Model for Children's School Travel Mode Choice

Accounting for Effects of Spatial and Social Interaction

Raghuprasad Sidharthan, Chandra R. Bhat, Ram M. Pendyala, and Konstadinos G. Goulias

Numerous programs aimed at enhancing the choice of bicycling and walking as modes of choice for children's trips to and from school are being implemented by public agencies around the world. Disaggregate models that can account for the myriad of factors that influence the school mode choice of children are needed to forecast the potential impacts of such programs and policies. This paper presents a model for school mode choice that can capture the unobserved spatial interaction effects that may influence household decision making in choosing a mode of transportation for children's trips to and from school. For example, households that are geographically close together in a neighborhood may interact or observe one another and be influenced by each other's actions. To overcome the computational intractability associated with estimating a discrete choice model with spatial interaction effects, the paper proposes a maximum approximated composite marginal likelihood approach for estimating model parameters. The model is applied to a sample of children in Southern California whose households responded to the 2009 National Household Travel Survey in the United States. Spatial correlation effects are statistically significant, and they arise from interactions among households that are geographically close to one another. The findings suggest that public policy programs aimed at enhancing the use of bicycle and walk modes may have a greater impact if directed toward the local neighborhood level as opposed to a more diffuse regional level.

Much attention is being paid to the analysis of factors contributing to the travel mode choice behavior of children for the trip to and from school (*1*). Major programs aimed at promoting walking and bicycling to school are in place, particularly in the United States, where a steady decline in the shares of walk and bicycle modes for school trips has been observed over the past few decades (*2–4*). Examples of these programs include the U.S. Department of Transportation Safe Routes to School program (http://www.saferoutesinfo.org) and the Walking School Bus initiative (http://www.walkingschoolbus.org). Much of this interest stems from a desire to promote active transportation mode use among children with a view that the choice of such modes would substantially help fight childhood obesity, which has become a serious public health concern in the United States and elsewhere (*5*). Several studies have shown that children who use active modes of transportation for the trip to and from school are likely to be more physically active during other periods of the day as well, thus increasing their overall physical and mental well-being (*6, 7*).

Many factors affect the choice of mode for children's trips to and from school. Studies of children's school mode choice show the important effects of home–school proximity, household socioeconomic attributes, neighborhood built environment characteristics, and parental or caregiver perceptions of neighborhood safety and vehicular traffic conditions on the path to and from school. A systematic review of the literature on this topic is provided by Pont et al. (*8*); some pertinent literature on this topic is reviewed in more detail in the next section of this paper.

The literature review reveals that many studies loosely acknowledge, but largely ignore or do not adequately account for, the spatial interaction effects that influence children's mode choice to and from school. Spatial interaction may occur in two ways: across spatial units (zones, neighborhoods, tracts, blocks), because units that are closer to one another share some common unobserved attributes, and across behavioral units (individuals, households), because behavioral units that are closer to one another in space may share common unobserved attributes that affect the way they behave. In the context of children's school mode choice, a household's decision related to a child's school trips may be influenced by the actions and choices of other households and individuals in the same spatial cluster (e.g., a neighborhood). For example, if many other children in the neighborhood walk to school, parents may feel comfortable with their own child walking to school. The Walking School Bus initiative was founded on this principle of social interaction effects among households in close proximity of one another.

Spatial interaction among individuals may arise in the context of children's mode choice to school in a number of ways. Similarly, social interactions between parents in a neighborhood or whose children attend the same school could lead to an exchange of information about characteristics of different modes, thus contributing to a dependence in the mode utility functions of different individuals. Another possible way such a correlation can arise involves other children in the same neighborhood using an active mode of transportation and creating a positive environment for the use of such modes by improving the safety of walking and bicycling in the neighborhood; this might persuade other children and their caregivers to adopt nonmotorized modes of transportation for the trip to and from school.

R. Sidharthan and C. R. Bhat, Department of Civil, Architectural, and Environmental Engineering, University of Texas at Austin, 1 University Station C1761, Austin TX 78712-0278. R. M. Pendyala, School of Sustainable Engineering and the Built Environment, Arizona State University, Room ECG252, Tempe, AZ 85287-5306. K. G. Goulias, Department of Geography, University of California, Santa Barbara, CA 93106-4060. Corresponding author: C. R. Bhat, bhat@mail.utexas.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2213, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 78–86.
DOI: 10.3141/2213-11

Finally, similarities in the built environment attributes across households and individuals who are located in greater proximity of one another may also create interactions in the modal utilities of individuals.

Previous attempts to study school mode choice for children have not accounted for such spatial and social interaction effects, although some attempts have been made to consider spatial attributes in mode choice decisions (*9*). Accounting for such effects requires methodologic advancements in the specification and estimation of discrete choice models; this paper presents a methodologic framework and estimation approach that makes it possible to estimate mode choice models with spatial and social interaction effects. Another major impediment to the development of mode choice models that account for spatial effects is that detailed spatial accessibility measures at small levels of geography are generally not available in most travel survey data sets. In this study, disaggregate census tract–level spatial accessibility measures are computed based on Chen et al. (*10*) for a survey sample drawn from the 2009 U.S. National Household Travel Survey (NHTS) and used to disentangle unobserved spatial correlation effects from observable built environment attributes associated with household location.

An overview of the literature is offered in the next section. The third section presents the modeling methodology adopted in this paper. The fourth section provides a description of the data set, and the fifth section summarizes model estimation results and study findings. The final section offers concluding thoughts and directions for further research.

ANALYSIS OF CHILDREN'S SCHOOL MODE CHOICE

There has been considerable research aimed at studying children's school trip mode choice behavior. Pont et al. (*8*) provided a systematic review of the literature on this topic and more broadly on the topic of active transportation among children. Studies on children's school mode choice span the globe, as it is clearly an issue of interest in metropolitan contexts around the world. In the United States, an analysis by McDonald (*2*) of the series of national travel surveys from 1969 through 2001 shows the substantial decline in active mode use over the past several decades. In 1969, about 41% of students bicycled or walked to school; by 2001, that proportion had decreased to about 13%. McDonald (*2*) indicated that the increased distance between home and school may account for about one-half of the decline in the use of active transportation modes to school.

Distance between home and school is a critical factor affecting the use of nonmotorized modes (*11*). Ewing et al. (*12*) analyzed data from Gainesville, Florida, and found distance to be one of the most important factors in the choice of bicycle and walking modes. Yeung et al. (*13*) reported a similar result in an analysis of data from Brisbane, Australia. However, unlike the U.S. study, they did not find a significant difference in the body mass index of children using active modes of transport versus those using motorized modes for travel to and from school. Loucaides and Jago (*7*), analyzing data from Cyprus, found that overweight children who walked to school were more physically active in general than overweight children who were driven to school. However, no such difference was observed in children of normal weight. Cooper et al. (*6*) analyzed a sample from Bristol, United Kingdom, and reported that boys who walk to school are likely to be more physically active in general after school than those who use motorized modes of transport. Such differences were not found among girls.

Several studies are dedicated to analyzing the influence of the built environment attributes and street configuration on school mode choice. The results are somewhat mixed possibly because of the difficulty in measuring built environment attributes and appending such variables to individual person and household survey records. For example, Yarlagadda and Srinivasan (*14*) found strong effects of socioeconomic attributes and distance but reported that the impacts of travel time and built environment attributes are statistically insignificant. Similarly, McMillan (*3*) reported that urban form variables had a modest impact on mode choice; these variables had a relatively smaller impact than other variables representing socioeconomic attributes, distance, and vehicular traffic conditions. On the other hand, Boarnet et al. (*15*), in analyzing the impact of the Safe Routes to School program, found that sidewalk improvements, crossing improvements, and traffic control enhancements improved the odds of children switching to walking and bicycling modes. Ewing et al. (*12*) also noted that street density and sidewalk connectivity are influential in facilitating walking to school.

Traffic safety and parental perceptions of crime against children (e.g., abduction, molestation) were found to be significant in a few studies. Timperio et al. (*16*), in an analysis of data from Melbourne, Australia, found that parental perceptions of the number of children walking to school in the neighborhood, the presence of lights and adequate crossings, and the presence of a busy roadway between the home and school affected mode choice. DiGuiseppi et al. (*17*), in a study of data from the United Kingdom, found that adults accompanied 84% of children to and from school. Only 3% of bicycle users were allowed to bicycle on main roads. Ninety percent of parents were very or quite worried about abduction or molestation and an almost identical percentage were very or quite worried about traffic safety. Parental concerns about safety were strong predictors of school mode choice.

Some studies have identified a few other factors influencing school mode choice. Weather conditions were cited as an important explanatory variable by Müller et al. (*18*) in a study conducted in Germany, while psychological and attitudinal factors were found to be significant by Black et al. (*4*), who reported on a study conducted using data gathered from 51 schools in the United Kingdom. Zwerts et al. (*19*), in a study of Belgian students, found that students viewed the walking and bicycling experience en route to school as an important factor in the attractiveness of those modes. Dellinger and Staunton (*20*) analyzed data from the U.S. National Health Survey (conducted by the Centers for Disease Control and Prevention). They reported that barriers to walking and bicycling among children were long distances, traffic danger, and adverse weather conditions. They found that 85% of those who reported no barriers ended up using nonmotorized modes of transportation.

The role of parental influence, intrahousehold interactions, and social networks is further brought out in other studies. For example, the study by Yarlagadda and Srinivasan (*14*) focused on the escort person for the school trip. They reported that the presence of multiple school-going children in a household increases the odds that the mother will drive the children to school. This finding is in contrast to that reported by McDonald (*21*), who noted that having siblings increases the likelihood of walking and reduces the likelihood of being driven. These findings point to the need to study further the role of intrahousehold interactions in school mode choice behavior. McMillan et al. (*22*) found that the odds of biking or walking to school were 40% lower in girls than in boys but noted that the relationship was moderated by the caregiver's walking propensity and behavior. Pooley et al. (*23*) examined Global Positioning System traces of school journeys of children in the United Kingdom and found great variability in the characteristics of school travel. They attributed this variability to complex household interactions, family responsibilities, personal commitments, and personal preferences. Zwerts et al. (*19*) noted that the social aspect associated with walking or bicycling together was very important, particularly for girls.

From the review of the literature, it is clear that several factors influence school mode choice for children. While some results are mixed, home-to-school distance (proximity), socioeconomic characteristics, built environment attributes, street configuration, land use density and mix, and attitudes and perceptions of safety and crime are important determinants of school mode choice behavior. Although these studies acknowledge the potential importance of interactions within and outside the household arising from neighborhood effects, and some studies attribute certain results to intrahousehold interactions and neighborhood social networks, the studies do not explicitly account for interaction and social network effects in the modeling of school mode choice. Mitra et al. (*9*) analyzed data from Toronto, Canada, and used spatial autocorrelation measures to identify zones with high walking rates. However, their study did not involve estimating a mode choice model in the presence of spatial interaction effects. Ulfarsson and Shankar (*24*) also attempted to capture correlation effects, but their model specification focused on accounting for correlations across alternatives using a covariance heterogeneity specification (as opposed to capturing interaction effects across behavioral units over space).

This paper aims to fill a critical gap in the study of children's school mode choice behavior by developing a model that accounts for spatial and social effects arising from interactions among household members and across households in geographic and social clusters, respectively.

MODELING METHODOLOGY

Spatial interaction effects may exist across discrete choice alternatives (*25, 26*) or across decision makers (*27, 28*). This paper focuses on spatial and social interaction across decision makers. In the context of spatial interaction across decision makers, earlier studies have focused on binary response models or ordered response models. In particular, spatial interaction across individuals has seldom been discussed in the context of unordered-response models. However, spatial interaction in data may occur in unordered-response models for the same reasons (e.g., diffusion effects, social spillover effects, and unobserved location-related effects) that these effects have been studied extensively in binary and ordered-response models.

In terms of estimating binary and ordered-response discrete choice models with a general spatial structure, the analyst confronts, in the familiar probit model, a multidimensional integral over a multivariate normal distribution, which is of the order of the number of observational units in the data. While a number of approaches have been proposed to tackle this enormous multidimensional integration problem (*29, 30*), none of these methods is practically feasible for moderate-to-large samples as they are quite cumbersome from a computational standpoint. In the context of unordered-response models, the situation becomes even more difficult: the likelihood function entails a multidimensional integral over a multivariate normal distribution of the order of the number of observational units factored up by the number of alternatives minus one. This situation, however, is relatively easily handled by using the maximum approximated composite marginal likelihood (MACML) estimation method proposed by Bhat (*31*).

Model Formulation

Consider a spatial lag model structure for unordered-response models as proposed by Bhat (*31*), where the dependencies in modal utilities across individuals is caused by a combination of direct spillover effects (utilities of individuals rubbing off on each other) and indirect unobserved spatial–social effects. In such a model structure, the utility (U) that an individual q associates with alternative i ($i = 1, 2, \dots, I$) is assumed to take the following form:

$$U_{qi} = \rho \sum_{q'} w_{qq'} U_{q'i} + b' x_{qi} + \xi_{qi} \qquad \xi_{qi} \sim N(0, 0.5), \; |\rho| < 1 \tag{1}$$

where

ρ = spatial correlation parameter,
x_{qi} = $(K \times 1)$-column vector of exogenous attributes,
b = $(K \times 1)$-column vector of corresponding coefficients,
$w_{qq'}$ = spatial weight corresponding to individuals q and q', with $w_{qq} = 0$, and
$\sum_{q'} w_{qq'}$ = 1 for each (and all) q.

It is also assumed that ξ_{qi} is independent and identically distributed across q and i. The above utility function may be equivalently written as follows:

$$U_{qi} = \left[\left(\text{iden}_Q - \rho W \right)^{-1} x_i b \right]_q + \left[\left(\text{iden}_Q - \rho W \right)^{-1} \xi_i \right]_q \tag{2}$$

where

iden_Q = identity matrix of size Q,
W = spatial weight matrix of size $Q \times Q$,
x_i = $Q \times K$ vector $(x'_{1i}, x'_{2i}, \dots, x'_{Qi})$,
ξ_i = vertically stacked vector of ξ_{qi} terms of size $Q \times 1$, and
$[.]_q$ = qth element of column vector [.].

By substituting $V_{qi} = [(\text{iden}_Q - \rho W)^{-1} x_i b]_q$ and $\epsilon_{qi} = [(\text{iden}_Q - \rho W)^{-1} \xi_i]$, Equation 2 may be written equivalently as follows:

$$U_{qi} = V_{qi} + \epsilon_{qi} \tag{3}$$

where $\text{Var}(\epsilon_i) = \tilde{A} = 0.5(\text{iden}_Q - \rho W)^{-1} (\text{iden}_Q - \rho W')^{-1}$, and ϵ_i = vertically stacked vector of the $Q \epsilon_{qi}$ error terms. Define $H_{qim_q} = V_{qi} - V_{qm_q}$, where m_q is the alternative chosen by individual q.

Then, the latent utility differentials, $y^*_{qim_q}$ $(= U_{qi} - U_{qm_q}, i \neq m_q)$, may be written as $y^*_{qim_q} = H_{qim_q} + (\epsilon_{qi} - \epsilon_{qm_q})$, $i \neq m_q$. Let $y^*_q = (y^*_{q1m_q}, y^*_{q2m_q}, \dots, y^*_{qIm_q}; i \neq m_q)'$, and let $y^* = (y_1^{*\prime}, y_2^{*\prime}, \dots, y_Q^{*\prime})'$. Thus y^* is an $(I - 1) * Q$ vector. Also, let $H_q = (H_{q1m_q}, H_{q2m_q}, \dots, H_{qIm_q}; i \neq m_q)$, which is an $(I - 1) \times K$ matrix. The likelihood of the observed sample (i.e., Individual 1 choosing alternative m_1, Individual 2 choosing alternative $m_2, \dots$, Individual Q choosing alternative m_Q) may then be written succinctly as $\text{Prob}[y^* < 0]$. To write this likelihood function, note that the mean vector of y^* is $B = [(H_1)', (H_2)', \dots, (H_Q)']'$.

Then one can write $y^* \sim \text{MVN}(B, \Sigma)$ (MVN = multivariate normal) and the likelihood function of the sample is

$$L_{ML}(b, \rho) = \text{Prob}(y^* < 0) = F_{(I-1)*Q}(-B, \Sigma) \tag{4}$$

where Σ is covariance matrix of y^* and Bhat (*31*) provided the equations for calculating it. $F_{(I-1)*Q}$ is the multivariate cumulative normal distribution of $(I-1)*Q$ dimensions. Of course, maximizing the above likelihood function requires the evaluation of an $(I-1)*Q$ integral. Integrals of this high dimensionality are clearly impractical to evaluate with the usual Monte Carlo simulation methods. However, the MACML estimation approach recently proposed by Bhat (*31*) can be used here. The MACML method is briefly described in the next section.

Maximum Approximated Composite Likelihood Approach

In contrast to approaches based on evaluating the multidimensional integrals in the true likelihood function using simulation techniques, the MACML estimation approach for cross-sectional unordered-response models with normally distributed mixing is based on analytic approximations to the multivariate normal cumulative distribution functions in the true likelihood function. The approximation adopted by Bhat (*31*) relies only on bivariate and univariate standard normal cumulative distribution function computations and is computationally efficient. The approximation is combined with the composite marginal likelihood (CML) estimation approach for estimation of unordered-response models with normally distributed mixing. The MACML approach can be applied with simple optimization software for likelihood estimation. It also represents a conceptually simpler alternative to simulation techniques and has the advantage of reproducibility of the results. The covariance matrix of the MACML estimator may be easily computed by using the inverse of Godambe's (*32*) sandwich information matrix [see Bhat (*31*) for complete details].

In the MACML estimation approach, a combination of the CML method and the approximation method for multivariate normal orthant probabilities is used. The pairwise CML function for the sample is given by the following expression:

$$L_{\text{CML}}(b,\rho)=\prod_{q-1}^{Q-1}\prod_{q'=q+1}^{Q}\text{Prob}\left(C_q=m_q,C_{q'}=m_{q'}\right)$$
$$=\prod_{q-1}^{Q-1}\prod_{q'=q+1}^{Q}\text{Prob}\Big[y^*_{qim_q}<0\,\forall i\neq m_q \text{ and } y^*_{qim_{q'}}<0\,\forall i\neq m_{q'}\Big] \qquad (5)$$

where individual q's choice is denoted by C_q and m_q is the observed choice of individual q.

Each multivariate orthant probability above has a dimension equal to $(I-1)\times 2$, which can be computed by using the approximation proposed by Bhat (*31*) in the MACML approach. The variances and correlations in the bivariate and univariate cumulative normal distribution expressions in the approximation can be obtained as appropriate submatrices of Σ. An issue that has a direct impact on computational time in the CML approach is the number of pairs $[Q(Q-1)/2$ pairs$]$ of $(I-1)\times 2$ multivariate probability computations.

The framework discussed above is extendable to include social and other forms of dependence as well. This is because the weight matrix W that forms the basis for spatial dependence can also be the basis for more general forms of dependence. In fact, W can be parameterized as a finite mixture of several weight matrices [as in Yang and Allenby's (*33*) application to the simple binary choice model], with each weight matrix being related to a specific covariate k—that is, $W=\sum_{k=1}^{K}\varphi_k W_k$, where φ_k is the weight on the kth covariate in determining dependency between individuals ($\sum_{k=1}^{K}\varphi_k=1$), and W_k is a measure of distance between individuals on the kth covariate.

DATA

The data used in this study were derived from the California add-on sample of the 2009 NHTS conducted from 2008 through 2009 in the United States. Within the California add-on sample, the survey subsample of respondents from the Los Angeles–Riverside–Orange County region was extracted and used for the model estimation effort. This selection process was done for several reasons. First, the use of a national sample for studying school mode choice behavior may be inappropriate given that there are likely to be substantive geographic differences across the country. Spatial correlation effects are likely to be more localized in nature, calling for the use of data drawn from a more limited geographic region for analysis and model development. Second, the use of a very large sample for model estimation would produce inflated test statistics, which would affect inferences drawn from the model results. Finally, the authors have access to census tract–level accessibility measures and land use data for the Los Angeles region in conjunction with an ongoing activity-based model development effort under way for the Southern California Association of Governments.

The survey collects detailed socioeconomic, demographic, and travel information for all household members in respondent households. The survey also collects information about usual travel characteristics by asking questions about travel undertaken in the past week. Extensive descriptive statistical analysis was conducted on the data to understand mode choice patterns for children's school trips and to identify explanatory factors that may influence such behavior. For the sake of brevity, all the analyses conducted are not described and presented here, but some highlights are noted to provide an overview of the data assembly process.

The survey sample included 1,192 children of 5 to 15 years of age for whom school mode choice behavior could be analyzed. Table 1 presents the average travel time to school, the average travel time by mode used, the median household income, the median household income by mode used, and descriptive statistics of other household characteristics to which these children belong. In general, the travel time to school ranges from 10 to 15 min, with an average of 12.4 minutes. Only the average bus travel time falls outside this range, with an average value of just over 25 min. Those who walk and use the school bus report lower median household incomes than other groups. Thus, it is clear that mode choice to and from school is correlated to income; perhaps lower car ownership in these households leads children to walk and ride the school bus. In general, the household characteristics show that households are larger than would be expected if one were analyzing the general population. This is consistent with the fact that the analysis sample here focuses on households with children going to school.

The importance of distance in school mode choice behavior has been highlighted in previous research. Table 2 presents modal split distributions by home-to-school distance bands. The association between home-to-school distance and modal split is readily apparent. While the overall mode split for car is 44%, the highest among all modes, it is clear that walking is the predominant choice of mode at very short distances. For distances less than ¼ mi, 60% of children walk to school and fewer than 25% ride to school in a car. However, 13% of children use a combination of car and walk (i.e., they

TABLE 1 Sample Demographic Characteristics

Characteristic	Value
Average travel time to school (min)	12.4
Average travel time to school by modal market segment (min)	
Car	10.9
School bus	25.8
Bicycle	14.0
Walk	12.1
Car–school bus	16.7
Car–walk	9.7
Median household income ($)	70,800
Median household income by modal market segment ($)	
Car	78,000
School bus	50,300
Bicycle	73,600
Walk	54,700
Car–school bus	82,700
Car–walk	68,400
Number of household members	4.3
Number of vehicles in household	2.4
Number of bicycle trips in past week	1.3
Number of walk trips in past week	4.0
Number of adults in household	2.3
Number of workers in household	1.5

take the car to school but walk home after school). There is a dramatic increase in car mode share as distance increases; the car mode share nearly doubles to 46% at distances more than ¼ mi but less than ½ mi. The car mode share continues to increase with distance and reaches nearly 75% at home-to-school distances in excess of 2 mi. The school bus mode share also increases with distance, consistent with expectations. The bicycle mode share shows some fluctuations, with higher shares observed for very short trips less than ¼ mi and midrange distances of ½ to 2 mi. The car–school bus combination shows a significant modal percent (6%) at longer distances, again consistent with expectations. Walk mode share dramatically drops off with increasing distances, with just about a 1% mode share for school trip distances greater than 2 mi. One factor affecting the choice of active modes of transportation is that nearly 40% of the children live more than 2 mi from their school. Only about 25% of the children live within ½ mi of their school location. As schools grow increasingly larger and cover larger boundary areas, this challenge may become more pronounced.

An analysis of the data showed that some children use a combination of modes to commute to and from school. In a cross-classification table of modes to and from school (not presented here due to space considerations), the diagonal elements of the table show the largest figures, as expected, signifying that a vast majority use the same mode to and from school. Of the 1,192 children, 1,041 (87%) use the same mode to and from school. More than 50% of the children use a car in both directions, and close to 20% walk in both directions. Of the modal transition segments, the largest one (with 71 students) involves using a car to go to school and walking home from school. Other modal transitions are rather small, although the walk–car and car–school bus segments cannot be ignored.

In preparing the final data set for model estimation, modes with very few observations were eliminated. They included "other," "school bus + walk," and "bicycle + car," which left 1,143 students in the sample. After further cleaning the data set, removing observations with missing information and clearly miscoded values and other reductions, 800 observations were retained for estimation.

In the survey, the walk travel time was reported for those children who walked to school. In addition, the distance between home and school was obtained for all the children in five distance bands (see Table 2). In examining the walk travel times and the distances to school for children who walked, a good bit of variation was found in walk times within the sample of children who were in the same distance band. So, it was decided, from an econometric efficiency perspective, to consider both travel time and travel distance in the specification. In doing so, the walk time to school was imputed for those children who did not walk to school by computing the mean walk travel time for children who did walk to school in the corresponding distance band. However, as reported later, walk travel distance did not turn out to be significant after controlling for walk travel time. For other modes, imputation procedures were similarly developed to construct travel time values for all individuals (whether or not they used the mode) and both travel times and distances (in the five distance bands) were considered. For all the nonwalk modes, the distance variable specification turned out to be better, presumably because of rounding and inaccuracy in trip time reporting for these relatively long trips.

As mentioned earlier, there may be household interactions that affect choice of mode for school trips. The bicycling and walking activity of adults in each household is reported in the survey as the number of bicycling and walking trips undertaken for various purposes in the previous week. For this study, adults (parents) were classified as active bicyclists or walkers if they made at least five trips using the corresponding mode in the previous week, with at least one

TABLE 2 School Mode Choice Distribution by Distance from Home to School

	Distance to School					
	<¼ mi	¼–½ mi	½–1 mi	1–2 mi	>2 mi	Total
Mode						
Car (%)	23.9	46.2	56.8	68.5	74.7	44.3
School bus (%)	NA	2.5	1.2	7.6	15.7	33.8
Bicycle (%)	2.5	1.7	4.9	2.0	0.5	1.2
Walk (%)	60.4	37.0	22.2	10.4	1.1	7.0
Car–school bus (%)	NA	NA	0.6	0.8	6.4	10.1
Car–walk (%)	13.2	12.6	14.2	10.8	1.6	3.5
Total children	159	119	162	251	439	1,130
Percentage by distance	14.1	10.5	14.3	22.2	38.8	100

NOTE: NA = not applicable.

trip being made for a purpose other than to escort children to and from home. In other words, if the sole reason an adult made bicycling or walking trips in the past week is to escort children, then the person is not considered an active bicyclist or walker (to avoid potential endogeneity problems).

The NHTS data set includes a set of attitudinal variables that capture individual attitudes and perceptions. In particular, the survey asks parents to rate a series of issues on a 5-point scale with a value of 1 meaning that the consideration is not an issue and 5 meaning that the consideration is a serious issue. Adults were asked to identify the extent to which each of the following considerations affected the decision to allow their child (or children) to walk or bicycle between home and school: distance between home and school, amount of traffic along the route, speed of traffic along the route, violence and crime along the route, and poor weather or climate in the area. A principal component factor analysis (without rotation) was undertaken to reduce these five attitudinal variables into a set of orthogonal factors. The factor analysis yielded two factors, one corresponding to objectively measurable attributes such as distance and speed and volume of traffic, and the other corresponding to more subjective measures of crime and weather. These factors were used in the model specification to capture the effects of parental attitudes on school mode choice.

MODEL ESTIMATION RESULTS

A simple probit model that does not account for spatial–social interaction effects and the spatial interaction model were estimated, and the results are presented in Table 3. A systematic procedure in which variables were entered in a stepwise manner and checked for their statistical significance and intuitive behavioral interpretation was followed to arrive at the final model specification. Various forms of explanatory variables and interaction effects among them were tested to arrive at the best possible model specification that is parsimonious and yet sensitive to a range of effects that one would expect to see in a mode choice model of the type developed here.

An examination of the alternative specific constants shows that, in general, the bicycle and car + walk combination modes are less preferred than other modes (although the constants also control for the range of exogenous variable values in the sample). It is also found that there are substantial differences in the alternative specific constants between the probit model with no spatial–social effects and the spatial interaction model. This is a first indication that ignoring spatial interaction effects when they are present results in inaccurate estimates of preferences for alternative modes. With respect to travel characteristics, findings are largely consistent with expectations. As the time to walk increases, the utility of walking decreases. For distances less than 2 mi, the utility of the school bus decreases; presumably, the bus is of greater value when distances to school are more than 2 mi. On the other hand, the utility of bicycle and car + walk combination modes is higher for distances within this range.

Age and gender of the student are statistically significantly associated with school mode choice. The utility of bicycling, walking, and using a combination of car and walking increases with the age of the child. In other words, older children are more likely to use nonmotorized modes of transportation than younger children, presumably because parents feel more comfortable letting older children use these modes. The coefficient associated with age is substantially higher for the bicycle mode than for the walk modes, suggesting that the utility for bicycle increases more than for walk modes with increasing age. A gender effect is apparent, with females less likely to choose

TABLE 3 Model Estimation Results

			Independent Probit		Spatial Model	
Variable Category	Variable	Mode Utility Equation	Coeff.	Est. SE	Coeff.	Est. SE
Alternative specific constant	School bus		0.393	1.864	1.413	4.819
	Bicycle		−3.721	−4.724	−1.942	−2.845
	Walk		0.570	2.112	1.420	4.307
	Car + school bus		−0.858	−5.424	1.133	0.869
	Car + walk		−2.000	−6.814	−1.103	−3.837
Trip characteristics	Time to walk	Walk	−0.061	−11.522	−0.063	−10.599
	Distance to school < 2 miles	School bus	−0.542	−3.740	−0.521	−3.635
	Distance to school < 2 miles	Bicycle, car + walk	1.099	6.458	1.123	5.908
	Distance to school < 2 miles	Car + school bus	−1.629	−0.285	−1.725	−0.724
Individual demographics	Age	Bicycle	0.157	2.537	0.165	2.575
	Age	Walk, car + walk	0.056	3.236	0.058	3.298
	Female child	Bicycle	−0.920	−2.164	−0.943	−2.519
Household demographics	Household income	School bus	−0.107	−5.270	−0.103	−5.244
	Household income	Walk	−0.065	−3.989	−0.063	−3.971
	Vehicles per capita in household	Car	0.440	2.545	0.447	2.514
	Adult nonworker present in household	Walk	0.301	2.186	0.318	2.194
	Adult nonworker present and no. cars > no. workers	Car	0.308	3.243	0.313	3.214
Parents' attitude	Attitude toward walk or bicycle mode	Walk	−0.102	−1.919	−0.108	−1.953
Accessibility of neighborhood	Total amount of retail employment that can be reached in 10 min	School bus	−0.055	−2.049	−0.044	−2.160
Spatial interaction parameter	ρ		NA	NA	0.844	6.447
	CML		−584,880.8		−580,600.8	

NOTE: Coeff. = coefficient; est. SE = estimated standard error; no. = number; NA = not applicable.

a bicycle than their male counterparts, a finding previously reported by McMillan et al. (*22*).

With respect to household demographics, higher household income and vehicle ownership is associated with a greater propensity to use a car and a lower utility for school bus and walking. This is consistent with previous research, which also reports that households with higher levels of vehicle ownership are less likely to depend on alternative modes for transporting children to and from school (*12, 21*). The presence of adult nonworkers in the household positively affects use of the walk mode, perhaps because the adult nonworker can accompany the child on the walk to and from school (alleviating safety concerns associated with having the child walk alone). However, when there are one or more adult nonworkers in the household with a spare automobile, then the utility of a car increases. The parental attitude is captured through the attitudinal factor that measures whether the parents considered distance and traffic conditions to be issues associated with having their child(ren) commute by walk or bicycle. If the attitudinal factor value increases, then it means the parents considered the issue to be more serious. As expected, in households where parents had issues with distance and traffic conditions, the utility of walking to and from school decreases. The subjective attitudinal factor (capturing weather and safety concerns) was not statistically significant. A physically active parent—an active bicyclist or an active walker—increases the probability of a child using the corresponding mode. However the relationship appears to be weak and the coefficients were insignificant at the 0.05 level of significance. So these parameters are not included in the final results presented in this paper.

Spatial factors play an important role in determining school mode choice. The accessibility of the neighborhood is measured by the total amount of retail employment that can be reached within a 10-min radius of the home location. These accessibility measures were computed at the tract level using block-level data about employment in different industry sectors obtained from the Southern California Association of Governments. In general, it is found that a higher level of neighborhood accessibility (measured in terms of retail employment) has a negative association with the school bus mode utility. It is possible that these households are in higher-density areas more conducive to walking and bicycling, or there are busy streets that motivate use of the car. This finding is consistent with that reported by Ulfarsson and Shankar (*24*), Yarlagadda and Srinivasan (*14*), and Ewing et al. (*12*).

Spatial interaction effects were tested by specifying the weight matrix using both geographic proximity and demographic closeness as potential measures of the correlation. For geographic proximity, alternative specifications of distance (e.g., inverse of distance between individuals, inverse of exponentiated distance) and membership in a county ($w_{ij} = 1$ if i and j belong to the same county; $w_{ij} = 0$ otherwise) were used. The distance between individuals was obtained as the distance separation between the centroids of the tracts of the household locations of individuals. For demographic closeness, alternative specifications of income and age similarity were created by using demographic distance measures. For each of these specifications, parameters were estimated independently with the MACML approach described here. The social interaction effects turned out to be statistically insignificant in all demographic distance-based weight matrix specifications. The spatial interaction parameter was significant (and positive) for all geographic distance-based weight matrix specifications, and the best CML was obtained for the specification using the inverse of distance as the spatial proximity measure.

The spatial correlation parameter ρ is positive, high in magnitude (0.844), and statistically significant; the attributes indicate a high degree of geographic interdependence in the choice of mode of travel to school. The spatial lag model seems more appropriate than the nonspatial independent multinomial probit (IMNP) model. Another way to demonstrate this is to use the adjusted composite maximum likelihood ratio test statistic, which follows a chi-squared distribution (*31, 34*). This statistic returned a value of 17.2 for comparing the spatial lag model with the IMNP model, which is higher than the corresponding chi-squared table value with 1 degree of freedom at any reasonable level of significance. However, and very importantly, the difference between the IMNP model and the spatial lag model is not simply a matter of data fit. The effects of a change in variable on aggregate mode shares will be quite different between the two models, because the IMNP model ignores interdependence, while the spatial lag model accommodates spillover effects due to interactions between decision agents and so may lead to relatively large changes in aggregate mode behavior despite small changes in the underlying primitives (or determinants) of the behavior. To demonstrate this difference in effects between the IMNP and the spatial lag model, this study examined the effect of a 5% decrease in walk time to school (e.g., due to better siting of schools relative to residences) and the impact of a 25% decrease in the level of negativity in parental attitude (in the context of distance and traffic conditions being deterrents) toward allowing children to travel to school by walking or bicycling. The decrease in walk time is estimated to lead to a 0.29% decrease in car mode share according to the IMNP model but to a decrease in car mode share by almost 12% according to the spatial lag model. Similarly, the improvement in parental attitude toward nonmotorized modes is estimated to decrease the car mode share by just 0.48% according to the probit model but by 3.2% as per the spatial lag model. Clearly, spillover effects are at work here, and the IMNP model provides estimates that are quite different than those from the spatial lag model.

In summary, the spatial interdependence means that, for any individual, the utility of each alternative is positively (negatively) influenced by an increase (decrease) in the utility of corresponding alternatives for his or her geographic neighbors. In other words, the spatial dependence in school mode choice appears to arise more from social interaction and neighborhood location effects associated with households geographically clustered closer together. It is possible that parents of households living in a zone, tract, or neighborhood interact with one another and share experiences about the school travel of their children. Households may band together to facilitate walking and bicycling in a safe and secure way, but this interaction among households is more due to geographic proximity considerations as opposed to socioeconomic similarity considerations (although it is plausible that households living within a neighborhood are at least somewhat homogeneous with respect to socioeconomic characteristics). When other children in the neighborhood use a mode such as bicycling or walking, this creates a positive externality by improving the safety of bicycling and walking in the neighborhood, thus enhancing the utility of these modes for households in the neighborhood. As households in a geographic cluster are likely to deal with the same or similar built environment, it is not surprising that the geographic distance-based spatial interaction parameter turned out to be statistically significant.

CONCLUSIONS

This research focused on modeling school trip mode choice behavior among children (less than 15 years of age) with a view to examine the presence of spatial and social interaction effects that may affect such behavior. These effects may arise due to interactions among households that are geographically or demographically similar to one another. When such interaction effects are present, the modal utilities of individuals become dependent, thus violating the basic assumption of traditional discrete choice models that assume independence of error terms across observations. The usual maximum likelihood estimate of a model that accounts for global spatial–social effects is quite complex as one must evaluate very high dimensional integrals of a multivariate normal distribution to compute the likelihood function (the order of the integral is the number of observations multiplied by the number of alternatives minus one; in the empirical context of the current study, this translates to 4,000 dimensional integrals). In this paper, a MACML approach recently developed by Bhat (*31*) was used to estimate a school mode choice model that accounts for spatial interaction effects.

In this study, the MACML approach was applied to a sample of children in the Southern California (Los Angeles and surrounding cities) region of the United States with data collected as part of the 2009 NHTS. The survey sample includes 800 children who provided detailed mode choice information for the journey to and from school along with information about household member use of bicycle and walk modes and parental concerns about the built environment in relation to their children's use of bicycling and walking to and from school. Both an independent probit model (that does not account for spatial interaction effects) and a spatial correlation model were estimated to determine whether the spatial interaction effects are significant and present. It was found that the spatial correlation, arising from interactions among households that are geographically clustered, is statistically significant.

The findings in this paper suggest that the consideration of spatial interaction effects is important in modeling mode choice behavior, particularly in the context of children's school mode choice, where residential proximity-based interaction among households and children is likely to be prevalent. This means that programs aimed at enhancing bicycle and walk as modes of choice for the trip to and from school (such as the Safe Routes to School program in the United States) should be focused to maximize the likelihood of interactions based on geographic proximity. That is, given that spatial interaction effects fade over distance (according to the inverse distance specification for spatial weights), one can use an optimization program to define the boundaries of fixed neighborhoods to maximize interaction effects.

This paper accommodates spatial dependence due to proximity in residential locations of children and social interaction effects. An avenue for future research would be to extend the dependence effects to include proximity in school locations of children, with the notion that peer effects at school may also affect children's school mode choice. This additional effect can be accommodated in a straightforward manner in the authors' methodology by defining another weight matrix W_k that corresponds to school location proximity and considering this weight matrix as one additional finite mixture dimension affecting the overall weight matrix W ($W = \sum_{k=1}^{K} \varphi_k W_k$). However, this would require identifying the school that each child in the sample attends, with a geocoding of these school locations. This information is not available in the NHTS data used in the current analysis but may be available in other activity-travel data sets in which each activity episode location is geocoded.

ACKNOWLEDGMENTS

The authors acknowledge support from the Southern California Association of Governments for facilitating this research and are grateful to Lisa Macias for help in typesetting and formatting this document. Three anonymous reviewers provided useful comments on an earlier version of this paper. The authors acknowledge support from the Sustainable Cities Doctoral Research Initiative of the Center for Sustainable Development at the University of Texas at Austin.

REFERENCES

1. Beck, L. F., and A. I. Greenspan. Why Don't More Children Walk to School? *Journal of Safety Research,* Vol. 39, No. 5, 2008, pp. 449–452.
2. McDonald, N. C. Active Transportation to School: Trends Among U.S. Schoolchildren, 1969–2001. *American Journal of Preventive Medicine,* Vol. 32, No. 6, 2007, pp. 509–516.
3. McMillan, T. E. The Relative Influence of Urban Form on a Child's Travel Mode to School. *Transportation Research, Part A,* Vol. 41, No. 1, 2007, pp. 69–79.
4. Black, C., A. Collins, and M. Snell. Encouraging Walking: The Case of Journey-to-School Trips in Compact Urban Areas. *Urban Studies,* Vol. 38, No. 7, 2001, pp. 1121–1141.
5. Koplan, J. P., C. T. Liverman, and V. I. Kraak, eds. *Preventing Childhood Obesity: Health in the Balance.* Institute of Medicine of the National Academies, Washington, D.C., 2005.
6. Cooper, A. R., A. S. Page, L. J. Foster, and D. Qahwaji. Commuting to School: Are Children Who Walk More Physically Active? *American Journal of Preventive Medicine,* Vol. 25, No. 4, 2003, pp. 273–276.
7. Loucaides, C., and R. Jago. Differences in Physical Activity by Gender, Weight Status, and Travel Mode to School in Cypriot Children. *Preventive Medicine,* Vol. 47, No. 1, 2008, pp. 107–111.
8. Pont, K., J. Ziviani, D. Wadley, S. Bennett, and R. Abbott. Environmental Correlates of Children's Active Transportation: A Systematic Literature Review. *Health and Place,* Vol. 15, No. 3, 2009, pp. 849–862.
9. Mitra, R., R. N. Buliung, and G. E. J. Faulkner. Spatial Clustering and the Temporal Mobility of Walking School Trips in the Greater Toronto Area, Canada. *Health and Place,* Vol. 16, No. 4, 2010, pp. 646–655.
10. Chen, Y., S. Ravulaparthy, K. Deutsch, P. Dalal, S. Y. Yoon, T. Lei, K. G. Goulias, R. M. Pendyala, and C. R. Bhat. Development of Opportunity-Based Accessibility Indicators. *Transportation Research Record: Journal of the Transportation Research Board,* Transportation Research Board of the National Academies, Washington, D.C., forthcoming.
11. Lawrence Frank and Company, Inc. *Youth Travel to School: Community Design Relationships with Mode Choice, Vehicle Emissions, and Healthy Body Weight.* Final Report. U.S. Environmental Protection Agency, Washington, D.C., 2008.
12. Ewing, R., W. Schroeer, and W. Greene. School Location and Student Travel: Analysis of Factors Affecting Mode Choice. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1895,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 55–63.
13. Yeung, J., S. Wearing, and A. P. Hills. Child Transport Practices and Perceived Barriers in Active Commuting to School. *Transportation Research, Part A,* Vol. 42, No. 6, 2008, pp. 895–900.
14. Yarlagadda, A. K., and S. Srinivasan. Modeling Children's School Travel Mode and Parental Escort Decisions. *Transportation,* Vol. 35, No. 2, 2008, pp. 201–218.
15. Boarnet, M. G., C. L. Anderson, C. Day, T. E. McMillan, and M. Alfonzo. Evaluation of the California Safe Routes to School Legislation: Urban Form Changes and Children's Active Transport to School. *American Journal of Preventive Medicine,* Vol. 28, Suppl. 2, 2005, pp. S134–S140.
16. Timperio, A., K. Ball, J. Salmon, R. Robers, B. Giles-Corti, D. Simmons, L. A. Baur, and D. Crawford. Personal, Family, Social, and Environmental Correlates of Active Commuting to School. *American Journal of Preventive Medicine,* Vol. 30, No. 1, 2006, pp. 45–51.
17. DiGuiseppi, C., I. Roberts, and L. Li. Determinants of Car Travel on Daily Journeys to School: Cross Sectional Survey of Primary School Children. *British Medical Journal,* Vol. 316, No. 7142, 1998, pp. 1426–1428.
18. Müller, S., S. Tscharaktschiew, and K. Haase. Travel-to-School Mode Choice Modelling and Patterns of School Choice in Urban Areas. *Journal of Transport Geography,* Vol. 16, No. 5, 2008, pp. 342–357.
19. Zwerts, E., G. Allaert, D. Janssens, G. Wets, and F. Witlox. How Children View Their Travel Behaviour: A Case Study from Flanders (Belgium). *Journal of Transport Geography,* Vol. 18, No. 6, 2010, pp. 702–710.
20. Dellinger, A. M., and C. E. Staunton. Barriers to Children Walking and Biking to School: United States, 1999. *Journal of the American Medical Association,* Vol. 288, No. 11, 2002, pp. 1343–1344.

21. McDonald, N. C. Children's Mode Choice for School Trip: The Role of Distance and School Location in Walking to School. *Transportation,* Vol. 35, No. 1, 2008, pp. 23–35.
22. McMillan, T. E., K. Day, M. Boarnet, M. Alfonzo, and C. Anderson. Johnny Walks to School—Does Jane? Sex Differences in Children's Active Travel to School. *Children, Youth, and Environments,* Vol. 16, No. 1, 2006, pp. 75–90.
23. Pooley, C., D. Whyatt, M. Walker, G. Davies, P. Coulton, and W. Bamford. Understanding the School Journey: Integrating Data on Travel and Environment. *Environment and Planning A,* Vol. 42, No. 4, 2010, pp. 948–965.
24. Ulfarsson, G., and V. N. Shankar. Children's Travel to School: Discrete Choice Modeling of Correlated Motorized and Nonmotorized Transportation Modes Using Covariance Heterogeneity. *Environment and Planning B,* Vol. 35, No. 2, 2008, pp. 195–206.
25. Bolduc, D., B. Fortin, and M. Fournier. The Effect of Incentive Policies on the Practice Location of Doctors: A Multinomial Probit Analysis. *Journal of Labor Economics,* Vol. 14, No. 4, 1996, pp. 703–732.
26. Bhat, C. R., and J. Y. Guo. A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling. *Transportation Research, Part B,* Vol. 38, No. 2, 2004, pp. 147–168.
27. Anselin, L. Spatial Externalities, Spatial Multipliers, and Spatial Econometrics. *International Regional Science Review,* Vol. 26, No. 2, 2003, pp. 153–166.
28. Bhat, C. R., and I. N. Sener. A Copula-Based Closed-Form Binary Logit Choice Model for Accommodating Spatial Correlation Across Observational Units. *Journal of Geographical Systems,* Vol. 11, No. 3, 2009, pp. 243–272.
29. LeSage, J. P. Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models. *Geographical Analysis,* Vol. 32, No. 1, 2000, pp. 19–35.
30. Fleming, M. M. Techniques for Estimating Spatially Dependent Discrete Choice Models. In *Advances in Spatial Econometrics: Methodology, Tools and Applications* (L. Anselin, R. J. G. M. Florax, and S. J. Rey, eds.), Springer-Verlag, Berlin, 2004, pp. 145–168.
31. Bhat, C. R. *The Maximum Approximated Composite Marginal Likelihood (MACML) Estimation of Multinomial Probit-Based Unordered Response Choice Models.* Technical Paper. Department of Civil, Architectural, and Environmental Engineering, University of Texas, Austin, 2010.
32. Godambe, V. P. An Optimum Property of Regular Maximum Likelihood Estimation. *Annals of Mathematical Statistics,* Vol. 31, No. 4, 1960, pp. 1208–1211.
33. Yang, S., and G. M. Allenby. Modeling Interdependent Consumer Preferences. *Journal of Market Research,* Vol. 40, No. 3, 2003, pp. 282–294.
34. Pace, L., A. Salvan, and N. Sartori. Adjusting Composite Likelihood Ratio Statistics. *Statistica Sinica,* Vol. 21, No. 1, 2011, pp. 129–148.

The Transportation Safety Management Committee peer-reviewed this paper.