Chapter 5

The Empirical Model of Household Travel Time

In this chapter the empirical model used to test the relationship between accessibility of defined economic centers and total household travel time is developed. First, the data used in the analysis and their sources are introduced. The empirical model used in the estimation is then described.

The Data

The data can be separated into three categories. Center accessibility data are measures of the travel time to the different economic centers for each household. Household specific data is data that are descriptive of and specific to each household. Neighborhood characteristics data are that that are descriptive of the neighborhood in which the household is located. These data are each discussed below.

Center Accessibility

The goal of this research is to determine whether sprawl causes increases in total household travel time. This goal is accomplished by examining the relationship between access to economic centers identified under monocentric and limited polycentric models and household travel times. A measure of access to economic centers is therefore critical to the analysis. Accessibility was determined using travel times for the journey to work from the U.S. Department of Transportation 1990 Census Transportation Planning Package in conjunction with household specific data was drawn from the 1994 COG/TPB Household Travel Survey for the Metropolitan Washington Region of the Metropolitan Washington Council of Governments National Capital Region Transportation Planning Board (the COG/TPB Survey).[1] The COG/TBP Survey data includes geocoded household locations for each household surveyed. These were matched to travel time data from the UTPP to

---

[1] An alternative measure of accessibility made possible by GIS is the distance by road network. The primary cost of travel is thought to be time expended in travel. As such commute time is a better measure of that cost. Even though travel times and distances are closely correlated, in urban areas travel times for equivalent distances over different roads may differ significantly due to congestion.

determine travel time accessibility of each center for each household. This was accomplished for each center by averaging the travel times to all of the zones in the center from each outlying zone.[2] To obtain travel times to the centers from each household location the average travel times from each transportation analysis zone centriod were converted to a GIS surface using the inverse distance weighting function of ArcView.[3] The travel time from each household to the center was then determined using the household's location on that surface. The surface generating procedure was useful for synthesizing travel time data for zones for which the UTPP provided no data. Census data is the most comprehensive and accurate data concerning commuting times available. Census data, however, is not without errors. The surface generating procedure had the added advantage of correcting for errors in the UTPP data. Table 5.1 is a full listing of the center accessibility variables used in the analysis.

Table 5.1 Center Accessibility Data
Data from the Census Transportation Planning Package *(by transportation analysis zone)*

Variable Name	*Description*
Hcbd	travel time the central business district (in hours)
H2cbd	square of travel time the central business district
Nrst	travel time to the nearest subcenter (in hours)

Household Characteristics

Household specific data was drawn from the COG/TPB Survey.[4] The Washington Metropolitan Council of Governments/Transportation Planning Bureau collected the data

[2] Extreme outliers were removed from these averages by deleting any average travel times that required a speed in excess of 28 miles per hour from transportation analysis zone centroid to the centroid of the identified center.
[3] The function interpolates a value for each point on the surface. To do so it uses a specified number of the nearest neighboring points for which data exists. The surface is interpolated by weighting the values of each point by a specified power of the inverse of the distance of the point. After trying a variety of different numbers of neighboring points and weights, it was decided to use 15 neighboring points and an exponent of one-half. This number and weight generated a consistent surface.
[4] A number of workers have jobs that require substantial travel outside of the office. Inclusion of these workers in the analysis would likely to bias results. To avoid this bias

using single day travel diaries for each household member of each household surveyed. Table 5.2 is a complete listing of household characteristics data used in the analysis of household transportation time.

Table 5.2 Household Characteristics Data

Data from the 1994 Metropolitan Washington Council of Governments Transportation Planning Board Transportation Survey *(by household)*

Variable Name	*Description*
Lhttime	log of total household travel time
Hinc12*	household income of less than $10,000
Hinc3*	household income of $10,000-20,000
Hinc4*	household income of $20,000-30,000
Hinc5*	household income of $30,000-50,000
Hinc7*	household income of $75,000-100,000
Hinc8*	household income of $100,000-125,000
Hinc91*	household income of more than $125,000
Kids	number of children under 17
Numveh	number of household vehicles
Numlic	number of licensed drivers
Twnhs*	townhouse
Apt*	apartment
Own*	household tenure
Mv8090*	moved to home between 1980 and 1990 (inclusive)
Mvpst90*	moved to home after 1990
Wkad	number of working adults in the household
Nwkad	number of working adults in the household
Htrnp	percent of trips by transit on survey date
Md*	Maryland household
Jur0nw*	Northwest Washington, D.C. household
Jur0ne*	Northeast Washington, D.C. household
Jur0sw*	Southwest Washington, D.C. household
Jur0se*	Southeast Washington, D.C. household

* Variable is a 0/1 dummy

Neighborhood Characteristics

Demographic data was drawn from the 1990 U.S. Census Transportation Planning Package: Urban Element (UTPP) and the 1990 U.S. Census (Census). The geographical unit

---

households with one or more workers making more than one business related stop other than at work were omitted from the data set.

for the UTTP data is transportation analysis zone; for the Census data it is the block group. The variables relied on in the analysis and their respective sources are listed in Table 5.3.

Table 5.3 Neighborhood Characteristics Data

Data from the U.S. Census Transportation Planning Package: Urban Element (1990) *(by transportation analysis zone)*

Variable Name	Description
Rptu18	percent of residents under 18
Rpdrval	percent of drive alone commuters
Rpsub	percent subway commuters
Rpnwht	percent nonwhite

Data from 1990 U.S. Census *(by block group)*

Variable Name	Description
Chage	average home age
Cavrms	average rooms per dwelling

The Empirical Model

The models estimated all take the form:

$$\log(t_{ij}) = x_i\beta_x + y_j\beta_y + HCBD_i\beta_{HCBD} + HCBD_i^2\beta_{HCBD^2} + NRST_i\beta_{NRST} + u_{ij}$$

where t_{ij} is the total time spent in travel in one day for the *i*th household located in the *j*th neighborhood, x_i is a vector of household specific characteristics of the *i*th household, y_j is a vector of neighborhood specific characteristics of the *j*th neighborhood, $HCBD_i$ is the travel time in hours to the CBD from the *i*th household, $NRST_i$ is the travel time to the nearest subcenter from the *i*th household[5] and u_{ij} is a mean zero, normally distributed error term. All models are estimated using ordinary least squares estimation.

The functional form of the model has been chosen for a variety of reasons. First, its exponential form is similar to the form most commonly used for estimating both monocentric (Clark, 1951; Mills, 1972; Mills and Tan, 1980; Small and Song, 1994) and

[5] Separate models were estimated that included a variable for the travel time to each of the subcenters and that included only a variable for the travel time to the nearest subcenter. In none of the models that included distance to all subcenters were any of those variables

polycentric (Griffith, 1981; Gordon, Richardson and Wong, 1986; Small and Song, 1994; Song, 1994; Song, 1996) density gradients. The form also has intuitive appeal since multiple centers may influence all locations. That influence, however, becomes "negligible at large distances" (Anas, Arnott and Small, at 1441-2, 1998). Such a form is appealing since subcenters may be expected to exert a strong influence on households that are nearby. Households with one or more centers close by, however, would not be expected to be influenced by distant centers. The form also proved to be the most robust to tests of the misspecification.[6] Reset tests of functional form supported use of this form over linear models, double log models and translog models.

The Variables

The goal of the research is to examine whether sprawl has a travel time cost by testing whether access to economic centers has an influence on household travel time. Other factors, however, may affect household travel time. For example, the theoretical model suggests that the amount of housing services purchased by the household will affect household travel time.[7] Households purchasing more housing services are predicted to live further from economic centers to obtain a lower price for those services. Choosing a more distant location is predicted to increase the household's travel. Since household services are a function of a variety of household and neighborhood characteristics (Simpson, 1992 at 17-22), household specific and neighborhood demographic variables are included in the models

---

significant. Consequently, the results of models that include only the distance to the nearest subcenter are reported.

[6] A battery of misspecification tests testing assumptions of normality, linearity, heteroskedasticity, autocorrelation and parameter stability were performed on each model reported. Results of those tests appear in Appendix 5A.

[7] In the theory's simplest formulation those services include only the land on which the housing is situated (Alonzo, 1964). Extensions and modifications of the model have considered housing services to include both land and qualities of the home itself (Straszheim, 1987). Some extensions have considered neighborhood amenities to be housing services. To maintain the spatial nature of the model, however, requires that at least some of these amenities (such as density) be a function of the proximity of the central business district. Alternatively, hedonic specifications may be used that incorporate neighborhood amenities, however, such models sacrifice the unifying spatial structure characteristic of the monocentric models (Straszheim. 1987).

to remove the effects of those factors.[8] To the extent that the effects of these variables on the purchase of housing services that relationship is discussed below.

Household travel time may also vary directly with variation in household and neighborhood characteristics regardless of the quantity of housing services purchased. For example, household travel time is likely to increase with the number of household members regardless of residential location. To discern the influence of access to economic centers on household travel requires that the influence of these factors be removed. The inclusion of household and neighborhood demographic variables removes these direct influences and also provides some insights into the how those variables influence household travel time.

The specification of the model clearly suffers from the joint determination of household location and household travel time. The household location choice (including both the access of economic centers and the characteristics of the neighborhood of the residence) must be expected to be a function a variety of household characteristics. For example, a household with school age children may choose an affluent suburban location to both assure good quality schools and escape the perceived danger of crime in the city. If one bears in mind, however, that the goal of the research is to determine how total household travel time varies with accessibility of centers holding all other factors constant, the model specification is adequate for measuring that influence.

To help clarify the specification of the empirical model and the affects of the different variables, following is a brief discussion of each variable included in the model. That discussion provides the background necessary for understanding how the variables may be expected to influence household travel time, as well as the interactions among the variables.

[8] At the extreme one set of models was estimated that included a dummy for each census block group observed in the data in an attempt to control for neighborhood effects. These models are less appealing theoretically since a block group not only has neighborhood characteristics but also a spatial location. The use of a dummy for a neighborhood may bias results if it removes the influence of the neighborhood's location. Intermediate models, the full results of which are reported, use neighborhood characteristics such as average number

Center Accessibility

The statistical significance of parameter estimates on travel time to the central business district, its square and travel time to the nearest subcenter will determine whether household travel time is influenced by access to economic centers. Under the theory household travel time is assumed to increase with travel time to economic centers. A decreasing price gradient is predicted since households further from centers are assumed to trade greater travel expenses for a decrease in housing prices. A decreasing density gradient is predicted because households that demand more housing services (i.e., larger lots and larger houses) save on those expenditures by living further from economic centers. Positive parameter estimates for these variables would support the assumptions of the theory.

If these measures of accessibility of economic centers are found to be insignificant (either statistically or economically) the validity of the monocentric and limited polycentric models must be questioned. The metropolitan area may instead have transformed into a more disperse form with significant economic activity outside of identifiable economic centers. This more disperse form may allow households to minimize total travel by choosing destinations distant from identifiable economic centers rather than by choosing housing locations close to centers, as embodied in the theory of the monocentric and limited polycentric models.

Household Characteristics

A vector of household specific characteristics (x_i) are included in the empirical model to adjust for variation in travel resulting from those characteristics. As suggested by the theoretical model travel costs are a function of the opportunity cost of travel time. Household travel time captures the majority of these costs. For that reason (and others discussed above) household travel time was used as the dependent variable in the empirical model. While travel time is useful to estimate travel expenditures, as the theoretical model demonstrates the value of time can be expected to vary with income. Including a set of income dummy variables for a range of incomes in the regression captures this income effect. Since the opportunity cost of time rises with income, the initial expectation may be that household travel time will decrease with income. Studies, however, have found that

---

of rooms per house and house age.

travel is a normal good (Pushkarev and Zuban, 1977). Housing services are also thought by some to be normal goods (Pines, 1977; Simpson, 1992, at 18). Under central place theory, one may be expected to incur additional travel to reduce the cost of satisfying the preference for additional housing services. These competing influences of income on travel prevent any clear expectations as to the signs of the income variables.

Three household composition variables are included in the empirical model to predict travel time, the number of working adults, the number of nonworking adults and the number of children. All three of these variables are expected to have positive parameter estimates, as a household will likely undertake more travel for each additional household member. A working adult may be expected to add the most travel, as commuting to work is thought to be the majority of household travel. Additional children may be expected to add the least to household travel, as adults will minimize the added travel they must undertake to accomplish tasks for their children and themselves.

Household composition is also likely to affect household travel indirectly through housing choice. Under the theory, a propensity to consume more housing services with each new household member suggests that the household will choose a housing location more distant from economic centers, which will also increase household travel time. The strength of this influence on location choice is thought to be greatest for children, who contribute to a household's preference for a suburban lifestyle, where schools are thought to be of higher quality and crime is thought to be less prevalent. Theory also suggests that additional wage earners in the household are likely to lead to a more central housing location to minimize travel costs (Simpson, 1992, at 18-19). Since the purpose for including household characteristics is to remove the influence of these variables their affect on household location choice is not troubling.

Household travel is also expected to increase with the number of vehicles and number of licensed drivers. The purchase of an additional vehicle or the insurance for an additional driver is a sunk cost the benefits from which can be maximized only by undertaking additional travel. Additional vehicles and drivers also reduce the marginal costs of travel to the household, so should result in the household undertaking more travel. These variables may be related to household location, as suburban households may have more vehicles and licensed drivers as those areas have less access to public transportation. Their inclusion in the model, however, is important as they are likely to contribute directly to household travel time.

Dummy variables for apartment dwellers, townhouse dwellers and homeownership are also included in the empirical model. Both apartment and townhouse dummies are expected to have negative coefficients as they represent the purchase of less housing services than single family detached housing. These households that purchase less housing services are expected to choose more central locations that minimize travel expenses. Homeownership is expected to contribute to household travel, as household that wish to obtain the benefit of ownership, may be expected to undertake additional to minimize the costs of ownership.

Two dummy variables are included that specify when the household moved to its current residence. Labor theorists would believe that more recent movers will do so to reduce travel time, all other influences held constant (Simpson, 1992). Waddell, et al. (1993), however, found that commutes changed little with household relocation, and suggested that households move for amenities rather than to reduce commuting time. The expectation for these variables is therefore ambiguous.

A variable for the percentage of the household's trips that are taken by public transportation is also included in the empirical models. Travel time is thought by many to vary with mode choice. In addition, economists and transportation experts have advocated greater use of public transportation to reduce not only pollution associated with auto use but also travel time associated with auto congestion (Cervero, 1986; Downs, 1992). Inclusion of a public transportation usage variable will provide some information on the relationship of travel time and mode choice as well as remove the influence of the mode choice on travel time from the model.

Neighborhood Characteristics

Theoretically, neighborhood characteristics, such as the average house size and house age, are determined by the proximity of economic centers. Models omitting neighborhood variables, therefore, may be most consistent with urban economic theory. Studies of rent gradients, however, have found that neighborhood characteristics add substantially to the explanatory power of monocentric and polycentric housing price models

(Waddell, Berry and Hoch, 1993). The significance of neighborhood characteristics in these models implies that people are willing to undertake additional travel to live in neighborhoods with certain characteristics that are not explained by the accessibility to economic centers. This conclusion is bolstered by previous work that has found that household are willing to extend commutes substantially to live in areas with desired neighborhood amenities (Cropper and Gordon, 1991). Inclusion of the neighborhood characteristics in the model will test whether this propensity to undertake greater travel to live in a neighborhood with specific amenities extends to all travel (or is just limited to commuting time) as well as remove the influence of these neighborhood characteristics from the model.[9]

A continuous variable measuring the percentage of the neighborhood that is nonwhite minorities is included in the model. The influence of this variable may enter the model directly, as households are thought to undertake additional travel to live in neighborhoods with fewer minorities (Cropper and Gordon, 1992). The influence of the variable may also be realized indirectly since housing prices are thought to be higher in neighborhoods with fewer minorities. According to urban economic theory, households desiring residences in neighborhoods with fewer minorities should undertake additional travel to compensate for the neighborhood premium that they pay.

Variables measuring the percentage of the neighborhood under the age of eighteen and the percentage of the housing in the neighborhood that is occupied are also included in the model. Both are thought to be proxies for the safety of the community. A major impetus for suburban development is thought to be the desire to accommodate a preference for safer neighborhoods. Household travel time therefore is expected to increase in the percentage of the neighborhood under the age of eighteen.

The percentage of homes in the neighborhood that are single family detached, the average house age and the average number of rooms in homes in the neighborhood are also included in the model. These variables are included as measures of the housing quality in

[9] A third set of models removed all observable and unobservable "neighborhood effects" using dummy variables for each block group. The results of *F*-tests comparing the models including neighborhood characteristics with these "neighborhood effects" suggest that the removal of "neighborhood effects" added no significant explanatory power to the model. Appendix 5B is a full report of the results of the *F*-tests.

the neighborhood. Household travel is expected to rise with neighborhood housing quality since neighborhoods with higher quality housing are thought to be safer, have better public services and retain housing values. Large homes and single family detached homes are thought to increase the quality of the neighborhood. Older housing is thought to decrease the quality of neighborhood housing.

As a further test of the influence of public transportation on household travel time variables representing the percentage of the neighborhood that commute to and from work driving alone and commute to and from work by subway are also included in the model. Urban planners and transportation experts believe that drive alone commuting congests roads and therefore should increase household travel times (Cervero, 1986; Downs, 1992). Subway use, on the other hand, is thought to reduce travel times by reducing neighborhood road congestion.

The use of household specific and neighborhood demographic data allows the examination of both the direct influence of those factors on household travel time and controls for these influences in the estimation of the effect of distance to centers on household travel time. Removing the influence of these variables is critical in revealing the relationship between the accessibility of economic centers and household travel time.

The Models

Pairs of models were estimated for three different samples of the data. In each pair, the first model included no neighborhood variables. As noted above, the theory underlying the model is that the neighborhood composition and demographics are determined by (or, at a minimum, are related to) the accessibility of economic centers. Accordingly, the inclusion of demographic variables in the model may bias results. To examine this influence one model for each sample excluded any neighborhood demographic variables.

The data were separated into three different samples, as tests of the stability of the parameter estimates indicated that tests using the entire sample would be unreliable. The first set of models (the city models) analyzed data from households in the city of Washington, D.C. only. These models are useful in that they provide an understanding of whether variations in household travel time exist within the city itself. If conditions in suburbs differ substantially from those of the core metropolitan area their inclusion in the model may bias parameter estimates. Tests of parameter stability suggested that such biases exist. The city models are useful, however, as they provide a basis by which to compare travel of city residents with those of suburban residents.

The second pair of models (the suburban models) analyzed data from the only suburban jurisdictions – Virginia and Maryland – in the immediate city surroundings. The sample for these models is limited to households within one-hour travel time to the central business district. Similar to the city models, the suburban models are particularly useful in that they provide insight into the travel behaviors within the developed area surrounding the city, but exclude the city proper. Models including either households from sparsely developed areas more distant from the central business district or households in the city itself were rejected by tests of parameter stability. The suburban model (unlike the city model) includes variables for the travel time to the central business district term and its square. For this models the squared term proved necessary in tests of functional form.

The third set of models (the outlying area models) analyzed data from households in excess of one-hour travel time from the central business district. Tests of misspecification indicated significant parameter shift for observations more than one-hour travel time from the central business district. Intuitively, such a shift should be anticipated. Residents of outlying regions may have less economic ties to the metropolitan area. Portions of these areas are unlikely to have been converted to metropolitan area. The squared distance term proved unnecessary in the outlying area models and is therefore omitted from these models.

Mean daily household travel time differs slightly across the different samples. In the city sample mean household travel time is approximately one and three-quarters hours. For the suburban sample it is slightly greater than two hours. Mean household travel time for the outlying area sample is approximately two and one-third hours. This fact alone suggests that household travel time increases with travel time to the central business district. A wide variation in household travel times, however, is evident in all samples. In all of the samples the standard deviation is approximately two hours and the travel times observed ranged from a minimum of .1 hours to a maximum of slightly greater than 14.5 hours. Complete summary statistics for each of the samples appear in Appendix 5C.