Title
The impact of residential neighborhood type on travel behavior: A structural equations modeling approach

Authors
Bagley, Michael N
Mokhtarian, Patricia L

Publication Date
2001-10-01

Ann Reg Sci (2002) 36:279–297

The Annals of
Regional Science

The impact of residential neighborhood type on travel behavior: A structural equations modeling approach

Michael N. Bagley[1], **Patricia L. Mokhtarian**[2]

[1] South Texas Community College, P.O. Box 9701, McAllen, TX 78502-9701, USA (e-mail: mnbagley@stcc.cc.tx.us)
[2] Department of Civil and Environmental Engineering, One Shields Avenue, University of California, Davis, CA 95616, USA (e-mail: plmokhtarian@ucdavis.edu)

Received: March 2001/Accepted: October 2001

Abstract. Using a system of structural equations, this paper empirically examines the relationship of residential neighborhood type to travel behavior, incorporating attitudinal, lifestyle, and demographic variables. Data on these variables were collected from residents of five neighborhoods in the San Francisco Bay Area in 1993 (final $N = 515$), including "traditional" and "suburban" as well as mixtures of those two extremes. A conceptual model of the interrelationships among the key variables of interest was operationalized with a nine-equation structural model system. The nine endogenous variables included two measures of residential location type, three measures of travel demand, three attitudinal measures, and one measure of job location.

In terms of both direct and total effects, attitudinal and lifestyle variables had the greatest impact on travel demand among all the explanatory variables. By contrast, residential location type had little impact on travel behavior. This is perhaps the strongest evidence to date supporting the speculation that the association commonly observed between land use configuration and travel patterns is not one of direct causality, but due primarily to correlations of each of those variables with others. In particular, the results suggest that when attitudinal, lifestyle, and sociodemographic variables are accounted for, neighborhood type has little influence on travel behavior.

JEL classification: C31, D12, R14

1. Introduction

It seems self-evident that residential location decisions profoundly influence urban travel patterns, but the precise nature of that influence is not completely

The data analyzed in this paper were collected under Contract Number A132-103 with the California Air Resources Board. The authors wish to express their thanks to the many contributors to this study, including Ryuichi Kitamura, Laura Laidet, Carol Buckinger, and Fred Gianelli of UC Davis; and Terry Parker, Fereidun Feizollahi, and Anne Geraghty of the Air Resources Board. Yoram Shiftan, Juan de Dios Ortuzar and an anonymous referee made some insightful suggestions on a previous draft.

understood. For example, numerous empirical studies (e.g., Frank and Pivo 1994; Kitamura et al. 1997) have demonstrated that living in higher-density, mixed-use neighborhoods (variously referred to as traditional, neo-traditional, urban, or New Urban) is associated with fewer vehicle trips and smaller distances traveled compared to living in typical low-density suburban environments. These encouraging results have supported the growing New Urbanism movement to adopt land use planning and design as a tool for reducing travel (see, e.g., Calthorpe and Richmond 1992; Fulton 1996). For example, the US Environmental Protection Agency is developing guidelines for allowing air quality improvement credit for developments considered to exhibit traditional or New Urban characteristics (see, e.g., Jack Faucett Associates and Sierra Research 1999; US EPA 2000). But whether the land use configuration itself prompts the observed travel patterns, or whether people with different prior travel propensities select themselves into residential neighborhoods that support those propensities, is impossible to determine from residential location and travel data alone. Statistical correlations between measures of each type demonstrate association, but do not identify the proper direction(s) of causality.

The distinction is important: if travel patterns are primarily a result of attitudinal and lifestyle predispositions and are not generated by the land use configuration itself, then we may overestimate the travel reduction potential of the widespread adoption of more compact land use patterns if that adoption results in a mismatch between residential preference and choice. Just as importantly, we may now be *underestimating* the disparity in travel patterns associated with the different types of land use, due to *current* mismatches between residential preference and choice that are factored into the average travel statistics for each land use type. In either case, we will be poorly equipped to predict the travel impacts of changing land use patterns, as residential location alternatives become more available for some people and more constrained for others.

Thus, to understand the extent to which travel-related predispositions influence residential location and then travel behavior given residential location, it is necessary (1) to obtain and incorporate data on individuals' attitudes and lifestyle preferences, and (2) to account for potential multiple directions of causality to the extent possible. This paper reports on an empirical study involving both of those elements. Specifically, we develop and estimate a structural equations model containing measures of travel behavior and residential location type among the endogenous variables, and measures of attitudes and lifestyle among the explanatory variables. To our knowledge, this is the first disaggregate structural equations model containing all of these types of variables.

The organization of this paper is as follows. Section 2 discusses the empirical setting of this research, and the data available to the study. Section 3 presents the structural equations model and the empirical results. The final section summarizes and comments on the key findings of the research.

2. Empirical setting and data available

2.1. Empirical context

The data used for this study were originally collected for a land use-travel behavior project sponsored by the California Air Resources Board in 1992. Microscale data on land use, the roadway network, and public transit were obtained from site surveys of five San Francisco Bay Area neighborhoods (selected sections of approximately one square mile within the cities or areas of Concord, Pleasant Hill, North San Francisco, South San Francisco, and San Jose). In addition, demographic, socioeconomic, attitudinal, lifestyle, and travel-related data were collected through mail-out surveys and travel diaries completed by residents in the same neighborhoods. The main objective of the original study was to examine the impacts of neighborhood type (i.e., land use) and individual attitudes on travel behavior (Kitamura et al. 1997). Thus, the neighborhoods were selected to represent a range of values on key characteristics of land use type, including public transit accessibility, land use mix, residential density, and employment mix.

About 18% of those initially contacted (randomly selected from address lists covering the study neighborhoods) agreed to participate, and 60% of those completed all three surveys involved. From the 963 households completing any of the surveys, 515 individuals from different households were selected for the analysis reported in this paper. The selection was based on relative completeness of the information on the key variables of the model, the elimination of unemployed and retired respondents from the sample (since their residential location decision processes may differ from those of employed individuals, and since commute distance is an important variable in the structural equations model), and other considerations described in Bagley (1999) and in Sect. 3 below. Thus, although the estimation sample cannot be asserted to be perfectly representative of the population as a whole, it is suitable for the purpose of this study, which is to model relationships among variables, not to project sample distributions to the population.

Table 1 provides descriptive and summary statistics for the final estimation sample on typical demographic and travel variables. Respondents tended to be professional, well-educated, and having moderate incomes. The average age was 45; the average household size was 2.3 people. Respondents were long-time residents of the Bay Area – 25 years on average. Each driver typically had a vehicle available, and the average one-way commute distance was 11 miles. The average 4.3 person-trips per day is consistent with travel diary results from other studies such as the 1995 Nationwide Personal Transportation Study (4.2 person-trips per day; FHWA, 1997).

2.2. *Key variables*

The information on lifestyle and attitudes constitutes an unusual feature of this data set. One section of one of the questionnaires used to collect the data contained 39 statements relating to attitudes toward the private automobile, ridesharing, public transportation, congestion and air quality, time use, housing preferences, and economic policies related to transportation. Respondents used a five-point Likert-type scale to express their level of agreement with each statement. Another section contained lists of more than 100 types of activities and interests, with the respondents variously indicating which subjects they read about within the last month, how they spent the last weekend, and what activities they conducted during the last year. These responses are indicators of lifestyle preferences.

Responses in each section were separately factor-analyzed to identify the underlying dimensions common to the correlated responses. The attitudinal

Table 1. Key characteristics of the final estimation sample[1] ($N = 515$)

Variable	NSF ($N = 95$) SSF ($N = 100$) CON ($N = 87$) PH ($N = 121$) SJ ($N = 112$)
Occupation[2]: Number (%)	
Manager/administration	98 (19.0%)
Professional/technical	238 (46.2%)
Administrative support	82 (15.9%)
Household composition: Mean (standard deviation)	
Household size	2.31 (1.12)
No. people 16 or over	1.85 (0.86)
No. people under 16	0.42 (0.76)
No. full-time workers	0.80 (0.40)
No. workers (part- and full-time)	1.45 (0.50)
Personal characteristics: Mean (standard deviation)	
Age	44.8 (9.4)
Education category[3]	4.15 (1.27)
Female (=1, Male = 0)	0.55 (0.50)
Household income category[3]	6.64 (1.20)
Years lived in Bay Area	24.7 (15.6)
Travel characteristics: Mean (standard deviation)	
No. of vehicles	1.94 (0.93)
No. of vehicles/driver	1.05 (0.44)
One-way commute distance (mi.)	10.81 (9.62)
Daily person trips	4.28 (1.80)
Daily vehicle-miles traveled	28.41 (24.94)
Daily transit-miles traveled	4.35 (11.79)
Daily walk/bike-miles traveled	0.22 (0.53)

[1] The values given in this table are based on imputed means replacing missing data.
[2] Not all job categories are presented, and thus percentages do not sum to 100%.
[3] For education, a value of 4 represents completion of 4-year degree; for household income, a value of 6 represents the range \$35,001 to \$50,000 per year.

variables grouped into 10 factors, such as pro-environment, pro-growth, pro-transit, and pro-driving. The lifestyle variables grouped into 11 factors, such as hobbyist, nest-builder, athletic, child-oriented, and couch potato. These attitudinal and lifestyle factors varied significantly between residents of the two neighborhood types, in expected ways (Bagley and Mokhtarian 1999). For example, on average, residents of traditional North San Francisco had significantly higher scores on the pro-high-density attitudinal factor, lower scores on the anti-urban life attitudinal factor, higher scores on the culture lover lifestyle factor, and higher scores on the athletic lifestyle factor than their suburban counterparts. The most important variables comprising the 7 lifestyle and 10 attitudinal factors significant in the final model are presented in Table 2; details on these and the other factors are available in Bagley (1999) and Bagley and Mokhtarian (1999).

It is also important to briefly describe the way in which residential location was measured for this analysis (a more detailed explanation appears in Bagley et al. 2002). A respondent in this study lives in one of five neighborhoods, each of which could be considered an indicator of residential choice. Indeed, some residential choice studies, such as Horowitz (1995), take census tracts or other location indicators as the dependent variable. However, to develop residential choice models that are robust and transferable, the generic characteristics of a

Table 2. Key variables loading on the lifestyle and attitudinal factors significant in the final model

Lifestyle factors[1]	
Adventurer	Went hunting; used off-road vehicle; went to a shooting range; participated in a motor cross
Culture-lover	Attended concert or symphony; attended ballet; read material on art or architecture; attended theater
Hobbyist	Read material on: science/nature; environment; outdoors; history; photography; humor; pets; spent last weekend doing hobbies
Homebody	Read material on: women's issues; fashion; cooking/recipes; decorating; sewed; did needlework
Nest-builder	Read material on: home improvement; gardening; made own house improvements; put in a garden; spent last weekend doing yard work
Outdoor enthusiast	Visited: national park/historic site; state park/historic site; local park/historic site; beach; went hiking/backpacking/camping
Relaxer	Spent last weekend: reading; at home relaxing; shopping; doing chores
Attitudinal factors	
Pro-alternatives	Shops and services within walking distance of home important; vehicle emissions increase the need for health care; I use public transportation when I can't afford to drive; should provide incentives to use electric/clean-fuel vehicles; more lanes should be set aside for carpools and buses
Pro-drive alone	I like someone else to do the driving (–); I am not comfortable riding with strangers; ridesharing saves money (–)
Pro-environment	Environmental protection costs too much (–); environmentalism hurts minority and small businesses (–); people and jobs are more important than the environment (–); environmental protection is good for the economy; stricter smog control laws should be enforced
Pro-growth	Need to build more roads to help decrease congestion; too many people drive alone (–); too much agricultural land is consumed for housing (–); getting stuck in traffic doesn't bother me too much
Pro-pricing	I'd be willing to pay a toll to drive on an uncongested road; should raise gas prices to reduce congestion and air pollution; congestion will take care of itself because people will adjust (–)
Time-satisfied	Would like to have more time for leisure (–); I am wasting time when I have to wait (–); getting stuck in traffic doesn't bother me too much
Work-driven	I like to spend most of my time working; when things are busy at work, I cut back on personal time
Pro-high density	Need to have space between me and my neighbors (–); would only live in a multiple family unit as a last resort (–); important for children to have a large backyard (–); high-density residential development should be encouraged
Pro-driving	Driving allows me to get more done; driving allows me freedom; would rather use a clean-fuel car than give up driving
Pro-transit	Public transportation is unreliable (–); costs more to use transit than to drive (–); buses and trains are pleasant to travel in; I can read and do other things on transit

[1] The time frame for these activities is as follows: "Read material on . . ." within the past month; all other activities occurred within the past 12 months except where noted to have taken place the past weekend.

neighborhood are of greater interest than a specific geographic location itself. The concept of "traditionalness" is the key characteristic chosen for this study (though many other traits such as aesthetic appeal could be used).

A number of studies have used a similar concept as the basis for defining residential location. In most cases (e.g., Friedman et al. 1994; Cervero and Radisch 1996; Bagley and Mokhtarian 1999), a dichotomous characterization is employed: a given neighborhood is homogeneously characterized as either traditional or suburban. However, there are several problems with this. First, the concept of traditionalness is not necessarily a single dimension, but rather is a composite of a number of different characteristics. Although these characteristics may be correlated, a neighborhood may be more stereotypically traditional on some characteristics than on others. Thus, a neighborhood's composite traditionalness rating presumably falls along a continuum rather than into one "bin" or another, and forcing it into one discrete category or the other distorts its true nature. Alternatively, if "hybrid" neighborhoods are discarded from analysis, useful data are ignored and the results are less generalizable. Further, even if a neighborhood's traditionalness rating is allowed to be continuous rather than discrete, assigning a single such rating to an entire neighborhood is too restrictive. Individuals within the same neighborhood may face different densities, and different proximities to commercial opportunities and to public transit, so the degree of traditionalness may vary considerably even within a single neighborhood.

To address these concerns, we wanted to develop a continuous, disaggregate measure of neighborhood traditionalness. A review of the literature on land use and travel (see, for example, Friedman et al. 1994; Cervero and Radisch 1996; Southworth 1997; and Tong and Wong 1997) identifies many characteristics associated with traditional neighborhoods (such as mixed land uses, high density, and transit access). Eighteen of these characteristics were available in our data set: 15 at disaggregate levels (e.g., perceived pleasantness of walking and cycling in the neighborhood, parking availability, distance to nearest public transit and grocery store, presence of sidewalks) and three only at aggregate levels (average speed limit, indicator of grid street system, and indicator of population density). We applied principal component analysis (Rummel 1970) to these eighteen measures (where for the three aggregate variables, each individual in a particular neighborhood had the same value).

We had hypothesized the existence of a single traditionalness construct, with the principal component analysis identifying the optimal weighting of each variable in determining the construct. Instead, two distinct dimensions emerged from the analysis: a *traditional* factor (with variables related to population density and public transit convenience loading positively, and variables related to home size, presence of a backyard, and parking availability loading negatively) and a *suburban* factor (with variables related to speed limit, distance to nearest grocery store and park, and ease of cycling loading positively, and the indicator of a grid street network loading negatively). Rather than traditionalness being a single "either-or" characteristic, neighborhoods could and did score high or low on both characteristics. For example, Pleasant Hill not only had the highest mean score on the suburban factor, but also the second highest mean score on the traditional factor. We also saw considerable variation in both factor scores across individuals within the same neighborhood, confirming the importance of using a disaggregate measure.

Thus, in the structural equations model presented in the next section, residential neighborhood type is measured by two variables: the individual's scores on the traditional and suburban factors. As continuous variables, these factor scores also have the advantage of being more tractable for use in a structural equations modeling context than discrete indicators of neighborhood type would be.

3. Structural equations model and results

3.1. Model specification issues

Structural equations models (SEMs) are useful for representing multiple relationships among a set of variables, where the same variable that is the outcome (dependent variable) in one set of relationships may be a predictor of outcomes (explanatory variable) in other relationships. Our context certainly involves multiple relationships among a set of variables. For example, we can hypothesize that attitudes and lifestyle characteristics affect both residential location and travel demand and in turn are affected by them, and that residential location and travel demand each affects the other. It is easy to see that a model that permits only one set of relationships to be estimated (say, the effect of residential location and other variables on travel demand), while ignoring other potentially important relationships, is likely to be a serious distortion of reality.

Hence, the structural equations model estimated in the current study represents a significant improvement over single-equation models. However, the current results cannot be taken as definitive since they, too, represent a simplification of reality. A model that allows each variable to impact every other variable is not statistically identifiable – that is, a "best" set of coefficients cannot be uniquely determined. Thus, it is necessary to impose some prior restrictions on the model. These restrictions in the best case involve the analyst's unhindered judgment as to which directions of causality are most plausible, and in the worst case are more or less dictated by the availability of relevant exogenous variables. The present situation falls in between those two extremes, closer to the best case end of the spectrum.

Based on an extensive review of the literature relating to residential choice and travel demand modeling, a reasonably comprehensive conceptual model was developed (Bagley 1999) for the relationships among residential location, travel demand and supply, job location, attitudes and lifestyle, demographic characteristics, and other variables. This model was not identifiable, and required simplification in order to be statistically estimable. The simplified model is the result of both conceptual reasoning and statistical experimentation, and succeeds in being acceptably realistic as well as econometrically identifiable.

The final model contains nine equations, that is, nine endogenous variables. Two variables represent residential location: the traditional and suburban factor scores discussed in Sect. 3. Three variables represent attitudes hypothesized to be endogenous to the system: pro-high-density, pro-driving, and pro-transit factor scores. Three variables represent travel demand: average daily miles traveled by personal vehicle, transit, and walk/bike respectively. The final variable represents job location: commute distance. To simplify the model structure, demographic and lifestyle variables were assumed to be exogenous. This is reasonable for the present cross-sectional data set. Over time, however, demographic and lifestyle measures might be influenced by residential location and other variables, and hence should be modeled as endogenous if longitudinal data were available.

3.2. Conformance with the assumption of multivariate normality

Using the software package AMOS (Arbuckle 1997), the SEM was initially estimated on a sample of 615 observations, with the endogenous variables in their original form. The parameters of the resulting equations were consistent with prior expectations. However, the validity of the estimated relationships (in particular, the validity of hypothesis testing used to determine whether a relationship is significant or not) is theoretically dependent on the SEM meeting the assumption of multivariate normality of its variables. When this assumption is true, the variance of the estimated parameters is consistently estimated by sample variances, but when it is false, the standard errors of parameter estimates can be substantially underestimated, leading to false conclusions of significance (West et al. 1995).

A review of the literature reveals that meeting this condition is a problem in many studies. Bentler and Dudgeon (1996, p. 566) stated that "in practice [for structural equation models], the normality assumption will often be incorrect." Micceri (1989) reviewed numerous data sets that were used in journal articles and found that a majority of the conclusions were based on data that were nonnormally distributed. Other researchers (e.g., Breckler 1989; Gierl and Mulvenon 1995) have noted that it is very common for practitioners to ignore the assumption of normality and to make conclusions as if the assumption were met.

We considered it important at least to test for departures from normality, and to attempt to achieve normality or come as close as practicable. Thus, we reviewed the Mardia statistic (a measure of multivariate kurtosis) associated with our original SEM. That statistic was equal to 313.15, with a critical ratio of 72.28 (a critical ratio above 1.96 signifies departure from multivariate normality with 95% confidence). Given this significant a failure of the assumption, modifications were in order.

The first step was to transform (taking the natural log or square root) the six variables (daily vehicle-miles, daily transit-miles, daily walk/bike-miles, adventurer, number of persons under the age of sixteen, and number of vehicles) that had high kurtosis values, as such transformations have been found to be potentially effective in making the distribution of a variable more normal (West et al. 1995). After re-estimating the previous model with the newly transformed variables, the resulting Mardia statistic was substantially improved, but it still slightly exceeded 100.

Since little is known empirically about the circumstances under which correct statistical inferences can be drawn even when the normality assumption is not met (Bentler and Dudgeon 1996), this degree of departure was still deemed large enough to warrant further corrective action. The next step was to find and remove outliers, as "extreme data points may affect the results of structural equation modeling" (West et al. 1995, p. 61). AMOS provided the Mahalanobis distance (see, e.g., Everitt 1993) for each case in the data set, where the greater the Mahalanobis distance, the greater the contribution to the departure from multivariate normality. Based on this information, cases were removed ten at a time until the remaining data set had a multivariate normal distribution. The removal of 100 cases led to the final sample of 515 respondents for which a structural equations model was estimated, resulting in a Mardia statistic of 0.55 with a critical ratio of 0.12.

While the removal of apparent outliers (especially so many of them) was not an appealing step, the alternative of egregiously violating the model assumptions was even more unattractive. In support of this step, it can be noted that the reduced sample was not substantially different from the larger sample in terms of mean values on key variables, and further, the findings from the model that met the assumption of multivariate normality were very similar to the results of the earlier model, on the larger sample, that did not meet the assumption (see Bagley 1999 for the estimated coefficients and discussion of that model).

These results, plus the closeness of the two models on several goodness-of-fit measures, suggest that the sizable departure from normality exhibited by the original model does not, in fact, materially affect the outcome in this case. On the one hand, it may be argued that the results of two models, each less than ideal for different reasons, would not necessarily be "true" just because they are similar. In response, it may be argued that when one set of equations is "wrong" for a certain reason (departure from normality), and fixing that problem (by selecting a large subset of cases exhibiting normality) yields very similar results, it does suggest a certain robustness. Had the removal of 100 cases yielded very different results, it would not have been surprising; we would have said that the 100 outliers that were removed had skewed the initial results, and the departure from normality made those initial results unreliable in any case. But when removing outliers and fulfilling the assumptions correctly does *not* change the results, it seems more plausible to think that both models are basically "right" than to believe that both models are wrong in just the same way on all the many coefficients in the equation structure.

The results of the transformed model (i.e., the model that met the multivariate normality assumption) are presented and discussed next.

3.3. Estimation results

The direct effects and total effects estimated for the final model are shown in Tables 3 and 4, respectively, with some customary goodness of fit measures presented in Table 5. The direct effects table illustrates the structure of the final model, with blank cells representing coefficients constrained to be zero. It is important to understand that the absence of a variable from the model can arise in two different ways: either (1) a relationship was *a priori* assumed to be insignificant for conceptual reasons or because it was felt to be the best way to achieve identification (as discussed in Sect. 3.1), or (2) a relationship was hypothesized to be (potentially) significant but was empirically found not to be. In particular, the residential location type variables were always *allowed* to enter the equations for travel demand; the fact that they do not appear in the final model is because they were statistically insignificant when allowed to enter, not because they were excluded in the first place.

The total effect of one variable on a second variable is the sum of the direct effect (represented by the coefficient of the first variable in the equation for the second) and all of the indirect effects arising from the first variable acting on

Table 3. Direct effects (t-statistics) for transformed[1] SEM

Endogenous variable →	Residential location		Attitudes			Travel demand			Job location
Explanatory variable ↓	Tradi-tional	Sub-urban	Pro-hi density	Pro-driving	Pro-transit	Ln vehicle miles[1]	Ln transit miles[1]	Ln walk/bike miles[1]	Commute distance (miles)
Constant	1.822 (8.3)	−1.034 (−3.8)	0.204 (3.9)	−1.407 (−4.7)	−0.046 (−1.1)	3.838 (4.3)	0.823 (3.1)	0.160 (11.6)	−3.313 (−1.2)
Sociodemographic									
Age	−0.018 (−5.06)	0.009 (2.61)							
Female				0.543 (6.03)		−0.812 (3.63)			
Household size	−0.081 (−1.92)								
Ln number of people under age 16			−0.328 (−4.86)						
Sq number of vehicles	−0.761 (−5.44)					1.652 (5.17)	−0.870 (−6.30)		
Years lived in Bay Area				0.004 (2.21)					
Lifestyle									
Ln adventurer						0.554 (2.40)			
Culture-lover		−0.178 (−4.48)	0.230 (5.80)						
Hobbyist		−0.099 (−2.75)							
Homebody		0.217 (5.50)							
Nest-builder	−0.143 (−3.53)		−0.140 (−3.61)						
Outdoor enthusiast			0.106 (2.87)						

Relaxer			0.042						
			(1.63)						
Attitudes									
Pro-alternatives		−0.070						0.026	
		(−1.57)						(2.13)	
Pro-drive alone					0.178				
					(4.21)				
Pro-environment					0.161				
					(3.73)				
Pro-growth							−0.058		
							(−2.07)		
Pro-pricing	0.108	−0.074							
	(3.11)	(−2.23)							
Time-satisfied				−0.109					
				(−3.19)					
Work-driven			0.092						
			(2.55)						
Pro-high density[2]	0.591								
	(5.00)								
Pro-driving[2]			−0.702					−0.082	
			(−4.41)					(−2.00)	
Pro-transit[2]							0.286		
							(2.56)		
Residential location									
Suburban[2]									2.149
									(2.84)
Travel demand									
Ln vehicle miles[2]				0.226					2.213
				(4.07)					(4.89)
Ln transit miles[2]									5.732
									(4.96)
Ln walk/bike miles[2]						−7.201			
						(−2.04)			
Job location									
Commute distance[2]		0.059					0.065		
		(2.62)					(3.24)		

[1] Ln = Natural log transformation, Sq = Square root transformation.
[2] Endogenous variable.

Table 4. Total effects for transformed[1] SEM

Endogenous variable →	Residential location		Attitudes			Travel demand			Job location
Explanatory variable ↓	Traditional	Suburban	Pro-hi density	Pro-driving	Pro-transit	Ln vehicle miles[1]	Ln transit miles[1]	Ln walk/bike miles[1]	Commute distance (miles)
Constant	0.052	0.015	0.053	0.101	−0.037	4.802	0.616	0.149	10.88
Sociodemographic									
Age	−0.018	0.011					0.002		0.038
Female	−0.172	−0.148	−0.292	0.416		−0.563	−0.161	−0.035	−2.489
Household size	−0.081								
Ln number of people under age 16	−0.194		−0.328						
Sq number of vehicles	−0.940	−0.091	−0.303	0.432		1.910	−0.969	−0.036	−1.523
Years lived in Bay Area	−0.002		−0.003	0.005		0.003	0.001	−0.005	0.013
Lifestyle									
Ln adventurer	−0.060	0.045	−0.102	0.145		0.641	0.184	−0.012	2.832
Culture-lover	0.136	−0.224	0.230				−0.050		−0.765
Hobbyist		−0.124					−0.028		−0.425
Homebody		0.272					0.060		0.930
Nest-builder	−0.226		−0.140						
Outdoor enthusiast	0.063		0.106						
Relaxer	−0.020	0.008	−0.034	0.048		0.029	0.008	−0.004	0.127
Attitudes									
Pro-alternatives	0.021	−0.146	0.035	−0.050		−0.220	−0.082	0.031	−1.272
Pro-drive alone		0.035			0.178		0.088		0.581
Pro-environment		0.031			0.161		0.080		0.528
Pro-growth		−0.040					−0.101		−0.664
Pro-pricing	0.108	−0.093					−0.021		−0.318
Time-satisfied	0.052	−0.020	0.089	−0.126		−0.076	−0.022	0.010	−0.334
Work-driven	0.054		0.092						
Pro-high density[2]	0.591								

Pro-driving[2]	−0.479	0.182	−0.811	0.156		0.692	0.198	−0.096	3.056
Pro-transit[2]		0.195					0.498		3.271
Residential location									
Suburban[2]		0.255					0.278		4.293
Travel demand									
Ln vehicle miles[2]	−0.108	0.304	−0.183	0.262		0.156	0.331	−0.022	5.111
Ln transit miles[2]		0.681					0.742		11.450
Ln walk/bike miles[2]	0.780	−2.189	1.321	−1.883		−8.327	−2.385	0.156	−36.804
Job location									
Commute distance[2]		0.119					0.129		0.997

[1] Ln = Natural log transformation, Sq = Square root transformation.
[2] Endogenous variable.

Table 5. Measures of fit for the structural equations model ($N = 515$)

Degrees of freedom	213
χ^2: measures discrepancy between the sample and population covariance matrices; the smaller the better.	554.100
χ^2/d.f.: a "relative" chi-square value corrected for degrees of freedom; values in the range 1 to 3 are indicative of an acceptable fit, with values closer to 1 being better.	2.601
Normed Fit Index (NFI): proportion of worst (independence) model χ^2 explained by the model of interest; varies between 0 and 1, with 1 being the best.	0.968
Relative Fit Index: NFI corrected for degrees of freedom; values close to one represent a good fit.	0.893
Incremental Fit Index: the incremental improvement of the model of interest over the worst (independence) model; values close to 1 indicate a good fit.	0.980
Comparative Fit Index: assumes a noncentral χ^2 distribution for the worst (independence) model discrepancy; values close to 1 represent a good fit.	0.979

intervening variables that in turn affect the second variable. To reduce the cognitive burden on the reader, Table 4 suppresses total effects that are very small (zero out to several decimal places). Thus, in contrast to Table 3, blanks in Table 4 may indicate either structural zeroes or non-zero but negligibly small effects. Even so, many more cells are filled in Table 4 than in Table 3, as is typical for total effects. Examining the direct effect only can be quite misleading (as comparing Tables 3 and 4 will indicate), since indirect effects may be quite strong and different from the direct effects alone. For brevity, the discussion of the results will focus on the main total effects of interest, although a careful comparison of direct and total effects is instructive.

3.3.1. Residential location

Turning first to the two residential location equations, we see a large number of total effects. With respect to demographic variables, consistent with expectations, age, household size, number of children, number of vehicles, and number of years lived in the Bay Area are all negatively associated with living in a traditional neighborhood. The lifestyle and attitudinal factor scores are also prominent in these two equations. Being a culture-lover, an outdoor enthusiast, pro-alternatives, pro-growth, pro-pricing, time-satisfied, work-driven, or pro-high density are associated with traditional residential locations, while being an adventurer, a homebody, a nest-builder, a relaxer, pro-drive alone or pro-driving, pro-environment, or pro-transit are associated with suburban residential locations. For the most part, these relationships are also stereotypical. The latter two (pro-environment and pro-transit attitudes associated with suburban residences) are somewhat surprising, and it is of interest to note that they arise only through indirect effects: the direct effects coefficients are not significant, as Table 3 indicates. This is a good example of the way in which complex interrelationships among a set of variables can lead to impacts that are obscured when only direct effects are examined, but that are revealed when those complexities are taken into account.

The impacts of the travel and job location variables on residential location have the expected signs. Longer commute distances, more vehicle-miles, and fewer walk/bike miles are associated with suburban locations. Transit-miles are also positively associated with suburban locations, due to the availability of the BART rail transit mode to several of the study neighborhoods, with substantial bus service as well. The imputation of causality here is somewhat indirect, however: current miles traveled by a given mode cannot strictly be said to have caused the (prior) choice of a particular residential location. In fact, it is the opposite direction of causality that is plausible in this context, and that is normally suggested by studies of this type: current residential location causes travel by a given mode. As will be seen below, however, this latter direction of causality is not at all supported by the empirical results, although it was quite consciously allowed in the original model specification. Thus, again, the usefulness of the SEM construction is apparent, in that it demonstrates that the causal inferences commonly drawn from the same associations between travel and residential location that we see here are not (at least in this case) statistically the most appropriate. Perhaps the most natural interpretation of the current results, then, is that the travel demand variables (together with the attitudinal variables) are serving as a proxy for travel predisposition (i.e., how much a person naturally desires to travel by various modes), and that the individual's choice of residence will be based on its ability to meet this predisposition. This also serves to illustrate the well-known difficulty in making true causal inferences based on cross-sectional data, since causal impacts are rarely instantaneous. Having a longitudinal data set, in which temporal precedence is clearly established, would strengthen the ability to infer causality.

3.3.2. Travel demand

Turning to the total effects on travel demand, several relationships are of interest. Among the demographic variables, number of vehicles is – not surprisingly – positively related to vehicle-miles and negatively related to miles by transit or walk/bike. All else equal, women travel less by all three modes, similar to the findings of other studies (while it is frequently noted that women use transit more than men, women who use transit still tend to travel shorter distances on it than men who use transit; Mauch and Taylor 1997). With respect to lifestyle and attitudinal variables: being an adventurer, pro-driving, not favoring transportation alternatives, or not being time-satisfied are associated with greater vehicular travel. The transit-miles variable shows some perhaps unexpected relationships with attitudes, e.g. being positively associated with the pro-driving attitudes. It should be noted, however, that most of those relationships are indirect only, and relatively small in magnitude (although the effects presented in Tables 3 and 4 cannot be directly compared due to differences in scale among the variables, this comment is supported by the estimation results – not shown here – obtained for the standardized variables). They are consistent with the observation that much of the transit use in this sample involves patronage of the BART rail system by suburban residents, who also tend to use autos extensively.

Along the same lines, the *only* significant effect of residential location on travel demand is the positive effect of a suburban location on transit miles. This is hardly the stereotypical relationship – it is supposed to be traditional neighborhoods that promote transit use – but is again consistent with the nature of transit use in our sample. Just as important is the conspicuous *lack* of any other effects (especially the expected ones) of residential location on travel demand.

4. Summary and conclusions

In analyzing the effect of residential location on travel demand, it is important to account for (1) multiple directions of causality to the extent possible, and (2) the role of attitudinal predispositions and self-selection in generating the observed results. This study attempted to do both of these things, through the development of structural equations models (SEM) of residential location and travel demand that included attitudinal and lifestyle explanatory variables. To our knowledge, this is the first such disaggregate SEM of residential location, particularly involving the range of variables included here.

The results showed that, in terms of both direct and total effects, attitudinal and lifestyle variables had the greatest impact on travel demand among all the explanatory variables. By contrast, residential location type had little impact on travel behavior. This is perhaps the strongest evidence to date supporting the speculation that the association commonly observed between land use configuration and travel patterns is not one of direct causality, but due primarily to correlations of each of those variables with others. In particular, the results suggest that when attitudinal, lifestyle, and sociodemographic variables are accounted for, neighborhood type has little influence on travel behavior.

Supporters of New Urbanism appropriately point out that even if the influence of residential location on travel demand is entirely due to self-selection, that does not abrogate the desirability of providing more neo-traditional location options for households (Levine 1999). To those who argue that market forces predominantly favor low-density suburban development, new urbanists respond that the tax break on mortgage interest and other policies distort the market, and that in any case there is a segment preferring higher-density, mixed use residential environments that is currently under-supplied. Without denying that possibility, an interesting recent stated preference study of Portland, Oregon residents (Shiftan and Suhrbier 2000) suggests that many people express a desire for having the best of both types of environment – but that in terms of willingness to pay, suburban features such as an attached garage, larger residences and larger lots dominate neo-traditional features such as nearby shops and access to transit.

This debate points to the need for better models of residential location preference, as a foundation to understanding travel demand given location preference. Disaggregate models providing behavioral insight, such as the ones presented here and in Bagley and Mokhtarian (1999) and Shiftan and Suhrbier (2000), should be complemented by aggregate forecasts of the demand for each type of neighborhood (not just housing) under realistic constraints.

The study presented here, although contributing to our knowledge on this issue, is by no means definitive. The econometric conditions of identifiability and the availability of data limited the number and kinds of relationships that could be tested. Thus, although the key relationships between travel demand and residential location were tested (among many others), some other relationships that could affect those main ones may have been overlooked or constrained from entering the model. Further, each conceptual variable could be operationalized in a number of different ways with our data, and while a great number of alternative specifications were tested, it is conceivable that a better one was overlooked. Another limitation is the focus on single individuals within households rather than on household interaction dynamics.

One particular alternate specification may be worth mentioning. It could be argued that our finding of a positive effect of suburban residence on transit use was a natural consequence of the choice of distance as the measure of travel demand, since suburban residents who use transit (especially rail transit, as was the case here) are likely to use it for longer journeys than would urban residents. We tested a number of specifications using daily trip rates by each mode rather than distance, but found their goodness of fit to be inferior to the specification presented here.

As a philosophical observation, it is unlikely that self-selection based on prior attitudes completely explains the observed results (of this and other studies). Rather, it seems quite plausible that the residential environment would have some impact on travel behavior. If nothing else, the spatial configuration of the residential neighborhood can impose constraints on the behavior to which one may be predisposed. An urbanite-at-heart may love to walk to that corner café for an espresso, but if she lives in a suburb with no such establishments within walking distance, she may drive to one in a strip mall or forgo the activity altogether. A suburbanite-at-heart may prefer driving everywhere, but if he is living in an urban neighborhood where parking is scarce and expensive while transit is excellent, he may reluctantly give up the automobile or reduce his use of it.

Thus, people do change, both their attitudes and their behavior, in response to external stimuli – the questions are, how many people, which kinds, how much, and how long does it take? A process of attitudinal and behavior adjustment, whether due to physical constraints as described above or due to a more subtle alteration of attitudes over time, comes into play most forcefully when people's predispositions and residential locations are mismatched, and the extent to which that is the case is unknown. The current study not only found little effect of residential location on (travel) behavior, it found *no* impact of residential location on attitudes (as indicated in Table 4). Travel behavior, on the other hand, showed a tendency to reinforce related attitudes: vehicle-miles positively affected the pro-driving attitude and negatively affected the pro-high-density attitude, and the converse was true for walk/bike miles. These results are consistent with theories and previous empirical findings relating to cognitive balance and mutual dependence between attitudes and behavior (e.g., Golob et al. 1979; Reibstein et al. 1980).

However, a major limitation of the current study is the inability of the available cross-sectional data to capture dynamic changes. Those changes can, of course, best be identified with a longitudinal data set that measures attitudes, lifestyle, demographics, and travel behavior for a panel of households across one or more residential moves, including an adjustment period after the move. But such a study must be careful to control for changes in demographic characteristics (such as household size or income) that may be related to a residential move and that may be confounded with the effect of the new residential location itself on travel behavior.

To conclude, evidence strongly suggests that land use characteristics have little independent impact on travel behavior. But a need still exists to provide a diversity of residential location types consistent with the diverse tastes found in the population (particularly when a segment that prefers modes of travel other than the automobile is under-served), and, through the use of more appropriate data and analysis techniques, to improve our understanding of the extent to which one's residential environment influences the attitudes and lifestyle that *do* affect travel demand.

References

Arbuckle JL (1997) Amos users' guide version 3.6. SmallWaters Corporation, Chicago
Bagley MN (1999) Incorporating residential choice into travel behavior – land use interaction research: A conceptual model with methodologies for investigating causal relationships. PhD Dissertation, Department of Civil and Environmental Engineering, University of California, Davis
Bagley MN, Mokhtarian PL (1999) The role of lifestyle and attitudinal characteristics in residential neighborhood choice. In: Ceder A (ed) Transportation and traffic theory: Proceedings of the 14th International Symposium on Transportation and Traffic Theory. Pergamon Press, Oxford pp 735–758
Bagley MN, Mokhtarian PL, Kitamura R (2002) A methodology for the disaggregate, multidimensional measurement of neighborhood type. Urban Studies 39(4):689–704
Bentler PM, Dudgeon P (1996) Covariance structure analysis: Statistical practice, theory, and directions. Annual Review of Psychology 47:563–592
Breckler SJ (1989) Application of covariance structure modeling in psychology: Cause for concern? Psychological Bulletin 107:260–273
Calthorpe P, Richmond H (1992) Land use and transportation planning: Sustainable growth. In: Green M (ed) Changing America: Blueprints for the new administration. Newmarket Press, New York, pp 699–710
Cervero R, Radisch C (1996) Travel choices in pedestrian versus automobile oriented neighborhoods. Transport Policy 3(3):127–141
Everitt BS (1993) Cluster analysis. Halstead Press, New York
Federal Highway Administration (FHWA) (1997) Our Nation's Travel: 1995 NPTS Early Results Report. Publication No. FHWA-PL-97-028, USDOT FHWA Office of Highway Information Management, Washington, DC, September
Frank L, Pivo G (1994) Impacts of mixed use and density on utilization of three modes of travel: Single-occupant vehicle, transit, and walking. Transportation Research Record 1466:44–52
Friedman B, Gordon S, Peers J (1994) Effect of neotraditional neighborhood design on travel characteristics. Transportation Research Record 1466:63–70
Fulton W (1996) The new urbanism. Lincoln Institute of Land Policy, Cambridge, MA
Gierl MJ, Mulvenon S (1995) Evaluating the application of fit indices to structural equation models in education research: A review of the literature from 1990 through 1994. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA, April
Golob TF, Horowitz AD, Wachs M (1979) Attitude-behavior relationships in travel demand modeling. In: Hensher DA, Stopher PR (eds) Behavioral travel modelling. Croom-Helm, London pp 739–757
Horowitz JL (1995) Modeling choices of residential location and mode of travel to work. In: Hanson S (ed) The geography of urban transportation, 2nd, Ed. Guilford Press, New York pp 219–239
Jack Faucett Associates and Sierra Research, Inc. (1999) Granting air quality credit for land use measures: Policy options. Report No. SR99-09-01, prepared for the US Environmental Protection Agency under Contract No. 68-C7-0051, Work Assignment No. 1-04, September
Kitamura R, Mokhtarian PL, Laidet L (1997) A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. Transportation 24(2):125–158
Levine J (1999) Access to choice. Access (magazine of the University of California Transportation Center) 14:16–19 (Spring)
Mauch M, Taylor BD (1997) Gender, race, and travel behavior: Analysis of household-serving travel and commuting in San Francisco Bay Area. Transportation Research Record 1607:147–153
Micceri T (1989) The unicorn, the normal curve, and other improbable creatures. Psychological Bulletin 105:156–166
Reibstein DJ, Lovelock CH, Dobson R de P (1980) The direction of causality between perceptions, affect, and behavior: An application to travel behavior. Journal of Consumer Research 6:370–376
Rummel RJ (1970) Applied factor analysis. Northwestern University Press, Evanston, IL
Shiftan Y, Suhrbier J (2000) The effects of land use policies on regional travel using a residential

choice stated-preference model and an activity-based travel model. Paper prepared for the 9th Meeting of the International Association for Travel Behaviour Research, Gold Coast, Australia, July

Southworth M (1997) Walkable suburbs? An evaluation of neotraditional communities at the urban edge. Journal of the American Planning Association 63(1):28–44

Tong CO, Wong SC (1997) The advantages of a high density, mixed land use, linear urban development. Transportation 24:295–307

United States Environmental Protection Agency (2000) Recognizing the air quality benefits of local and state land use policies and projects in the air quality planning process. Public Comment Draft of Report EPA420-P-00-002, June. Office of Air and Radiation and Office of Transportation and Air Quality, US EPA

West SG, Finch JF, Curran PJ (1995) Structural equation models with nonnormal variables: Problems and remedies. In: Hoyle R (ed) Structural equation modeling: Concepts, issues, and applications. Sage Publications, Thousand Oaks, CA pp 56–75