Title
Analyzing the Determinants of Transit Ridership Using a Two-Stage Least Squares Regression on a National Sample of Urbanized Areas

Authors
Taylor, Brian D.
Miller, Douglas
Iseki, Hiroyuki
et al.

Publication Date
2003-09-01

ANALYZING THE DETERMINANTS OF TRANSIT RIDERSHIP USING A TWO-STAGE LEAST SQUARES REGRESSION ON A NATIONAL SAMPLE OF URBANIZED AREAS

Submitted for presentation at the
2004 Annual Meeting of the Transportation Research Board
July 2003

7,464 words

By

Brian D. Taylor and Douglas Miller

With

Hiroyuki Iseki and Camille Fink

UCLA Institute of Transportation Studies
3250 Public Policy Building
Los Angeles, CA 90095-1656
Telephone: (310) 825-7442
Telefax: (310) 206-5566
Email: btaylor@ucla.edu

ABSTRACT
Understanding the factors influencing transit ridership is central to decisions on transportation system investments and the pricing and deployment of transit services. Yet most previous analyses of transit ridership have examined one or just a few systems, have not included many of the control variables thought to influence transit use, and have not addressed the simultaneous relationship between transit supply and demand. This study addresses the shortcomings in the previous research by (1) conducting a cross-sectional analysis of transit use in 265 urbanized areas, (2) testing an array of variables measuring transit system characteristics, auto system characteristics, geography, metropolitan economy, and population characteristics, and (3) constructing two-stage least squares regression models to account for simultaneity between supply and demand. We find that most of the variation in transit ridership can be explained by (1) the size (population and area) of the metropolitan area, (2) the vitality of the regional economy (median housing costs), and (3) the share of the population with low levels of private vehicle access (carless households). We find further that transit patronage is to a lesser extent, explained by transit service levels and fares. The observed influence of fares on ridership is consistent with the literature. Likewise, the relative influence of transit service levels on ridership is greater than the influence of transit fares. Finally, our separation of transit service supply into two variables – an instrumental control variable and a residual policy variable – allows for more nuanced assessments of the ridership effects of changes in transit service.

INTRODUCTION

What explains transit ridership? Why does public transit carry a relatively large share of metropolitan trips in New York City, but such a small share in places like Houston, Atlanta, and Indianapolis? The answers to these simple questions are both obvious and elusive**.**

Public transit systems carry larger shares of person travel in older, larger metropolitan areas around the globe, but in most places – old and new, large and small – transit is losing market share to private vehicles. Figure 1 shows that both annual transit ridership and annual ridership per capita in the U.S. peaked during the 1940s. But, with the exception of the fuel, tire, and steel rationing years during and immediately following the Second World War, per capita transit use began to decline significantly during the 1930s and, despite four decades of increasing public subsidy in the U.S., has remained essentially constant since 1970.

[FIGURE 1]

In terms of the market share of metropolitan travel, public transit has for decades been losing customers to private vehicles. Nationally, only 2.1 percent of all trips were on public transit in 2001, compared to 85.8 percent by private vehicles, 9.9 percent by foot and bicycle, and 2.2 percent by other means (*1*). But consumption of transit service varies dramatically from place to place. Transit use is highest in the centers of the oldest and largest metropolitan areas, and virtually non-existent in many smaller cities and towns. In the U.S., New York City is the 800-pound transit gorilla – nearly 4 in 10 transit trips nationally in 2000 (38 percent) (*2*) were made in the greater New York City area.

Even the most casual observer of cities can offer informed speculation on why, for example, the share of year 2000 commuters using public transit in metropolitan San Francisco (19 percent) was nearly five times higher than in metropolitan Atlanta (4 percent) (*3*). Population density, levels of private vehicle ownership, topography, freeway network extent, parking availability and cost, transit network extent, service frequency, transit fares, transit system safety and cleanliness, and so on all surely play a role. The relative importance of these various factors, and the interaction between them is far from obvious. Yet understanding the influence of these factors is central to public policy debates over transportation system investments and the pricing and deployment of transit services. The research literature on explaining transit ridership, however, is surprisingly uneven – in some cases poorly conceived – with results that are often ambiguous or contradictory.

This paper presents an analysis that attempts to address some of the shortcomings present in previous studies of the determinants of transit patronage. We begin by developing a simple causal model hypothesizing the collective influence of a wide range of factors on transit ridership. Given this model, we briefly review and critique the previous research, emphasizing both the principal supportable findings and identifying many of the methodological problems plaguing this research. We then describe the national data set developed from the National Transit Database (NTD) and several other sources. We use this data in a cross-sectional regression analysis of transit ridership with a two-stage least squares model. Through this approach, we identify an array of factors thought to significantly influence transit ridership, while taking into account both the small samples sizes and the simultaneity conundrum of transit supply and transit demand common to many previous studies. We then present our models results and conclude with a discussion of the implications for policy.

In a nutshell, we find that most of the variation in transit ridership between urbanized areas – in both absolute and relative terms – can be explained by (1) the size (both population and area) of the metropolitan area, (2) the vitality of the regional economy (measured in terms of median housing costs), and (3) the share of the population with low levels of private vehicle access (measured in terms of the percent of zero-vehicle households). We find further that transit patronage is to a lesser, but still significant extent, explained by transit service levels and fares. Consistent with research on transit service elasticities, we find the relative influence of transit service levels on ridership to be greater than the relative influence of transit fares. Finally, by separating the service supply variable into an instrumental control variable and a residual policy variable, we estimate that large changes (particularly increases) in transit service are likely to have far less influence on transit ridership than many of the previous aggregate models of transit patronage would suggest.

What Determines Transit Demand?

Basic consumer economics theory tells us that a person consumes a good when the utility of consuming the good is higher than the disutility of its cost. A basic demand function presents the relationship between the cost (or price) of a good and the level of demand. As long as the cost of consuming a good is lower than an individual's willingness to pay, the good is consumed (*4,5*). While the demand for transportation is often viewed as derived from the demand for other goods,services, and activities, the application of basic consumer economics theory still holds (*6,7,8*). Thus the demand for a transit trip can be viewed as a function of both the utility of the trip and its costs: time (access time, wait time, travel time), money (transit fare), and uncertainty (schedule adherence, safety).

Estimating transit demand functions is complex, however, because the perceived utility and disutility of transit trips varies significantly from person to person and from trip to trip, even for the same person. First, the utility of a transit trip is to a large extent a function of the utility of the activity from which the demand for a transit trip is derived. While the utility, and hence, demand for a particular good, service, or activity can be ascertained, transit is likely just one of many possible ways to reach the desired good, service, or activity. Second, the perceived disutility of transit trip costs varies dramatically. Numerous studies have found that travelers perceive out-of-vehicle time (walking to and from transit stops, transferring, waiting at transit stops) as more onerous (and therefore more costly) than in-vehicle time (*9,10,11,12*). Therefore, someone who lives and works near transit stops on a particular line will likely perceive lower costs for a peak-hour, peak-direction trip than will a person traveling between the same two stops, but who lives and works farther from the stops and who is traveling at night. This is because the former person has a shorter out-of-vehicle time than the latter person, although the entire trip takes both of them the same total time. Third, some people do not have realistic substitutes for transit trips, but most do. Relatively fast, flexible private vehicles dominate metropolitan travel, and as noted earlier, even walking now far exceeds the number of trips made on public transit. Thus, most travelers find the relative utility of traveling by other modes (particularly private vehicles) to be greater than of public transit.

The characteristics of the transit service obviously affect the perceived costs of transit travel. Basic economic theory tells us that the actual consumption of goods is determined by the equilibrium point between the demand and supply curves under free market conditions (*4,5*). Put simply, in the absence of transit service, no service will be consumed, regardless of transit demand. On the other hand, increasing the network density, reducing headways, and/or lowering fares all lower the perceived cost of transit travel, move the demand and supply equilibrium point, and result in increased transit patronage. Further, if buses and trains are packed full and service supply is insufficient to accommodate demand, increased service supply will lead to increased consumption of transit trips by accommodating demand previously suppressed due to inadequate supply.

Thus, we can think of aggregate demand for public transit as a function of the collective characteristics of travelers, the physical and economic characteristics of metropolitan areas, the availability of substitute modes for travel, and the price, quantity, and quality of transit services (Figure 2).

[FIGURE 2]

Empirically, the level of service supply (usually measured in terms of vehicle service hours or vehicle service miles) is highly correlated with the consumption of transit trips in an area. Nationwide, 95 percent of the variation in transit trips in 2000 was explained by level of vehicle service hours. Obviously, the level of transit demand largely determines the supply of transit service. And just as obviously, the level of transit service supplied largely determines the level of consumption of transit trips. Given ongoing efforts to increase public transit patronage around the U.S., the nature and significance of this circular causality is especially relevant to public policy. For example, if service supply is largely a function of transit demand, increasing transit service may prove a costly and relatively ineffective way to increase transit use. If, on the other hand, transit demand is strongly influenced by transit service supply, then increasing transit service – by, for example, reducing headways – may be a more cost-effective way to increase ridership than reducing transit fares. Because the levels of transit supply and demand are jointly determined, it is impossible to consider one in isolation from the other. While this conundrum is well understood in economic theory (*13, 14*), it has largely gone unaddressed in the literature examining the factors explaining transit ridership. We now turn to this previous research.

Previous Research on the Factors Affecting Transit Ridership

Studies of the determinants of transit ridership can be grouped into two general categories: (1) research that focuses on traveler attitudes and perceptions, with both travelers and operators as the units of analysis, and (2) studies that examine the environmental, system, and behavioral characteristics associated with transit ridership. In general, the studies of attitudes and perceptions are descriptive in character, while system-focused studies tend to be structured as causal analyses. Both descriptive and causal analyses examine a host of factors related to transit ridership. These elements can be broadly divided into two categories: (1) internal (or policy) factors, and (2) external (or control) factors. Internal factors are those over which transit managers exercise some control, such as fares and service levels. External factors are largely exogenous to the system and its managers and include factors such as service area population and employment.

Descriptive analyses generally use survey and interview data of transit system managers and transit patrons to assess perceptions of the factors affecting ridership (*15, 16, 17, 18, 19, 20, 21*). Of the studies of operator perceptions and views, most, not surprisingly, emphasize internal factors. In general, transit managers report five general categories of strategies, programs, and initiatives affecting transit ridership: service improvements and adjustments; fare innovation and changes; marketing and information; new planning approaches and partnerships; and service quality and coordination.

Causal analyses are an attempt to posit and test hypotheses about the factors influencing transit ridership (*22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36*). While this is clearly an important area of inquiry for transportation policy research, studies of this type share surprisingly little in terms of data, methods, or findings. Many, but not all, of these studies use multivariate regression analysis to identify the factors most strongly related to changes in transit ridership. One unfortunate commonality between many of the previous studies – small sample sizes – raises questions about generalizability and statistical significance of findings (*23, 24, 27, 31, 32, 34, 35*). Furthermore, the broad conceptual factors hypothesized to influence ridership and the variables operationalized in these models vary widely. Not surprisingly, these models tend to find that a combination of internal and external variables explains transit ridership. Of the two, external factors (e.g., income, parking policies, development, employment, fuel prices, car ownership, and density levels) are found to have greater effects on ridership than internal factors. Of the internal factors, service quality is often found to be more important than low fares.

Descriptive and causal analyses both offer a range of various advantages and disadvantages. Descriptive analyses are based on sets of often interesting and rich qualitative data from surveys of and interviews with transit operators. This is a valid attempt to identify the factors experts believe affect ridership. However, these data pose methodological and interpretive concerns. This information is subjective and dependent on respondents' perceptions and assumptions about internal and external factors related to ridership (*19, 20, 21*). The data are subject to biases based on limited or incorrect information. Other descriptive studies fail to outline the specific data collection processes used to obtain information (*16*). In addition, the causal linkages between perceived factors and actual ridership are often simply asserted. Many of these studies are relatively old, and most of them do not specifically ask about perceptions of causality or the relative influence of internal and external factors. Some identify commonalities among agencies with ridership growth and conclude that they are related to ridership increases (*18,37*).

Causal analyses have the advantage of being more sophisticated empirical studies – of one, a few, or many agencies – that allow researchers to obtain better quality and a wider array of data than those found in descriptive studies. The generalizability of studies looking at a small number of systems is limited, but there is more opportunity for the conceptual development of models. In empirical causal analyses of many agencies, the use of data from a large number of agencies and outcomes produces more robust results. In addition, the results are more likely generalizable to other places and systems. However, these data sets have their own limitations. Some studies include data that are the most readily available, particularly Census data (*35*) to measure external variables. Most studies look only at unlinked trips (*23, 24, 26, 27, 28, 30, 32, 35*) rather than linked trips, a more accurate measure of transit ridership, because linked trips are more difficult to measure. Several studies consider only work trips in the models (*26, 33*). Finally, data aggregation and collinearity of variables can lead to spurious conclusions.

Moreover, the models developed in these studies are often not fully specified and there is inconsistency in the variables included in the models. For example, the studies vary widely in the modes examined; some focus specifically on rail or bus (*23, 27, 28*) and others consider multimodal systems (*24, 26*). Additionally, some factors, such as driver friendliness and other service-quality issues, are potentially important, but are difficult, if not impossible, to quantify. Auto access and operating costs measures are clearly important to transit use, but difficult to operationalize at the level of the transit system or metropolitan area. Further, measures of transit

service quality – including reliability, comfort, and convenience – and service efficiency are widely viewed as important, but difficult to quantify in the aggregate.
As discussed in the previous section, serious endogeneity problems between service supply variables and transit demand can arise, and most studies employing multiple regression analysis do not take into account the simultaneity between transit supply and demand (*23, 24, 25, 27, 28, 30, 31, 32*). The causality arrow points in both directions. In practice, transit operators typically respond to observed changes in service demand by adjusting service supply, which in turn further influences demand.

Only a few previous studies have tried to account for the simultaneity of transit supply and demand (*31, 38, 39, 40, 41, 42, 43*). Gaudry (*39*) uses a recursive model in which only the ridership level of the previous year, not the current year, affects the level of supply. Peng, et al. (41) use three-stage least squared estimation models to solve both demand and supply equations separately. Alperovich et al. (*38*) and Kemp (*42*) use structural equation models to relate variables of demand, supply, and service quality. Kyte, et al. (*44*) use simultaneous equation transfer function (STF) models to conduct a time series analysis. Interestingly, Liu (*31*) compares coefficients from structural equation ridership models with those obtained in single equation models and concludes that the simultaneous effect between transit demand and supply is likely small. Most of these studies consider the simultaneity of transit demand and supply to develop time-series analyses for particular agencies with the goal of developing projections of future transit demand in order to modify service levels and routes (*38, 39, 40, 41, 43*). While such time series analyses of individual transit operators are certainly relevant to service planning, the findings of these studies may not generalizable.

Thus, most previous aggregate analyses of the factors influencing transit ridership have examined one or just a few systems, have not included many of the external, control variables thought to influence transit use, and have not addressed the simultaneous relationship between transit service supply and transit patronage demand. This study has attempted to simultaneously address each of these shortcomings in the previous research by (1) conducting a cross-sectional analysis of transit use in 265 urbanized areas, (2) testing a wide array of variables measuring transit system characteristics, auto/highway system characteristics, regional geography, metropolitan economy, and population characteristics, and (3) constructing two-stage regression models to account for simultaneity between transit supply and demand. How we address each of these issues is detailed in the following section.

DESCRIPTION OF THE DATA AND MODEL FORMULATION

In the following discussion, we will show how simultaneity of demand and supply leads to inefficient and inconsistent estimate of coefficients in the ordinary least squares (OLS) model. Following Berechman (*45*), we can write the demand and supply functions as follows:

The general form of the demand function can be written as:

$$D = D(P, T, Y, Q, I, V, Z, R) \tag{1}$$

The general form of the supply function can be written as:

$$Y = Y(D, E) \tag{2}$$

D: a transit demand function for the service area
P: transit fare

T: a vector of travel times
Y: a vector of outputs (service supply)
Q: a vector of service attributes
I: a vector of passenger characteristics
V: a vector of prices of alternative modes
Z: a vector of urban characteristics
R: a vector of regional characteristics
E: a vector of exogenous factors to determine service supply

Assume that each vector has only one variable. Y has only one variable measured in terms of vehicle hours such as Y_{VHi}, and there are not any other endogenous variables. Then a simultaneous equation model is expressed by the following:

$$D = b_0 + b_1 * Y_{VHi} + b_2 * P_i + b_3 * T_i + b_4 * Q_i + b_5 * I_i + b_6 * V_i + b_7 * Z_i + u_i \quad (3)$$
$$Y_{VHi} = c_0 + c_1 * D + c_2 * E_i + v_i \quad (4)$$

When D_i is regressed on Y_{VHi} and the other variables in equation 3 without taking into account the endogeneity of two variables, the estimated coefficient for Y_{VHi} will be biased and inconsistent.

In order to understand intuitively, simplify equations (3), (4)

$$D_i = b_0 + b_1 * Y_{VHi} + u_i \quad (5)$$
$$Y_{VHi} = c_0 + c_1 * D_i + c_2 * E_i + vi \quad (6)$$

Substitute (5) into (6) to get:

$$Y_{VHi} = c_0 + c_1 * (b_0 + b_1 * Y_{VHi} + u_i) + c_2 * E_i + vi \quad (7)$$
$$Y_{VHi} = \{1/(1 - b_1 c_1)\}\{c_0 + c_1 * b_0 + c_2 * E_i + c_1 * u_i + vi\} \quad (8)$$

Therefore,

$$Cov(Y_{VHi}, u_i) = E(Y_{VHi} u_i) \neq 0 \quad (9)$$

Equation 9 indicates a violation of one of the conditions of OLS, and therefore leads to inefficient and inconsistent estimation. In order to understand this problem, assume that u_i increases as Y_{VHi} increases. That is, Y_{VHi} and u_i (or D_i and u_i) are positively correlated. Then OLS will produce a slope for the regression model in equation 5 that is larger than the actual slope.

In order to address this simultaneity problem, we need use a two-stage least squared (2SLS) regression method.

- Step 1: Regress Y_{VHi} on all exogenous variables for Y_{VHi}, ignoring D_i.
- Step 2: Obtain estimated values for Y_{VHi}, Y_{VHi}.
- Step 3: Regress D_i on Y_{VHi} and other exogenous variables for D_i.

Model Variables

Data were assembled from a variety of sources. The primary source for transit-related data was the National Transit Database (NTD), which is compiled annually by the Federal Transit Administration (*46*). All transit variables listed in this paper are for the year 2000. The data are compiled individually for each system operator. However, because we are unable to match other characteristics associated with the actual service areas of each transit provider, we chose to aggregate transit variables to a common level of geography. Because of federal funding strategies that use urbanized areas (UZAs), we chose this as the applicable level of geography.

Most of the demographic and other control variables were compiled from the 2000 U.S. Census, Summary File 3 (SF3), and which were also aggregated to UZAs. Examples of these variables include median rent, median household income, total population, and total land area. The bulk of the variables in the models either came from the Census or the NTD. A few others were taken from other sources. Among these are gas prices (*47*), and sprawl (*48*).

Some of the variables required construction from other, simpler variables within each data set. For example, our measure of route coverage (rtdens) is the total annual service hours, divided by the land area of the urbanized area. Variable construction details are summarized in the following table.

[TABLE 1]

As the table shows, we did not include some conceptual variables in our models. We considered these variables important in determining transit ridership, but the variables could not be found or constructed for a variety of reasons. For example, we know that restricted parking tends to result in higher transit usage. However, such data are not readily available for the 265 urbanized areas included in this study. Other variables are not included for a similar reason. That is, we expect them to be important for predicting transit ridership, but we were unable to find adequate measures. Most of our conceptual variables measuring urban form fall into this category.

Model Specification

Our first step is to test a model similar to Kain and Liu (*28*) in which the simultaneity is not addressed. Like Kain and Liu, we also use the natural logarithm to transform variables on both sides of the equation. This transformation is necessary due to the extreme skewness of the distributions in key variables such as revenue hours and UZA population. Other variables are transformed to allow easy interpretation of their coefficients.

This single stage model is summarized in equation 10:

$$D\hat{} = f(Y,Q, V, Z, I, R)+ \varepsilon \qquad (10)$$

In this and following models, all terms are as previously described in equations 1 and 2. The stochastic term is denoted ‘ε.’ In testing this model, we initially included as many of the key measures identified in Table 2 as possible. We then used this as a basis for creating a parsimonious model, removing terms that were either collinear with other terms, or that were not significantly correlated with the outcome measure.

[TABLE 2]

After confirming that our results were consistent with previous studies, we constructed a two stage model. The first stage uses exogenous variables to estimate the supply of transit; this term is then fed into the second stage. We also include the residual from the first stage (ε_1) as a policy variable in the second stage:

$$Y = g(E) + \varepsilon_1 \quad (11)$$
$$D\hat{} = f(Y, Q, V, Z, I, R, \varepsilon_1) + \varepsilon_2 \quad (12)$$

We produced two forms of this general model. The first was based on total supply and demand. Because the size of the metropolitan area was such a dominant term, we also produced a second set of models using per capita measures.

Model Results
We present our model results here in two parts. The first examines the factors influencing total urbanized area transit ridership, and the second examines the factors influencing urbanized area transit ridership per capita. In these two parts we present models that test a wide array of external and internal factors hypothesized to influence transit patronage – both without and with instrumental variables to predict transit service levels. We conclude each part by presenting and discussing our most parsimonious models to estimate total and per capita urbanized area transit ridership.

Total Urbanized Area Transit Ridership
Model 1 below presents the results of our initial regression of a wide array of external and internal factors hypothesized to influence aggregate transit ridership (Table 3). As with most other analyses of this sort, the results indicate service levels (measured here as vehicle hours of service (lnv31)) is – by far – more strongly associated with transit ridership than any of the other variables tested. The relationship is so strong that it leaves little unexplained variance to be accounted for by the other variables. In fact, a simple one variable regression finds that, in this sample of 265 urbanized areas, vehicle hours of service explains 95 percent ($R^2 = 0.9503$) of the variation in transit patronage.

[TABLE 3]

Among the other *Transit System Characteristics* tested, transit fare (lnfare) exhibited the expected negative and significant relationship with ridership. The dominance of a single transit operator in area (domin) was also positively and significantly related to patronage, though like all independent variables other than vehicle hours of service, the magnitude of the effect was relatively small. Two service quality variables – route network density (lnrtdens) and service intensity (lnsrvlv) – exhibited unexpected negative relationships with ridership. But these unexpected results are due to multicollinearity with the vehicle hours of service variable. In fact, a simple, two-variable regression model using route network density and service intensity as independent variables explains 55 percent ($R^2 = 0.5529$) of the variation in transit ridership. In this model, both service quality variables are positively and significantly related to ridership, with service intensity (Std Est = 0.67575) explaining about 50 percent more variation than route network density (Std Est = 0.42086).

For the *Auto/Highway System* variables tested, the percent of zero-vehicle households (pct_nocar) was positively and significantly related to transit patronage, as expected. In contrast, regional gasoline prices (ln_gas) were not significantly related to ridership. This is likely due to the relatively low levels of variation of average fuels prices (less than $0.30 for 95 percent of the urbanized areas in our sample) between one urbanized area and another.

Among the *Regional Geography* variables tested, population density (lndens) had the expected positive and significant effect on ridership, while population (lnpop) had an unexpected negative effect; this was due primarily to the high correlation between urbanized area population and service hours (R^2 = .893). For the *Metropolitan Economy* variables tested, median household income (ln_inc) and median monthly housing cost (ln_rent) exhibited the expected positive and significant relationships with ridership. None of the three *Population Characteristics* tested – percent of households below 150 percent of the poverty line (pct_lt150pl), percent of recent immigrants (pct_recimm), or percent African American (pct_blk_alone) – were significantly associated with ridership, at least at the metropolitan level analyzed here. The insignificance of the poverty variable is likely due to the fact that the highest apparent poverty rates among urbanized areas are in college towns, like Gainesville, Florida, and Iowa City, Iowa, which have relatively high levels of transit ridership, and in smaller, agriculturally-based cities, like Brownsville, Texas, and Bakersfield, California, which tend to have relatively low levels of transit ridership. The percent of recent immigrants is highly, positively correlated with urbanized area population.

Given the obvious simultaneity between transit service supply (measured here as vehicle service hours) and transit service demand (measured as passenger boardings), interpreting the results of Model 1 is problematic. To address this issue, we use a simultaneous equations approach to first develop a model to predict transit service supply, and then to use the predicted service supply variable from this first model as an instrumental variable in a second model to predict transit service demand.

Table 4 presents the second-stage results of the two sequential models. While a variety of models to predict total vehicle service hours were tested, we settled on a simple one-variable model for the first stage using urbanized area population (lnpop) which explains about 80 percent (R^2 = 0.7968) of the variation in vehicle hours of service.

[TABLE 4]

The predicted vehicle service hours variable (lv31hat) from the first-stage model was then used as an instrumental variable in a second model to predict transit patronage. In addition to this predicted vehicle service hours variable, a second policy service hours variable (v31rsd) was created by subtracting the predicted level of service from the actual vehicle service hours in an urbanized area. The logic here is that, at the margin, metropolitan areas choose to provide more or less transit service than would otherwise be predicted by overall levels of transit demand. Some areas, like Honolulu, Hawaii, and Ithaca, New York, provide substantially more transit service than would be predicted by urbanized area population, while others, like Montgomery, Alabama, and Nashua, New Hampshire, provide substantially less transit service. Thus, this second variable can be interpreted as a policy variable which measures the effects of transit service supply on transit ridership at the margin. Table 2 lists the ten urbanized areas where actual transit service levels most exceed predicted levels (apparent oversupply), and the ten urbanized areas where actual transit service levels are furthest below predicted levels (apparent undersupply).

In examining the urbanized areas that diverge most dramatically from the expected values of transit supply, several things jump out. Among urbanized areas with more transit than would be expected, many are dominated by large universities, which frequently have substantial transit systems designed to serve the university community. Others, like San Francisco, have urban densities conducive to transit and/or have restricted parking. The cities with less transit than predicted tend to be relatively small metropolitan areas.

Among the *Transit System Characteristics* variables, the predicted vehicle service hours variable (lv31hat), not surprisingly, explains most (Std Est = 0.83890) of the variation in transit ridership. However, the policy vehicle service hours variable (v31rsd), which measures how much the actual supply of transit service varies above or below the level of service predicted for a given urbanized area, also explains a very large share of the variation in transit ridership (Std Est = 0.50900), suggesting relatively elastic responses to changes in transit service at the margin.

In addition to these predicted and policy variables for transit service supply, we also included the same set of variables representing *Transit System Characteristics*, *Auto/Highway System*, *Regional Geography*, *Metropolitan Economy*, and *Population Characteristics* tested above. As expected, the all of the other variables perform identically to Model 1 above.

Given the high level of collinearity between many of the independent variables tested, we tested several models in developing parsimonious two-stage regression models to predict overall transit ridership in an urbanized area (Table 4, Model C). The models developed include variables for three (external) factors outside of the control of transit systems: (1) the ambient level of transit service demand (measured by the predicted level of vehicle service hours estimated in the first model), (2) the economic vitality of the region (measured by median housing costs), and (3) the proportion of the population with little or no access to private vehicles (measured by the percent of zero-vehicle households). The models also include two (internal) factors over which transit systems exercise control: (1) the level of transit service provided above or below what would otherwise be predicted by the size of the urbanized area, measured as the actual vehicle service hours provided less the predicted level of vehicle service hours, and (2) the transit fare, measured by total fare revenues minus total boardings.

Collectively, these five variables explain 97 percent (R^2 = 0.9702) of the variation in overall transit boardings between the urbanized areas analyzed. While the external, control variables (predicted service level, housing prices, and zero-vehicle households) in the second model account for about three-fourths of the variation in ridership, about a quarter of the variation is explained by the two internal, policy variables (service supply and fare levels). Importantly, and consistent with previous studies of service and fare elasticities (*22*), variations in service supply appear to have substantially more influence on transit patronage than do transit fare levels.

Per Capita Urbanized Area Transit Ridership

Because analyses of overall levels of transit ridership are so strongly influenced by the overall urbanized area population, we conducted a second analysis of the factors influencing per capita levels of transit ridership.

Reliably predicting vehicle service hours per capita proved far more challenging than predicting overall levels of transit service (Table 5). Our final first-stage model (Model D) to estimate vehicle service hours per capita was comprised of three independent variables: population density (lndens), proportion of zero vehicle households (pct_nocar), and percent African American (pct_blk_alone). Collectively, these three variables explained 28 percent (R^2 = 0.2756) of the variation in vehicles service hours per capita, substantially less than the 80 percent explained in first stage of the total ridership models.

[TABLE 5]

The less robust first-stage prediction of vehicle service hours per capita almost certainly explains the stronger influence of policy service hours (v31rsd) variable in the second-stage model. Here, the influence of both predicted (Std Est = 0.70954) and policy service levels (0.75633) are estimated to be approximately equal. We speculate, however, that, were we able to construct a more robust first-stage model, the estimated influence of varying service levels above or below predicted levels would decline. With respect to other *Transit System Characteristics*, all operate similarly what we observed in the overall ridership models. Transit fares and the relative dominance of a single transit system operate as expected, while collinearity with vehicle service hours causes the two other service level measures to produce unexpected signs. As discussed above, when vehicle service hours are excluded, both variables operate as expected.

Among the *Auto/Highway System*, *Regional Geography, Metropolitan Economy*, and *Population Characteristics* variables, land area (lnarea), median household income (ln_inc), percent of population below 150 percent of the poverty line (pct_lt150vl) all operate as expected, while median housing costs (ln_rent), average fuel prices (ln_gas), percent of trips made by means other than private vehicles or transit (pct_not_trans_sov), and percent of recent immigrants (pct_recimm) were not statistically significant.

To address the problem of collinearity among many of the independent variables evaluated, we again tested a variety of more parsimonious models, settling on the final model listed in Table 5-Model E. While these per capita transit ridership models control for urbanized area population size, the results of this final model are quite similar to the final total ridership model discussed above. The model explains 90 percent ($R^2 = 0.8945$) of the variation in transit patronage per capita among the urbanized areas analyzed. The only significant difference between the total and per capita ridership models is that the size (land area) of the urbanized area (lnarea) replaced the proportion of zero-vehicle households (pct_nocar) in the per capita model.

As with the final total ridership model above, three external control factors – ambient levels of transit demand (predicted vehicle service hours), urbanized area size (land area), and economic vitality (housing costs) – explain most of the observed variation in transit patronage. But the two internal, policy factors – service supply (actual less predicted vehicle service hours) and fare levels – are substantially related to per capita transit ridership. However, given the lack of robustness in the model estimating the predicted vehicle service hours variable, we do not feel confident in estimating the relative influence of the two policy variables on per capita transit ridership.

CONCLUSION: IMPLICATIONS FOR POLICY

Most previous aggregate analyses of the factors influencing transit ridership have examined one or just a few systems, have not included many of the external, control variables thought to influence transit use, and have not addressed the simultaneous relationship between transit service supply and transit patronage demand. This study has attempted to address each of these shortcomings in the previous research by (1) conducting a cross-sectional analysis of transit use in 265 urbanized areas, (2) testing a wide array of variables measuring transit system characteristics, auto/highway system characteristics, regional geography, metropolitan economy, and population characteristics, and (3) constructing two-stage simultaneous equation regression models to account for simultaneity between transit supply and demand.

Aggregate analyses like these clearly have limitations. While using urbanized areas, rather than individual transit systems, as our unit of analysis allowed us to include and test a wide array of regional, economic, and demographic variables on aggregate transit ridership, such a relatively coarse unit of analysis does not allow us to meaningfully evaluate a wide array of factors – such as personal safety, schedule reliability, and parking availability and costs – thought to significantly influence transit use. Further, aggregating to the urbanized area allows for between-group comparisons, but ignores the significant within group variation in nearly every record in our sample. This is particularly important for public transit because the transit use varies dramatically across metropolitan areas; in many areas, a substantial share of transit ridership is concentrated on just a few lines in and around the central parts of central cities, with far lower levels of patronage elsewhere. Such variation is not captured in this analysis. Our planned next steps with this research are to (1) include additional auto/highway system variables – such as roadway extent and congestion levels – in future models, (2) develop better aggregate measures of service quality (such as cumulative levels of schedule adherence), and (3) develop a series of models analyzing changes in transit ridership during the 1990s.

To conclude, we find that most of the variation in transit ridership between urbanized areas – in both absolute and relative terms – can be explained by (1) the size (both population and area) of the metropolitan area, (2) the vitality of the regional economy (measured in terms of median housing costs), and (3) the share of the population with low levels of private vehicle access (measured in terms of zero-vehicle households). We find further that transit patronage is to a lesser, but still significant extent, explained by transit service levels and fares. The observed influence of fares on ridership is consistent with the literature. And, consistent with research on transit service elasticities, we find the relative influence of transit service levels on ridership to be greater than the relative influence of transit fares. Finally, separating the service supply variable into two parts – an instrumental control variable and a residual policy variable – makes clear that large changes in transit service – such as a doubling of service supply – will not result in a near doubling of patronage, as many of the models developed in earlier research would imply.

ACKNOWLEDGEMENTS
This research was funded by a grant from the U.S. Department of Transportation and the California Department of Transportation through the University of California Transportation Center. The authors are grateful for this support from the University Transportation Centers research program. Any errors or omissions are the responsibility of the authors and not of the funding agencies. Thanks to Brent Boyd for his help in constructing the data set used for this analysis, and for his assistance in reviewing the literature. Thanks also to Norman Wong for his assistance with data collection, construction of several figures and tables, and with the production and layout of this document.

REFERENCES

1. 2001 National Household Travel Survey, 2001. http://nhts.ornl.gov/2001/index.shtml
2. American Public Transportation Association. Unlinked Passenger Trips and Passenger Miles Transit Agency Data, 2001. http://www.apta.com/research/stats/ridershp/
3. Census of Population, SF-4. U.S. Census Bureau, 2001. http://www.census.gov.
4. Varian, H..R. *Intermediate Microeconomics: A Modern Approach*, W. W. Norton & Company, Inc., New York, 1990.
5. Dawson, A.J. Segmentation of the Transit Market. *Transportation Quarterly*, Vol.37, No.1, 1983, pp.73-84.
6. Kanafani, A.K. *Transportation Demand Analysis*, McGraw-Hill, New York, 1983.
7. Train, K. *Qualitative Choice Analysis: Theory, Economics, and an Application to Automobile Demand*, third printing, the MIT Press Cambridge, Massachusetts, 1993.
8. Ben-Akiva, M. and S.R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*, The MIT Press Cambridge, Massachusetts, 1985.
9. Small, K.A. Urban Transportation Economics, 1992, p.44.
10. Small, K.A., R. Noland, X. Chu and D. Lewis. Valuation of Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation. In *National Cooperative Highway Research Program Report 431*, Washington, D.C.: Transportation Research Board, 1999.
11. Wardman, M. A Review of British Evidence on Time and Service Quality Valuations. *Transportation Research E*, Vol 37, 2001, pp. 107-128.
12. Horowitz, J.L., F.S. Koppelman, and S.R. Lerman. *A Self-Instructing Course in Disaggregate Mode Choice Modeling*. Report DOT-T-93-18. U.S. Department of Transportation, December 1986. http://ntl.bts.gov/DOCS/381SIC.html.
13. Kelejian, H.H. and W.E. Oates. *Introduction to Econometrics: Principles and Applications.* Harper & Row, Publishers, New York, 1981.
14. Kennedy, P. *A Guide to Econometrics*. The MIT Press, Cambridge, Massachusetts, 1998.
15. Abdel-Aty, M.A. and P.P. Jovanis. The Effect of ITS on Transit Ridership. *ITS Quarterly*, Fall 1995, pp. 21-25.
16. Bianco, M.J., K.J. Dueker, and J.G. Strathman. *Parking Strategies to Attract Auto Users to Transit*. Transportation Research Board Report. Washington, D.C.: National Academy Press, 1998.
17. Brown J.R., D.B. Hess, and D. Shoup. Unlimited Access. *Transportation*, Vol. 28, 2001, pp. 233-267.
18. Sale, J. Increasing Transit Ridership: The Experience of Seven Cities. Prepared for the Urban Mass Transportation Administration, U.S. Department of Transportation, November 1976.
19. Jenks, C. Transit Ridership Initiative. *TCRP Research Results Digest*, No. 4, February 1995. [http://gulliver.trb.org/publications/tcrp/tcrp_rrd_04.pdf]
20. Jenks, C. Continuing Examination of Successful Transit Ridership Initiatives. *TCRP Research Results Digest*, No. 29, August 1998. [http://gulliver.trb.org/publications/tcrp/tcrp_rrd_04.pdf]
21. Dueker, K.J., J.G. Strathman, and M.J. Bianco. Strategies to Attract Auto Users to Public Transportation. *TCRP Report Number 40*, TRB, National Research Council, Washington DC, 1998. [http://gulliver.trb.org/publications/tcrp/tcrp_rpt_40.pdf]

22. Cervero, R. Transit Pricing Research: A Review and Synthesis. *Transportation*, No. 17, 1990, pp. 117-139.
23. Chung, K. Estimating the Effects of Employment, Development Level, and Parking Availability on CTA Rapid Transit Ridership: From 1976 to 1995 in Chicago. In *Metropolitan Conference on Public Transportation Research: 1997 Proceedings*. University of Illinois, Chicago, May 30, 1997, pp. 255-64.
24. Gomez-Ibanez, J.A. Big-city transit, ridership, deficits, and politics. *Journal of the American Planning Association,* Vol. 62, No. 1, 1996, pp. 30-50.
25. Hartgen, D.T. and M.L. Kinnamon. *Comparative Performance of Major U.S. Bus Transit Systems: 1988-1977*, Charlotte, NC: Center for Interdisciplinary Transportation Studies, University of North Carolina at Charlotte, 1999.
26. Hendrickson, C. A Note on the Trends in Transit Commuting in the U.S. Relating to Employment in the CBD. *Transportation Research Part A: General*, 20A(1), 1986, pp. 33-37.
27. Kain J. and Z. Liu. Ridership Models for San Diego and Houston. In *Secrets of Success: How Houston and San Diego Transit Providers Achieved Large Increases in Transit Ridership*. Prepared for the Federal Transit Administration, U.S. Department of Transportation, May 25, 1995.
28. Kain J. and Z. Liu. *Econometric Analysis of Determinants of Transit Ridership: 1960-1990*. Prepared for the Volpe National Transport Systems Center, U.S. Department of Transportation, May 25, 1996.
29. Kitamura, R. A Causal Analysis of Car Ownership and Transit Use." *Transportation*, Vol. 16, No. 2, 1989, pp. 155-73.
30. Kohn, H.M. Factors Affecting Urban Transit Ridership. In *Bridging the Gaps* Conference, Canadian Transportation Research Forum, Charlottetown, Prince Edward Island, Canada, June 6, 2000.
31. Liu, Z. *Determinants of Public Transit Ridership: Analysis of Post World War II Trends and Evaluation of Alternative Networks*. Doctoral Thesis, In Architecture, Landscape Architecture, and Urban Planning, John F. Kennedy School of Government, Harvard University, Cambridge, Massachusetts, 1993.
32. McLeod, Jr., M.S., K. Flannelly, L. Flannelly, and R. Behnke. Multivariate Time-Series Model of Transit Ridership Based on Historical, Aggregate Data: The Past, Present, and Future of Honolulu. In *Transportation Research Record No. 1297*, TRB, National Research Council, Washington, DC, 1991, pp. 76-84.
33. Morral, J. and D. Bloger. The Relationship Between Downtown Parking Supply and Transit Use. *ITE Journal*, Vol. 66, No. 2, 1996, pp. 32-36.
34. Nelson and Nygaard, 1995 (cited in TCRP 1996)
35. Spillar, R.J. and G.S. Rutherford. *The Effects of Population Density and Income on Per Capita Transit Ridership in Western American Cities*. Presented at the 60th Meeting of the Institute of Transportation Engineers' Compendium of Technical Papers, pp. 327-331, August 5-8, 1988.
36. Syed, S.J. and A.M. Khan. Factor Analysis for the Study of Determinants of Public Transit Ridership. In *Journal of Public Transportation*, Vol. 3, No. 3, 2000, pp. 1-17.
37. Taylor, B.D., P. Haas, B. Boyd, D.B. Haas, H. Iseki, and A. Yoh. *Increasing Transit Ridership: Lessons from the Most Successful Transit Systems in the 1990s*. Norman Y. Mineta International Institute for Surface Transportation Policy Studies, 2002.

38. Alperovich, G., M.A. Kemp, and K.M. Goodman. *An Econometric Model of Bus Transit Demand and Supply*. The Urban Institute Working Paper No. 5032-1-4, Washington, DC, 1977.
39. Gaudry, M. An Aggregate Time-Series Analysis of Urban Transit Demand: The Montreal Case. In *Transportation Research*, 9, 1975, pp. 249-258.
40. Kemp, M.A. *Bus Service in San Diego: A Study of Patronage Growth in the Mid-1970s*. Urban Institute Working Paper 1470-1, June 1981
41. Kemp, M.A. *Planning for Fare Changes: A Guide to Developing, Interpreting, and using Fare Elasticity Information for Transit Planners*. Urban Institute Working Paper 1428-5, May 1981.
42. Kemp, M.A. *A Simultaenous Equations Analysis of Bus Route Patronage Growth in San Diego*. Presented at the 61st Meeting of the Transportation Research Board, Washington, DC, 1981.
43. Peng, Z., K. J. Dueker, J. Strathman, and J. Hopper. A simultaneous route-level transit patronage model: demand, supply, and inter-route relationship. In *Transportation* 24, 1997, pp. 159-181.
44. Kyte, M., J. Stoner, and J. Cryer. A Time-Series Analysis of Public Transit Ridership in Portland, Oregon, 1971-1982. In *Transportation Research A*, 22A/5, 1988, pp. 345-359.
45. Berechman, J. *Public Transit Economics and Deregulation Policy*. Elsevier Science Publishers, New York, 1993
46. National Transit Database. Federal Transit Agency. http://www.Fta.dot.gov/ntl/database.html.
47. U.S. Department of Labor. Bureau of Labor Statistics. http://www.bls.gov/data/
48. Burchell, R.W., N.A. Shad, D. Listokin, H. Phillips, A. Downs, S. Seskin, J.S. Davis, T. Moore, D. Helton, and M. Gall. Costs of Sprawl—Revisited. In *TCRP Report 39*, TRB, National Research Council, Washington, DC, 1998. 159-181.

List of Tables

List of Figures

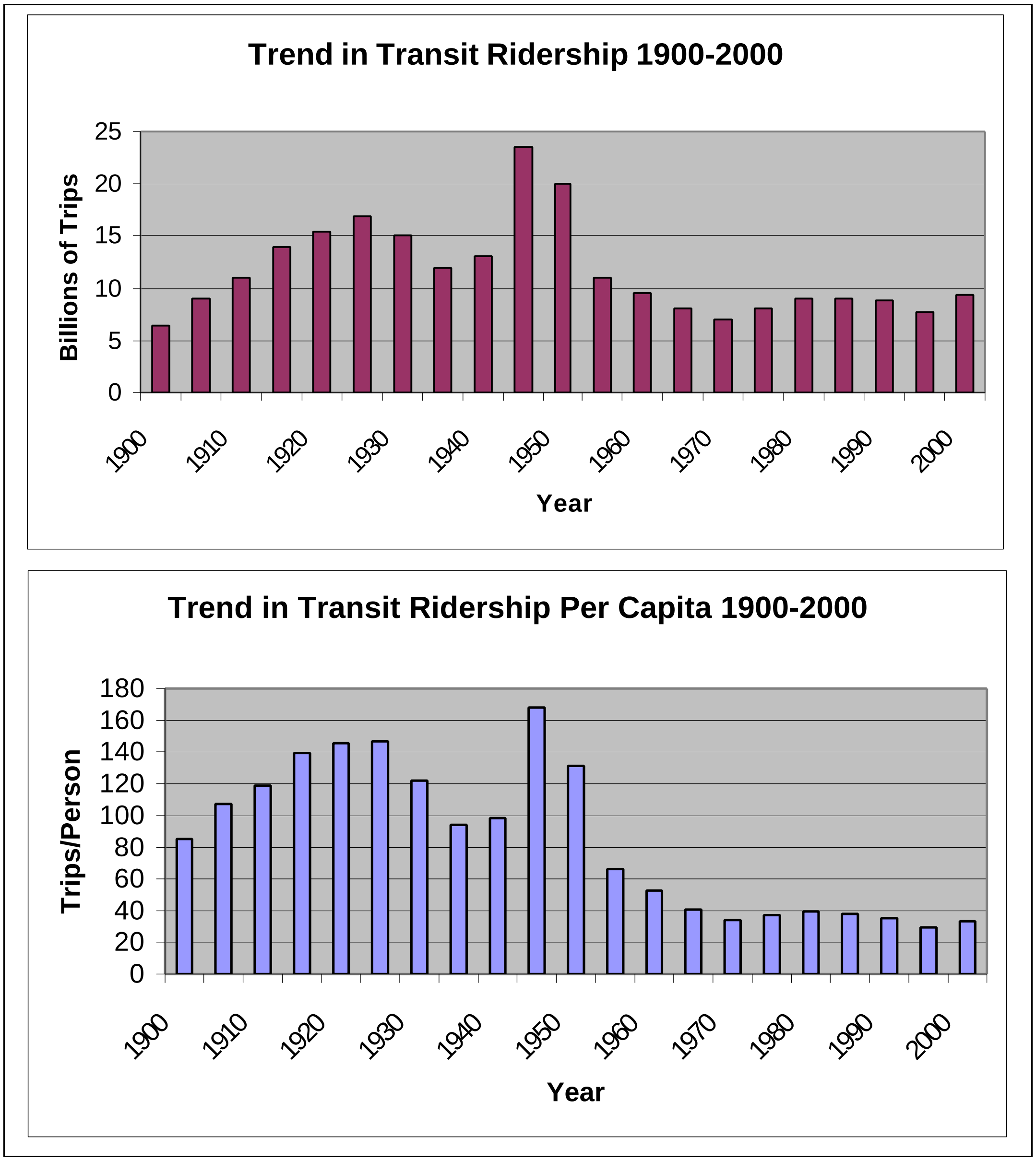

FIGURE 3 Trends in Transit Ridership.
(Source: American Public Transportation Association, 2000).

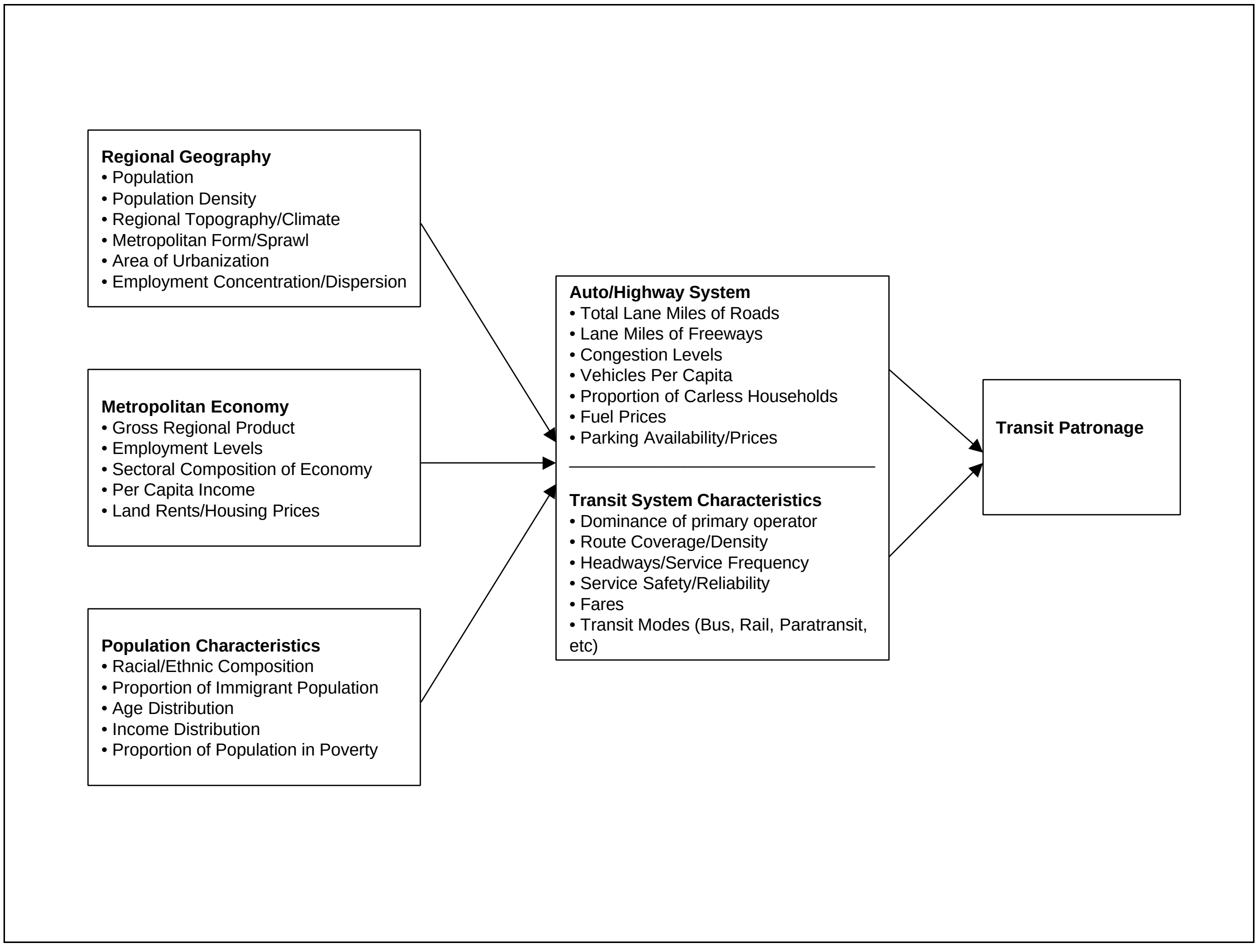

FIGURE 4 Conceptual Model of the Factors Influencing Aggregate Transit Demand.

TABLE 3 Conceptual Variables and Their Operationalization

Category	Variable	Used ?	Source	Variable Construction	Expected Relationship
Regional Geography					
	Population	Yes	Census 2000 SF3	Total Population	+
	Population Density	Yes	Census 2000 SF3	Population÷Geographic Area	+
	Regional Topography/Climate	No	Undetermined		
	Metropolitan Form/Sprawl	Yes	TCRP	Sprawl Index	-
	Area of Urbanization	Yes	Census 2000 SF3	Land Area	+
	Employment Concentration/Dispersion	No	Undetermined		
Metropolitan Economy					
	Gross Regional Product	No	Undetermined		
	Employment Levels	No	Bureau of Labor Statistics		
	Sectoral Composition of Economy	No	Bureau of Labor Statistics		
	Personal/Household Income	Yes	Census 2000 SF3	Median Household Income	-
	Land Rents/Housing Prices	Yes	Census 2000 SF3	Median Rent	+
Population Characteristics					
	Racial/Ethnic Composition	Yes	Census 2000 SF3	Single Race Population ÷ Total Population	?
	Proportion of Immigrant Population	Yes	Census 2000 SF3	Immigrant Population ÷ Total Population	+
	Age Distribution	Yes	Census 2000 SF3	Age Group Population ÷ Total Population	?
	Income Distribution	No	Undetermined		
	Proportion of Population in Poverty	Yes	Census 2000 SF3	Poverty Population ÷ Total Population	+
Auto/Highway System					
	Total Lane Miles of Roads	No	FHA Highway Statistics 2000		
	Lane Miles of Freeways	No	FHA Highway Statistics 2000		
	Congestion Levels	No	FHA Highway Statistics 2000		
	Vehicles Per Capita	No	Census 2000 SF3		
	Proportion of Carless Households	Yes	Census 2000 SF3	Carless Households ÷ Total Households	+
	Fuel Prices	Yes	Bureau of Labor Statistics	Average Gas Price	+
	Parking Availability/Prices	No	Undetermined		
Transit System Characteristics					
	Dominance of Primary Operator	Yes	NTD 2000	VRH of Largest Operator ÷ Total VRH	+
	Route Coverage/Density	Yes	NTD 2000	Route Miles ÷ Land Area	+
	Headways/Service Frequency	Yes	NTD 2000	VRM ÷ Route Miles	+
	Service Safety/Reliability	No	Undetermined		
	Fares	Yes	NTD 2000	Average Fares	-
	Transit Modes	No	NTD 2000		

TABLE 4 Urbanized Areas with the Greatest Deviations from Predicted Values

	Name	Vehicle Revenue Hours	Unlinked Trips	Total Population
Undersupply	Denton--Lewisville, TX Urbanized Area	15202	92743	299823
	Montgomery, AL Urbanized Area	9657	21363	196892
	Kingsport, TN--VA Urbanized Area	5957	53872	95766
	Nashua, NH--MA Urbanized Area	13950	257895	197155
	Panama City, FL Urbanized Area	9036	66482	132419
	Palm Bay--Melbourne, FL Urbanized Area	30895	302322	393289
	Benton Harbor--St. Joseph, MI Urbanized Area	3899	27805	61745
	Springfield, MA--CT Urbanized Area	59797	2262306	573610
	Greenville, SC Urbanized Area	33015	578508	302194
	Fayetteville--Springdale, AR Urbanized Area	20188	1178999	172585
Oversupply	Duluth, MN--WI Urbanized Area	139374	3016317	118265
	San Francisco--Oakland, CA Urbanized Area	5847653	433108429	2995769
	San Juan, PR Urbanized Area	4303756	131998311	2216616
	Honolulu, HI Urbanized Area	1245021	66602820	718182
	Santa Cruz, CA Urbanized Area	234757	6333449	157348
	Seaside--Monterey--Marina, CA Urbanized Area	189351	4016332	125503
	State College, PA Urbanized Area	106193	5331947	71301
	Champaign, IL Urbanized Area	216932	8724038	123938
	Ithaca, NY Urbanized Area	115688	2571605	53528
	Iowa Falls, IA Urban Cluster	45716	1256482	4908

TABLE 3 Single Stage Regression Model

Model A			
		Adj R-Sq	0.9713
Variable	**Parameter Estimate**	**Pr > |t|**	**Standardized Estimate**
Intercept	-6.68605	0.0124	0
Revenue Hours (lnv31)*	1.53424	<.0001	1.20614
Population Density (lndens)*	0.48863	0.0113	0.08545
Total Population (lnpop)*	-0.39029	0.0341	-0.24650
Median Income (ln_inc)*	0.45850	0.0650	0.04705
Median Rent (ln_rent)*	0.33781	0.0871	0.03828
Average Gas Price (ln_gas)*	0.59816	0.1368	0.02231
Percent Carless Households (pct_nocar)	3.88929	<.0001	0.06673
Percent Low-Income Population (pct_lt150pl)	0.24833	0.7398	0.00855
Percent Recent Immigrant (pct_recimm)	3.16463	0.1417	0.02701
Percent African American (pct_blk_alone)	0.26281	0.2871	0.01624
Transit Fare (lnfare)*	-0.32602	<.0001	-0.09921
Route Density (lnrtdens)*	-0.37499	0.0331	-0.13801
Service Level (lnsrvlv)*	-0.35368	0.0652	-0.09771
Dominant Operator (domin)	0.39954	0.0764	0.02505
*** Natural Log**			

TABLE 4 Total Ridership Models—Second Stage, Full and Parsimonious

Model B			
		Adj R-Sq	0.9713
Variable	**Parameter Estimate**	**Pr > |t|**	**Standardized Estimate**
Intercept	-7.48552	0.0051	0
Predicted Revenue Hours (lv31hat)*	1.18581	<.0001	0.83890
Residual Policy Variable (v31rsd)	1.53424	<.0001	0.50900
Population Density (lndens)*	0.48863	0.0113	0.08545
Median Income (ln_inc)*	0.45850	0.0650	0.04705
Median Rent (ln_rent)*	0.33781	0.0871	0.03828
Average Gas Price (ln_gas)*	0.59816	0.1368	0.02231
Percent Carless Household (pct_nocar)	3.88929	<.0001	0.06673
Percent Low-Income Population (pct_lt150pl)	0.24833	0.7398	0.00855
Percent Recent Immigrants (pct_recimm)	3.16463	0.1417	0.02701
Percent African American (pct_blk_alone)	0.26281	0.2871	0.01624
Transit Fare (lnfare)*	-0.32602	<.0001	-0.09921
Route Density (lnrtdens)*	-0.37499	0.0331	-0.13801
Service Level (lnsrvlv)*	-0.35368	0.0652	-0.09771
Dominant Operator (domin)	0.39954	0.0764	0.02505
Model C			
		Adj R-Sq	0.9702
Variable	**Parameter Estimate**	**Pr > |t|**	**Standardized Estimate**
Intercept	-3.81447	<.0001	0
Predicted Revenue Hours (lv31hat)*	1.17894	<.0001	0.83404
Residual Policy Variable (v31rsd)	1.17808	<.0001	0.39084
Median Rent (ln_rent)*	0.67348	<.0001	0.07632
Percent Carless Household (pct_nocar)	4.58217	<.0001	0.07861
Transit Fare (lnfare)*	-0.31441	<.0001	-0.09567
*** Natural Log**			

TABLE 5 Per Capita Ridership Models—Second Stage, Full and Parsimonious

Model D			
		Adj R-Sq	0.9005
Variable	**Parameter Estimate**	**Pr > |t|**	**Standardized Estimate**
Intercept	-5.31810	0.0482	0
Predicted Revenue Hours (lv31hat)*	1.99194	<.0001	0.70954
Residual Policy Variable (v31rsd)	1.43284	<.0001	0.75633
Total Land Area (lnarea)*	0.16485	<.0001	0.17307
Median Income (ln_inc)*	0.77061	0.0022	0.14526
Median Rent (ln_rent)*	0.30720	0.1689	0.06395
Average Gas Prices (ln_gas)*	0.17836	0.6429	0.01222
Non-Car/Transit Commute (pct_not_trans_sov)	0.94843	0.3290	0.04073
Percent African American (pct_lt150pl)	1.11407	0.0958	0.07047
Percent Recent Immigrant (pct_recimm)	2.55803	0.2477	0.04011
Transit Fare (lnfare)*	-0.30447	<.0001	-0.17020
Route Density (lnrtdens)*	-0.26991	0.0120	-0.18248
Service Level (lnsrvlv)*	-0.24126	0.0458	-0.12243
Dominant Operator (domin)	0.40178	0.0776	0.04627
Model E			
		Adj R-Sq	0.8945
Variable	**Parameter Estimate**	**Pr > |t|**	**Standardized Estimate**
Intercept	-0.38590	0.5686	0
Predicted Revenue Hours (lv31hat)	1.61076	<.0001	0.57376
Residual Policy Variable (v31rsd)	1.17944	<.0001	0.62258
Total Land Area (lnarea)*	0.19332	<.0001	0.20296
Median Rent (ln_rent)	0.46022	<.0001	0.09580
Transit Fare (lnfare)	-0.27415	<.0001	-0.15325
*** Natural Log**			