would be useful to set up a system of simultaneous equations, with variables such as trip distance, auto ownership, and transit provision endogenously determined, along with VMT and mode choice. The transit provision information is difficult to obtain at a regional level, but should be helpful. Messenger and Ewing (1996) managed to set-up this kind of model for almost 700 zones in Dade County, Florida[22]. Transit time and cost data (as well as walk, carpool, and bike data, for example) would also be useful in mode-choice models by allowing for a more detailed description of competing modes between distinct O-D pairs.

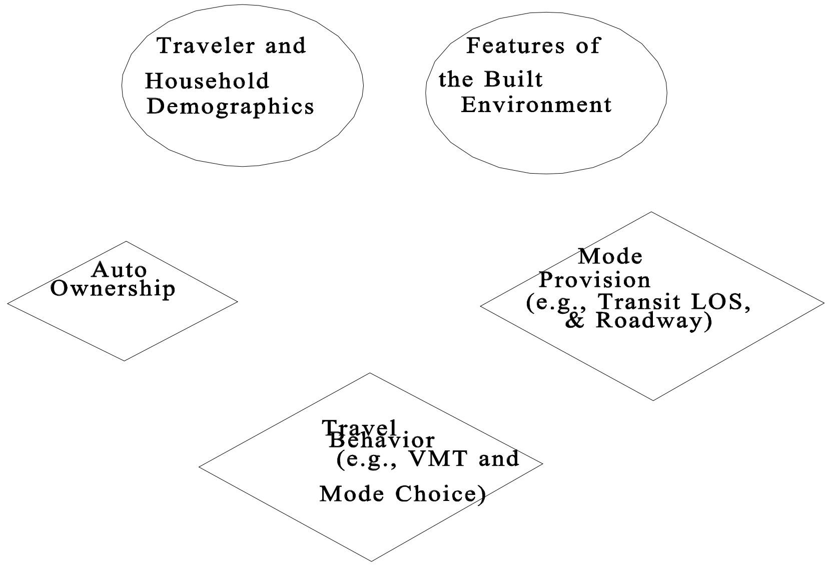


FIGURE 3. – Diagram of Relationships between Travel Decisions and the Built Environment

Furthermore, in any model of the types explored here, there is a fundamental question of causation. While attributes may be associated with the dependent variable and do well in models of prediction, statistical results can mask underlying linkages that are at the heart of the transportation-land use connection. For example, attitudes may be more important than local environment in determining how one travels. And people may be choosing their environments according to how they expect to travel. Thus, environment may not be exogenously determined. Kitamura *et al.* (1995) have taken some steps toward an exploration of this possible relation, which is very similar in spirit to Tiebout's (1964) theories on household location.

In this same vein, there is the question of whether the built-environment determines travel behaviors, or whether the reverse is more valid. In order to better observe the direction of causation, serial models are generally preferable to models relying on simply a single cross-section of data. This is a weakness of the vast majority of empirical land use-transportation models; Handy (1992a) noted this as a "glaring gap" in her review of the relevant literature. However, a very sizable hurdle in data acquisition of this type is that few descriptors of the built environment change within a few years time. Other variables affecting travel behaviors, such as automotive and transit technologies, are likely to change far more rapidly, "drowning out" any built-environment changes that may occur over time. Ideally, paired-sample analyses can overcome this hurdle, but there are hardly any neighborhood pairs where all but one of the attributes are the same.

Finally, the lack of more than a single weekday's data is a troubling aspect in the data. Single-day reporting adds variability, particularly for NWHB trip-making. And weekend trip-making is important for non-work travel. Holtzclaw's (1990 & 1994) use of odometer readings or, at least, use of actual network travel distances for VMT estimation is far preferable to the method used here (*i.e.*, centroid-to-centroid Euclidean distance estimates). (Note, however, that Holtzclaw had access only to very limited and aggregate demographic data.)

CONCLUSIONS:

Measures of the built environment centering on the intensity, balance, and mix of land uses proved to be of substantial use in the models of travel behavior tested here, lending credible support to neo-traditionalists' claims that land-use integration and compact development reduce automobile reliance. Moreover, if a societal objective is the reduction of automobile use (while maintain or adding to accessibility levels), these results provide evidence in favor of several land-use policies. For example, they lend support to policies of mixed and flexible land-use zoning, the removal of zoning development intensity caps (which can be exclusionary of small and low-income households if set too low), and the constraint of urban growth through the setting of regional boundaries. Unfortunately, in regards to this third policy suggestion, few institutions are currently in place to deal with regional land-use issues. Councils of Government do not generally have authority to supersede municipal controls, and few municipalities, if any, are generally willing to relinquish this power. But these results do indicate a need for such institutions.

The results also indicate that one should avoid the use of purely local measures of the built environment in models of travel behavior, given that many may be correlated with more regional variables (such as regional accessibility and population density, whose correlation across the sampled Bay Area households was +0.69) and that tract and other zonal boundaries are rather arbitrarily determined and may not represent neighborhoods well at all. However, local measures of the built environment proved very important in the auto-ownership and walk/bike mode-choice models pursued here; so these variables do appear to affect some very relevant aspects of travel behavior.

Finally, it is evident that regional land-use data sets, coupled with the computing power of geographical information systems, provide a wealth of data to travel modelers at a relatively low cost. Such data should be considered very seriously for inclusion in future modeling efforts, particularly when data requirements are large and preclude on-site in-depth surveys.

REFERENCES:

Attaran, M. (1986). "Industrial Diversity and Economic Performance." **The Annals of Regional Science**, 20: 44-54.

Bendavid-Val, Avrom (1991). **Regional and Local Economic Analysis for Practitioners** (Fourth Edition). Westport, Conn.: Praeger.

Bookout, Lloyd W. (1992). "Neotraditional Town Planning: Cars, Pedestrians, and Transit." **Urban Land**. February 1992: 10-15.

Calthorpe Associates (1992). "Transit-Oriented Development Impacts on Travel Behavior". San Francisco: Calthorpe Associates, Inc.

Cambridge Systematics, Inc. (1994). "The Effects of Land Use and Travel Demand Management Strategies on Commuting Behavior." Prepared for the FHWA, /U.S.DOT, with subcontract assistance from Deakin, Harvey, Scabardonis, Inc. Contract DTFH61-91-C-00085. July 1994.

Cervero, Robert (1989). **America's Suburban Centers: The Land Use-Transportation Link**. Boston: Unwin-Hyman.

Cheslow, Melvin D., and J. Kevin Neels (1980). "The effect of urban development patterns on transportation energy use." Presented at the 59th Annual Meeting of the TRB, Jan. 1980.

Frank, Lawrence D. (1994). **An Analysis of Relationships Between Urban Form (Density, Mix, and Jobs: Housing Balance) and Travel Behavior (Mode Choice, Trip Generation, Trip Length, and Travel Time)**. Washington State DOT & USDOT, July 1994.

Friedman, Bruce, Stephen P. Gordon, & John B. Peers (1994). "Effect of Neotraditional Neighborhood Design on Travel Characteristics." **Transportation Research Record No. 1466**: 63-70.

Garrison, C.B. and A. Paulson (1973). "An Entropy Measure of the Geographic Concentration of Economic Activity." **Economic Geography**, 49: 319-324.

Green, William H. (1993). **Econometric Analysis**. Second Edition. New York: Mcmillan Publishing.

Hamilton, B.W. (1982). "Wasteful Commuting." **Journal of Political Economy**. Volume 90, #5: 1035-1053.

Handy, Susan (1992a). "How Land Use Patterns Affect Travel Patterns: A Bibliography." Council of Planning Librarians Bibliography #279, Chicago, IL.

Handy, Susan (1992b). "Regional Versus Local Accessibility: Neo-Traditional Development and its Implications for Non-work Travel." **Built Environment**, Vol. 18, No.4: 252-267.

Handy, Susan (1993). "Regional Versus Local Accessibility: Implications for Nonwork Travel." **Transportation Research Record No. 1400**. August. National Research Council, TRB. Washington, D.C.: 58-66.

Hansen, Walter (1959). "How Accessibility Shapes Land Use." Journal of the American Institute of Planner, Vol. 25: 73-76.

Holtzclaw, John (1990). "Manhattanization vs. Sprawl: How Urban Density Impacts

Auto Use Comparing Five Bay Area Communities." Paper presented at the Eleventh International Pedestrian Conference.

Holtzclaw, John (1994). "Using Residential Patterns and Transit to Decrease Automobile Dependence and Cost." Paper prepared for California Home Energy Efficiency Rating Systems. San Francisco: Natural Resources Defense Council.

Hopper, Kevin G. (1989). "Travel characteristics at large-scale suburban activity centers." Transportation Research Board, NCHRP #323. Washington, D.C.: National Research Council.

Jacobs, Jane (1961). **The Death and Life of Great American Cities**. New York, Random House (Vintage Book).

Kelbaugh, Doug (Ed.) (1989). **The Pedestrian Pocket Book**. New York: Princeton Architectural Press.

Kitamura, Ryuichi, Patricia L. Mokhtarian, and Laura Laidet (1995). "A Micro-Analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area." Paper presented at the 74th Annual Meeting of the Transportation Research Board. Washington, D.C.

Kockelman, Kara (1995). "Which Matters More in Mode Choice: Density or Income?: The Relative Effects of Population Density and Income on Modal Split in Urban Areas." Compendium of papers presented at the 1995 ITE District 6 Conference in Denver, Colorado.

Krumm, Ronald J. (1980). "Neighborhood Amenities: An Economic Analysis." **Journal of Urban Economics**. Vol. 7: 208-224.

Kulash, Walter (1990). "Traditional Neighborhood Development: Will the Traffic Work?" Prepared for ASCE's "Successful Land Development: Quality and Profits" Conference. March 1990.

Levinson, David (1995). "Location, Relocation, and the Journey to Work." Paper presented at the Western Regional Science Conference (November 1995), San Diego.

Levinson, David and Ajay Kumar (1995). "A Multimodal Trip Distribution Model: Structure and Application." **Transportation Research Record No. 1446**. TRB, NRC. Washington, D.C.

Levinson, Herbert S. and F.H. Wynn (1962). "Some Aspects of Future Transportation in Urban Areas." **Highway Research Board Bulletin No. 326**: 32-36

.Levinson, Herbert S. and F.H. Wynn (1963). "Effects of Density on Urban Transportation Requirements." **Highway Research Record No. 2**: 38-64.

Lynch, Kevin (1981). **Good City Form**. Cambridge, Massachusetts: MIT Press.

McMillan, Velville L., Bradford G. Reid, and David W. Gillen (1980). "An Extension of the Hedonic Approach of Estimating the Value of Quiet." **Land Economics**. Vol. 56, No. 3 (August): 315-328.

McNally, Michael G., and Sherry Ryan (1993). "Comparative Assessment of Travel Characteristics for Neotraditional Design." **Transportation Research Record No. 1400**. TRB, NRC. Washington, D.C.

Messenger, Todd, and Reid Ewing (1996). "Transit-Oriented Development in the

Sunbelt: Get Real (and Empirical)!" Presented at the 75th Annual Meeting of the Transportation Research Board. Washington, D.C.

Miller, Vincent P. & John Quigley (1990). "Segregation by racial and demographic group: Evidence from the San Francisco Bay Area." University of California Transportation Center, Paper #14.

Newman, Peter & Jeffrey Kenworthy (1989). **Cities and Automobile Dependence: An International Sourcebook**. Aldershot, England: Gower Publishing.

PBQD (Parsons Brinckerhoff Quade Douglas, Inc.) (1993). **The Pedestrian Environment - Volume 4A**. Portland, OR: 1000 Friends of Oregon.

Pivo, Gary E. (1988). "The Intrasuburban Pattern of Office Suburbanization: Form and Impact in the San Francisco Bay Area." **The City of the 21st Century**. Madis Pihlak (Ed.)

Procos, Dimitris (1976). **Mixed Land Use: From Revival to Innovation**. Stroudsburg, Pennsylvania: Dowden, Hutchison, & Ross.

Pushkarev, Boris S. & Jeffrey M. Zupan (1977). **Public Transportation and Land Use Policy**. Bloomington, Indiana: Indiana University Press.

Rodenborn, Shirley A., and Charles L. Purvis (1991). "1990 Bay Area Travel Surveys: Data collection and Data Analysis." Presented at the 70th Annual Meeting of the Transportation Research Board. Washington, D.C.

Rodgers, A. (1957). "Some Aspects of Industrial Diversification in the United States." **Economic Geography**. 33:16-30.

Rosenbloom, Sandra, and Elizabeth Burns (1993). "Gender Differences in Commuter Travel in Tucson: Implications for Travel Demand Management Programs." **Transportation Research Record No. 1404.** TRB, NRC. Washington, D.C.

Rosenbloom, Sandra, and Elizabeth Burns (1994). "Why Working Women Drive Alone: Implications for Travel Reduction Programs." **Transportation Research Record No. 1459.** TRB, NRC. Washington, D.C.

Schneider, Jerry B. (1981). "Transit and the Polycentric City: Final Report." Departments of Civil Engineering and Urban Planning, University of Washington. UMTA. Washington, D.C.

Siegel, Paul B., Thomas G. Johnson, and Jeffrey Alwang (1995). "Regional Economic Diversity and Diversification." **Growth and Change**, 26 (Spring 1995): 261-284.

Sosslau, A.B., A.B. Hassam, M.M. Carter, & G.V. Wickstrom (1978). "Quick-Response Urban Travel Estimation Techniques and Transferable Parameters: User's Guide." **NCHRP Report 187**. TRB, NRC. Washington, D.C.

Tauer, L.W. (1992). "Diversification of Production Activities Across Individual States." **Journal of Production Agriculture**, 5: 210-214.

Tiebout, Charles (1964). "A Pure Theory of Local Expenditures." **Journal of Political Economics**. Vol. 64: 416-424.

Tomazinis, Anthony R. (1979). "A Study on Factors Affecting Success of Suburban Mass Transit Lines: Final Report." USDOT, UMTA, Washington, D.C.

Tress, R.C. (1938). "Unemployment and the Diversification of Industry." **The Manchester School**, 3: 140-152.

Voith, Richard (1991). "Transportation, Sorting and House Values." **AREUEA Journal**. Vol. 19, No. 2: 117-137.
Wilson, Alan Geoffrey (1970). **Entropy in Urban and Regional Modelling**. London: Pion.

APPENDIX

SUMMARY STATISTICS OF DATA:

TABLE A-1-1. Summary Statistics of Household Data

HOUSEHOLD DATA:

N = 8,050	Mean	Median
Household Size	2.53	2
#Household Members >5 years of Age	2.32	2
Total # Trips on Survey Day	7.24	6
Total Travel Distance (Euclidean miles)	37.9	26.4
Total Travel Time (hours)	2.32	1.83
Personal-Vehicle Miles Travelled (Euclidean miles)	28.4	19.2
Vehicles per HH Member >5 years of Age	0.919	1
Income per HH Member >5 years of Age ($1990)	$24,600	$18,750
Entropy (confined to tract)	0.325	0.314 (Max.=0.884)
Mean Entropy (over 1/2-mile-radius neigh.)	0.314	0.317 (Max.=0.808)
Non-Work Entropy (confined to tract)	0.372	.371
Mean Non-Work Entropy	0.362	.365
General Mix Index	0.113	0.106 (Max.=0.482)
Detailed Mix Index	0.120	0.113 (Max.=0.485)
Accessibility to All Jobs (exponential form)	86,600	82,400(Max.=288,870)
Access. to All Jobs (expon., 30 min.)	85,960	82,900(Max.=291,360)
Sales & Service Accessibility (exponential)	18,460	17,970 (Max.=62,843)
Population Density of Home Tract (per developed acre)	15.2	10.8 (Max.=232)
Jobs Density of Home Tract (per developed acre)	6.17	2.7 (Max.=689)

Statistics for Households that had made at least one NWHB Trip on the Survey Day:

N = 6,310		
# Non-Work Home-Based Trips	4.20	3
NWHB Travel Distance (Euclidean miles)	17.3	8.94
NWHB Personal-Vehicle Miles Travelled (Euclidean miles)	11.6	5.95

TABLE A-1-2. Summary Statistics of Adults' Individual Trips Data

ADULTS' INDIVIDUAL TRIP DATA:

N = 52,650 (Trips by Persons 20 yrs & older only)	Mean	Median
Trip Distance (Euclidean miles)	5.52	2.77
Trip Time (minutes)	19.8	15.0
% Trips by Personal Vehicle	88.4	--
% Trips by SOV	61.0	--
% Trips that are HBNW	39.9	--
% Trips by Males	48.8	--
% Trips by Licensed Persons	96.0	--
% Trips by White Persons	77.9	--
% Trips by Employed Persons	74.0	--
% Trips by Professional Persons	39.5	--
Age of Tripmaker	42.6	40
Vehicles per HH Member >5 years of Age	0.935	1
Income per HH Member >5 years of Age	$24,216	$18,750
(Household Size)$^{-1}$	0.4652	0.5
Mean Entropy (over 1/2-mile-radius neigh.)	0.345	0.353
Mean Non-Work Entropy	0.389	0.401
General Mix Index	0..119	0.114
Accessibility to All Jobs (exponential form)	94,000	88,000
Access. to All Jobs (expon., 30 min.)	93,800	87,800
Sales & Serv. 30 min. Walk Accessibility	986	660 (Max.=10,734)
Sales & Serv. 10 min. Summed-Jobs Access.	41,231	35,032 (Max.=131,100)
Population Density (per developed acre)	13.6	9.82
Jobs Density (per developed acre)	5.33	2.48

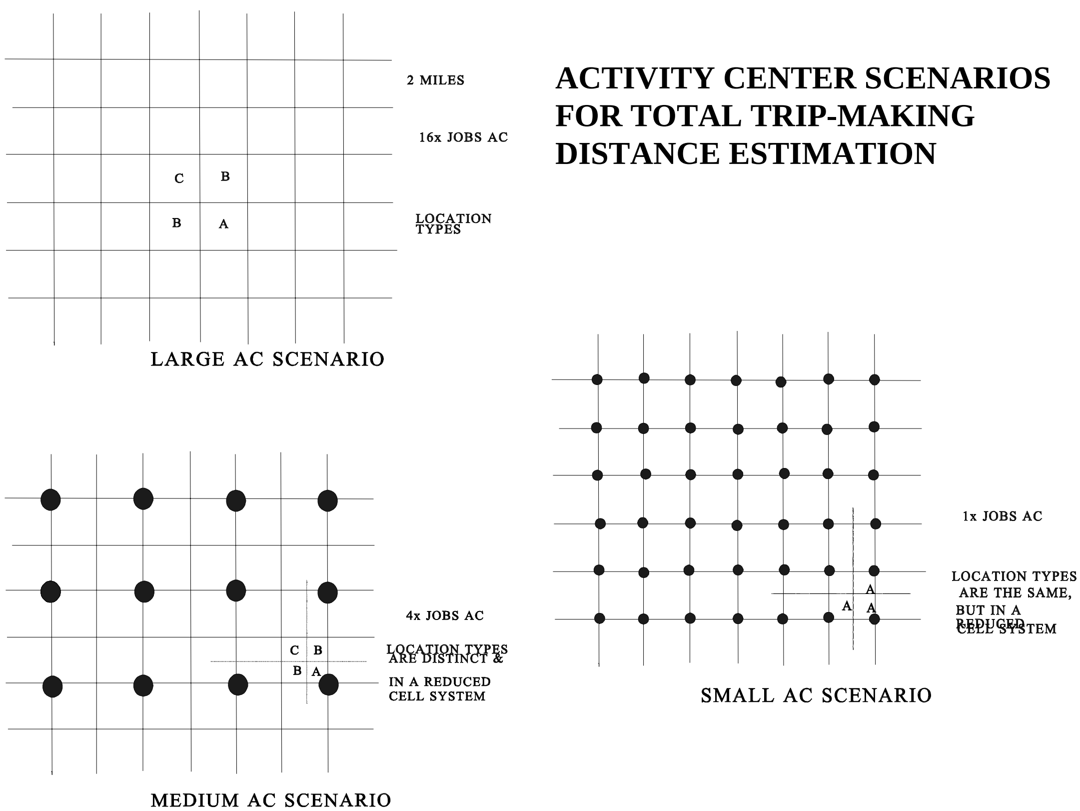
Note: Built-environment statistics for the adults trips dataset are listed with respect to Origin and Destination zones, not trip-makers' home zones.

ACTIVITY CENTER DISPERSION AND THE ESTIMATED IMPACT ON TRAVEL DISTANCES: A HYPOTHETICAL MODEL

Technique for Total Travel Distance Estimation:

Activity centers (ACs) were arranged in three symmetric patterns on a pure grid of roadways. In the "large-AC" scenario, ACs with a jobs or attractiveness weighting of 16 were located at every 16th intersection (or every fourth intersection east to west and every fourth intersection north to south). In the "medium-AC" scenario, ACs with a weighting of four were located at every fourth intersection (or on a two-by-two grid system). In the "small-AC" scenario, ACs of weight one were located at every intersection.

The shown rectilinear grid’s spacing is two miles so that the minimum-path AC spacings are 8 miles, 4 miles, and 2 miles. However, travelers are assumed to travel at "any" rectilinear grid spacing, so the shown grid system is only for the purpose of illustration. A maximum of only three household locations were considered under any one scenario with the "average" household assumed to exist at each cell’s centroid; in the small- and medium-AC scenarios a one-mile spacing has been added to the basic two-mile grid to further divide the residential area. The remaining cells (and their "average" households) are symmetrically located.



ACTIVITY CENTER SCENARIOS FOR TOTAL TRIP-MAKING DISTANCE ESTIMATION

Travel times between home cells and AC sites were estimated under two different assumptions. The first assumed a constant speed, regardless of distance, of 0.4 miles per minute, or 25 miles an hour. The second allowed speed to increase (and thus marginal travel times to decrease) with trip distance and incorporated the following assumption:

Speed [in miles per minute] = 0.25 + (0.015 x Distance [in miles])

So, for a distance of 30 miles, the average trip speed, under this second speed scenario, was assumed to be 42 mph; and, for a 50-mile trip, the speed was assumed to average 60 mph.

Trip generation per residential zone/cell was assumed constant and uniform, and trips were distributed according to the common assumptions of the gravity model. Three decay functions were applied, two of which come from Levinson and Kumar's (1995) coefficient estimates for trip distribution as a function of time by SOV: these are for non-work trips ($f(t)=exp(.39+.16\times Time)$) and work trips ($f(t)=exp(.97+.08\times Time)$). The third relies on a simpler functional form, with time raised to 2.2 (as recommended for general analyses by Sosslau *et al.* in NCHRP Report 187 [1978]).

Results:

As expected, the relative accessibility (as measured here) of home zones closest to the largest ACs fell following AC dispersion; however, those of more distant home zones rose sufficiently to more than offset the fall under all six cases studied. Thus, in Table A-2-1, one finds only negative changes in accessibility in moving from the small-AC to the large-AC scenario. Weighted-average total trip distances (from summing the product of O-D probabilities and trip lengths [along the grid network]) across all home zones consistently rose as ACs grew more concentrated -- with the strongest increases computed for the exponential-decay non-work trips and for the simplified decay across all trip types (under both speed assumptions). Following AC concentration the average percentage rise in travel distance (from the small-AC to large-AC scenario) across all six set-ups was over 16 percent. (Note, however, that the most realistic distance estimates per trip [in miles] appear to come from the non-work exponential decay set-ups, where the average rise in per-trip distance was 11.4 percent.)

Moreover, as ACs concentrate, one expects a lower percentage of a household's trips to fall within a reasonable walking distance. For example, trips no more than two miles in length became a substantially smaller proportion of household trips when moving from the small-AC to the large-AC scenario, as tabulated in Table A-2-1. The average percentage fall of this proportion was roughly 22%. With shortened trip lengths, walking and other modes become more competitive and this increasingly likely mode switch should induce a still greater drop in VMT-per-household than trip-distance estimates alone may indicate. Note the importance of distance as illustrated by its coefficient level and statistical significance in the mode-choice models discussed in the main body of this thesis. In fact, the elasticity of the estimated probability of walk/bike mode choice is
-1.30 for the distance variable, far exceeding the impact of most other independent variables.

TABLE A-2-1. AC Scenarios' Accessibility and Travel Results

	ACTIVITY CENTER SIZE			%Rise
	SMALL	MED.	LARGE	Sm. to Lg.
Non-Work Exponential Decay:				
Speed Increasing with Distance:				
Average Accessibility	4.248	4.248	4.097	-3.6%
			(Big AC's: A, B, & C: 5.84, 3.82, 2.90)	
Average Travel Distance	14.26	14.38	15.64	+9.7
			(10.7, 15.9, 20.1)	
%Trips ≤ 2 Miles*	18.9%	17.9%	14.8%	-21.7%
			(59.1, 0, 0)	
Speed Constant at 25 mph:				
Average Accessibility	4.536	4.536	4.372	-3.6%
			(6.68, 4.13, 2.55)	
Average Travel Distance	5.272	5.340	5.956	+13.0%
			(3.87, 5.96, 8.04)	
%Trips ≤ 2 Miles*	24.2%	22.8%	18.8%	-22.3%
			(75.1, 0, 0)	
Work Exponential Decay:				
Speed Increasing with Distance:				
Average Accessibility	32.23	32.23	32.14	-0.3%
			(33.4, 32.0, 31.0)	
Average Travel Distance	41.22	41.22	41.37	+0.4%
			(39.7, 41.5, 42.8)	
%Trips ≤ 2 Miles*	2.2%	2.2%	2.6%	+18.2%
			(10.2, 0, 0)	
Speed Constant at 25 mph:				
Average Accessibility	10.25	10.25	10.16	-0.9%
			(11.6, 10.1, 8.80)	
Average Travel Distance	10.45	10.45	10.64	+1.8%
			(9.27, 10.6, 12.0)	
%Trips ≤ 2 Miles*	8.1%	8.0%	8.9%	+10.0%
			(35.6, 0, 0)	
Time Simply Raised to 2.2 in Gravity Model's Denominator:				
Speed Increasing with Distance:				
Average Accessibility	0.8514	0.8514	0.8137	-4.4%
			(0.94, 0.78, 0.75)	
Average Travel Distance	41.22	43.64	45.63	+10.7%
			(39.4, 46.9, 49.2)	
%Trips ≤ 2 Miles*	9.4%	8.7%	5.6%	-40.6%
			(22.6, 0, 0)	
Speed Constant at 25 mph:				
Average Accessibility	0.5124	0.5124	0.4070	-20.6%
			(0.72, 0.33, 0.25)	
Average Travel Distance	10.45	13.52	17.09	+63.6%
			(9.09, 18.1, 23.0)	
%Trips ≤ 2 Miles*	40.8%	31.7%	17.6%	-56.9%
			(70.4, 0, 0)	

* Note that the percentage of trips under any distance less than two miles would have left the large AC scenario with zero percent of such trips, given the way the idealized home centroids were set up (to

minimize computation time).

ENDNOTES:

1. Kulash (1990) predicts that for making trips *internal* to a community VMT will fall by 43%, thanks to the more direct access provided by a grid network. And McNally and Ryan's experimental model (1993) predicts that *total* VMT will fall by 10.6% when one moves from a "typical" PUD to an area with a "highly interconnected street system," a mixture of land uses, and a pedestrian- and bicyclist-friendly street design.

2. A Euclidean version of VMT (*i.e.*, "as the crow flies") was estimated using each trip's origin and destination tracts' centroid coordinates and was adjusted for vehicle occupancy levels. Trips were recorded only for household members of age five or older. Therefore, it is expected that VMT estimates for households with very young children are biased low since children under the age of five may have added to vehicle occupancies (and thus reduced VMT) but recorded no trips (and so could not add back to the household's VMT estimate).

If a trip left the nine-county region, its distance could not be estimated so the trip was not considered in the mode-choice models and the trip-maker's household's information was not used in the models of Euclidean VMT (adjusted for occupancy) per household. If a trip's origin and destination census tract were the same, the trip was assigned a distance estimate of 0.2 Euclidean miles (which typically translates to a quarter mile).

Additionally, if no home-based trips were recorded by the household on the survey day, the household's data were not included because the general survey data on home tract were insufficiently detailed to know the six-digit home-tract number. While four-digit tract numbers were provided for all surveyed households' location of residence, six-digit tract numbers were provided only for the origins and destinations of trips recorded.

3. Note that *association* is not causation. While cross-sectional analysis can indicate associations, the direction of causation remains obscured. For example, does high transit service primarily follow or contribute to high levels of transit demand? Does a variety of shops encourage walking more than a high number of pedestrians encourages commercial enterprise? Moreover, due to limited neighborhood data sets, many of the cited studies present results only as gross point averages of mode choice for different locales or environments, without having the degrees of freedom to control for many environmental factors at once.

4. When there is correlation between a regressor and unobserved yet relevant variables, which are carried by the error term of a model, the regressor's coefficient estimate is biased and inconsistent. (Greene 1993) Positive correlation between density and a variety of typically unincluded (and unobserved) variables that have an effect on travel decisions (such as parking fees and transit service) increases the coefficient of the density term; a more correctly specified model, with all relevant variables made explicit, would place

less (or no) emphasis on such a regressor. However, density arguably can be expected in and of itself to have some effect. For example, high transit service frequencies are more viable (and thus more common) when densities in trip-making exist; and higher population densities may reduce crime rates and allow people to feel safer walking outside of daylight hours by increasing the "eyes on the street" (Jacobs 1961).

5. "Developed area" is measured as that area under industrial, commercial, and/or residential use. A typical total-area density measure of periphery and other substantially undeveloped census tracts could seriously underestimate the densities that residents and others experience in that zone.

6. This measure of accessibility is commonly used by practicing transportation planners when they perform the second stage of the "Four-Stage" regional transportation modeling process, distributing "generated" trips across destination zones by their relative accessibility.

7. In the literature, one can also find accessibility being coarsely proxied by a distance-to-CBD variable (*e.g.*, McMillan *et al.* 1980, Krumm 1980, and Voith 1991).

8. In previous research by the author, a wide variety of accessibility indices (using different functional forms of time and maximum trip-time radii) were constructed and incorporated into models of housing price. Many of these performed almost as well in explaining housing price as the relatively few indices used as independent variables here. From this previous research, it appears that "accessibility" is relatively robust to functional specification -- at least for purposes of home/land valuation.

9. For examples of jobs-housing balance variables, see Messenger and Ewing (1996) or Frank (1995).

10. Note that the homogenous distribution of the characteristics across all zones is all that is necessary for this to be maximized, given the constraint that the individual characteristics may not be represented equally in general. For example, the population may be 10% elderly, 15% youth, and 75% otherwise; given this regional distribution, the index is maximized only if these proportions are represented in *each* of the zones.

11. The ABAG land-use data are sufficiently detailed that one could pursue a double-summation entropy index at the tract (or neighborhood) level here by dividing tracts into sub-areas and computing an entropy across the sub-areas (as well as across the many use types). However, practically speaking, this calculation requires relatively large tracts (or neighborhoods) so that there is differentiation of developed uses within the sub-areas and that there are at least several sub-areas contained in the tract to make for a reasonably stable average entropy across them.

12. Messenger and Ewing (1996) specify their "degree-mix" variable as the following:

"{[housing units x $\log_{10}$(housing units)] + [retail jobs x $\log_{10}$(retail jobs)] + [service jobs x $\log_{10}$(service jobs)] + [industrial jobs x $\log_{10}$(industrial jobs)]/(housing units + retail jobs + service jobs + industrial jobs)". While this represents another method of evaluating balance of activities (using housing units and jobs, rather than land area), it is fundamentally flawed since there is no normalization for magnitude of activity. For example, if all units (job types and housing) were doubled (as they could easily be in an area twice the size), the index level would change -- when in fact the balance across activities can not be considered different. Furthermore, this index's specification prompts the question of whether jobs and housing should compared in a one-to-three fashion such as this.

13. One feasible weighting scheme might depend on deviations from the *regional* balance of proportions across the J use types. For example, the weighting of $P_i * ln(P_i)$, or "w_i," could be equal to one minus the absolute difference between the zonal "P_i" and the regional P_i. In this way, those zones which come closest to having land uses represented in a way that is compatible with the regional land pattern achieve maximal entropy indices.

Two Land-Use Patterns, where Entropy stays constant yet Mix changes dramatically

14. The index of land-use balance, entropy, remains constant when distinct land-use types remain in constant relative proportions; yet *mixing* or integration of land uses can change dramatically. The two distinct land-use patterns shown here receive the same entropy value, while mix (or the "dissimilarity index") can be expected to have increased many, many times as a result of the two uses distributing themselves more homogeneously

15. The eleven use types distinguished in the "general mix" index are the following: residential, educational, commercial outdoor recreational (such as stadiums and golf courses), parks, religious uses, general and retail commercial uses, travel uses (such as hotels and convention centers), office uses and research parks, industrial uses, airports, and military uses. Note that a single use surrounded by non-uses (such as roadways or

marshlands) is awarded no dissimilarity points, so it is highly unlikely that hectares housing uses like airports and racetracks received any points.

16. A Lorenz curve is often used to look at the percentage of a total that corresponds to a percentage of observational units (*e.g.*, the percentage of income as earned across all households, starting from the poorest quantile). The Gini coefficient is the percentage of area under the straight equal-distribution line which lies above the Lorenz curve; thus, it is a measure of disparity (or imbalance) in distribution.

The Gini coefficient is similar to the Ogive Index, which is the sum of the squares of the differences from equal proportions over any trait, multiplied by the number of distinct levels/attribute types considered (Tress 1938).

Closely related to the Gini coefficient as well as the entropy index is the Herfindahl index, which is the sum of the squares of the proportions of distinct attributes and which has been used as a measure of economic diversity (Tauer 1992). Siegel *et al.* (1995) have found that these three, besides all varying between zero and one, produce fairly similar rankings.

17. Regressor endogeneity implies correlation between regressors and error terms, thereby violating the assumptions needed for consistent ordinary least squares (OLS) estimation of parameters (Green 1993). Consistency of estimates implies that any biases in parameter estimation converge to zero as the number of observations increases; thus, this is a highly desirable -- and critical -- property of regression analysis.

18. The following were included as personal vehicles: car, truck, taxi, limo, van, and motorcycle. Note that all of these are motorized modes.

19. Several socio-economic factors have been normalized with respect to the number of households members over age five since a household's *absolute* levels of vehicle ownership and income are not as important the availability of autos and income to each trip-making member. Unfortunately, the number of members over other age groups was not known fully since no further explicit questions of this sort were made of the surveyed households and not all household members made trips or were available to make trips in the Bay Area on the survey day.

20. The negative and positive coefficients for males in the PV and walk/bike mode-choice models, respectively, may be unexpected for some readers. However, the relative reliance of women on PVs is supported in research focusing on other urban areas and is likely due to child care responsibilities and the running of household errands; for examples, see Rosenbloom and Burns (1993 & 1994). However, like the income variable, the coefficient's sign changes in the San Francisco Bay Area for models of SOV choice.

21. The ABAG land-use data are relatively gross if one really wants to look at a spatially

detailed land-use mixture. One hectare, the area for which ABAG judges the "dominant" land use, is basically the size of two football fields, whereas land-use mixing can and does occur at a very local level, oftentimes within a single building.

22. Messenger and Ewing's (1996) structural equations had percentage of commute trips by transit, a binary auto-ownership-per-household variable, and bus-route frequency as endogenously determined.