7.0 CONCLUSIONS

The analysis described in this report, summarized in Table 16, suggests strongly that light rail and commuter rail transit performs better when there is a large Central Business District. In fact, light rail may not work at all when CBDs get too large since ridership may outstrip the mode's carrying capacity. For commuter rail, the larger CBDs produce more effective services, but are slightly less cost-efficient.

The density of the CBD is particularly important for commuter rail, probably because there is usually only one terminal station and lower density CBDs may put some jobs beyond easy reach of the terminal station. Light rail, in contrast, is less affected by the density of the CBD since there are likely to be multiple stations to serve lower density CBDs.

The importance of both the size and density of Central Business Districts suggest that corridors that do not pass through or terminate in a Central Business District would be harder pressed to be cost-effective.

Residential density itself matters for light rail and commuter rail but, in the latter case, density is confounded by the effect of income, since commuter rail's higher fares attracts more riders with higher incomes, who also tend to live at lower densities.

The length of the rail line assumes some importance for both light rail and for commuter rail. Longer light rail lines are both slightly more cost-efficient and effective. But the effects diminish with length. Commuter rail lines are much more cost-efficient when they are longer and their effectiveness declines beyond 50 miles. At short distances there often is not enough riders to justify even minimal service.

The availability of parking and feeder bus service can help to achieve higher performance levels, all else being equal; feeder buses more strongly affect light rail and parking more strongly affects commuter rail.

Among the more interesting findings in this research is the distinctly different characteristics of light rail and commuter rail. It is clear that they serve different markets and different land uses patterns. Indeed, there are more dissimilarities than similarities. This does not imply that in any one metropolitan area they both may not have a niche, only that they have different niches.

Not accounted for here but worthy of serious exploration is a fuller consideration of costs, including those saved as a result of other modes not used, if the rail line is put in place. To accomplish this is it would desirable to assign the rail ridership to the modes from which riders would be diverted — auto and bus — and estimate the appropriate savings in operating, capital and full environmental costs. Beyond that, the application of the full costs of both transit and highway modes can balance the burden that rail transit must now bear in proving its value.

Finally, this effort should not be viewed as a substitute for a careful examination of all transportation alternatives in all types of corridors including those that do not end in a CBD, accounting for site-specific conditions and preferences. Rather, it should be seen as a means to understand the role that land uses in a corridor play in determining costs.

Further, it makes clear the need to integrate transit planning with land use planning at the earliest possible stage, a finding that is reinforced in the case studies prepared for another report of this project, Public Policy and Transit Oriented Development: Six International Case Studies.

Table 16. Summary of Findings on Cost Efficiency and Effectiveness for Hypothetical Rail Corridors

Factor	Cost Efficiency	Effectiveness
Light Rail		
Residential density gradient	highly positive	highly positive
CBD employment	moderately negative at high CBD job levels may not be possible	highly positive
CBD employment density	slightly positive	moderately positive greater impact for larger CBDs
Feeder bus	unclear	highly positive
Parking availability	unclear (site-specific)	moderately positive
Line length	slightly positive	slightly positive
Commuter Rail		
Residential density gradient	not significant	not significant
CBD employment	slightly negative, for smaller CBDs may have insufficient riders, especially for shorter line lengths	highly positive
CBD employment density	highly positive	highly positive
Feeder bus	unclear	moderately positive
Parking availability	unclear (site-specific)	highly positive
Line length	strongly positive, insufficient riders for shorter lengths	varies, best at 50-mile length

APPENDIX A

This page left intentionally blank.

Ridership Models for the Bay Area Rapid Transit System: Influences of Built Environment and Other Factors on Station Passenger Trips

1.0 Introduction

This analysis complements the main analysis of Topic 1 by conducting a in-depth analysis of how the land-use environment shapes transit demand for a single transit system — the Bay Area Rapid Transit (BART). While other modeling conducted under Topic 1 has concentrated on studying light rail and commuter rail services, this analysis is carried out for a heavy rail system. Along with the Chicago metropolitan area, the San Francisco Bay Area is being used as a case context for conducting fairly in-depth analyses for most topics of the Phase II study. This modeling also includes detailed information that could not be included in the United States model because it was not available for all systems. This includes information about the type and mix of land uses near stations and employment densities near all stations, not just in the central business district.

In estimating demand models, stations represent individual cases. The BART system presently has 34 stations (see Figure 1), so there are 34 data points used in the analysis. Ridership data compiled from the Spring of 1990, representing turnstile entries and exits, serve as the main dependent variable for the analyses. Besides station-by-station ridership totals, analyses are also conducted for ridership rates — specifically, ridership per 1,000 population in the catchment area, and ridership per square mile of the catchment area. Each of the 34 BART stations has a catchment area, and these are identified in the Topic 3 report. A catchment area is a contiguous area that captured 90 percent of all access trips to and egress trips from a BART station. The average catchment area is around 90 square miles with a radius of around 7 miles, though as noted in the Topic 3 report, there is considerable variation around these averages. In general, findings of *ridership rates* (e.g., trips per 1,000 population of the catchment area) are likely to be most useful from a policy-making standpoint since such statistics are more easily transferable to other transit properties.

2.0 Research Methods

Multiple regression was used for modeling the influence of various land-use, pricing, service, and other factors on the demand for BART trips. Both linear and log-log functional forms were estimated and are presented. The overall fits of both function forms were fairly good. One advantage of presenting log-log models is that regression coefficients represent elasticities. Thus the relative sensitivity of transit demand to land-use variables, price, service levels, and other factors can be easily compared using elasticities. Elasticities are generally well-understood by practitioners, and can be directly applied to ridership models developed by local transit properties.

In addition, graphs that reveal the sensitivity of transit trips to land-use variables and other control factors are presented in this report. These graphs are generally more intuitive and accessible than the regression results.

The process of developing regression models involved a fairly data intensive effort. Initially, associative relationships between variables were postulated (as discussed below). This was followed by generating various correlation matrices and scatterplots among variables. Attempts were made to estimate fully specified models that incorporated many of the variables that classical economic theories suggest influence travel demand. Because of multi-collinearity among many of the candidate variables and in an effort to present reasonably intuitive and parsimonious models, variables were selected that, in combination, provided a good predictive model that met underlying assumptions of ordinary least squares estimation. Various residual diagnostics were also employed in estimating models.

2.1 Variables Used in Analysis

Three dependent variables, listed and defined in Table 1, were used in the analysis. Models estimating total weekday passenger trips are, by design, scaled to the unique ridership characteristics of the BART system. Thus, the two rate measures of ridership demand — passenger trips per 1,000 population and passenger trips per square mile — were estimated, using the defined catchment area of each station as the geographic region for gauging population and land area.

Five sets of predictor variables were used: land use, pricing, bus service, transportation supply, and demographic. These sets and the variable in each set are also listed and defined in Table 1.

Land Use Variables

The land use variables are the ones of primary policy focus of this report. Employment and population density are the principle land-use variables examined in this study. Employment density was measured as employees per acre for a one mile radius from downtown stations and two mile radius around all other stations. A fairly restrictive geographic area was chosen for measuring employment density (as opposed to catchment areas) because of the nature of access trips to and egress trips from work sites. Since at the work end of trips most employees do not have access to cars, trips tend to be short. Thus, as discussed in the Topic 3 report, limiting the analysis of employment densities to the area reasonably close to rail stations is appropriate. Population density, however, is expressed for the entire catchment area — specifically,

Table 1. Dependent and Candidate Independent Variables Used for Modeling Transit Demand for the BART System

Dependent Variables

ENTRIES — Average weekday entries through station turnstiles, Spring 1990. See Appendix B. Source: BART Planning Department.

EXITS — Average weekday exits through station turnstiles, Spring 1990. See Appendix X. Source: BART Planning Department.

RIDERS — Total weekday passenger trips entering and exiting station, Spring 1990. ENTRIES EXITS

RIDEPOP — Weekday passenger trips (entries and exits) at station per 1,000 population in station catchment area. Catchment area is contiguous area that captures 90% of all access trips to and egress trips from station.

RIDEAREA — Weekday passenger trips (entries and exits) at station per square mile of station catchment area.

Land Use Variables

EMPDENS — Employment density, in employees per acre in 1990. Measured for census tracts and block groups that encompass a one mile radius around downtown stations (Embarcadero, Montgomery, Powell, Civic Center, Oakland 12th St., Oakland 19th St., and Berkeley) and a two mile radius around other stations. GIS used to create buffers for estimating employment within these radii. Source: 1990 Census Transportation Planning Package, Part II, Metropolitan Transportation Commission.

POPDENS — Population density, in 1990 population per square mile of station catchment area. Source: 1990 census, STF-3A.

AREA — Area of catchment zone, in square miles. GIS used for computing area.

ENTROPY — Index of land-use mixture. Entropy = $\{;_i[p_i{}^{*}\ln(p_i)]\}/\ln(k)$ where p_i = proportion of land area in land-use category i, and k = number of land-use categories; ranges between 0 and 1, where 0 signifies land devoted to a single use and 1 signifies all land area is evenly spread among all uses.

COMMERCI, INDUSTRI, RESIDENT — Proportion of land area in commercial use for one-mile radius around station. Source: 1990. Association of Bay Area Governments land use inventory.

Table 1 (Continued). Dependent and Candidate Independent Variables Used for Modeling Transit Demand for the BART System

Pricing Variables

AVGFARE	Average oneway adult cash fare from each station to all other stations, Spring 1990. Calculated from BART fare matrix. Source: BART Planning Department.
CBDFARE	Oneway adult cash fare from station to the downtown San Francisco Montgomery Street Station, Spring 1990.

Complementary and Competing Bus Service Variables

FEEDER	Route miles of all bus services in station catchment area that feed into BART station (i.e., bus stop is within 200 feet of station entrance). These are complementary services. Bus services include AC Transit, San Francisco Muni (diesel and electric trolley), Golden Gate Transit, SamTrans, Central Contra Costa County Transit, and municipally sponsored shuttle services. (Light rail and cable car feeder mile in San Francisco is not included.) Data files were obtained from the Metropolitan Transportation Commission containing geocoded points for bus stops; route mile distances were calculated by overlaying a street layer on the bus route layer and using GIS distance measuring functions. Source: Data files provided by the Metropolitan Transportation Commission and local transit agencies.
PARALLEL	Route miles of all bus services in station catchment area that do not fee into BART station and that essentially parallel BART lines within the catchment zone. These are competitive services. Bus services include AC Transit, San Francisco Muni (diesel and electric trolley), Golden Gate Transit, SamTrans, Central Contra Costa County Transit, and municipally sponsored shuttle services. (Light rail and cable car feeder mile in San Francisco is not included.) Data files were obtained from the Metropolitan Transportation Commission containing geocoded points for bus stops; route mile distances were calculated by overlaying a street layer on the bus route layer and using GIS distance measuring functions. Source: Data files provided by the Metropolitan Transportation Commission and local transit agencies.

Table 1 (Continued). Dependent and Candidate Independent Variables Used for Modeling Transit Demand for the BART System

Transportation Supply Variables

PARKING	Park-and-ride spaces at station, surface and structured. Source: BART *Systemwide Parking Inventory, 1990.*
TERM_NT	Terminal or Near-Terminal Station (0 = no, 1 = yes). Near-terminal stations are those toward the end of a line that function like terminals because they are closer to freeways than actual terminals and thus serve a larger catchment area. BART's near terminal stations, El Cerrito del Norte and Pleasant Hill, have larger supplies of parking than terminal stations since they are easier to reach by freeway.
FWYPX	Freeway proximity. Lineal distance of the nearest limited-access freeway to the station entrance, in feet.

Demographic Variable

HHINCOME	Annual household income for households within station catchment area, 1990. Measured using GIS. Source: 1990 census STF-3A.

An entropy index of land-use mixture was also used as a candidate predictor variable. Using data from the Association of Bay Area Governments, the shares of land area within a one (CBD stations) or two mile (all other stations) radius of stations that is devoted to commercial, office, residential, industrial, and institutional uses were determined. (Data are broken down by hectare grid-cells; using GIS buffers, the proportions of total area within each land use were determined.) An entropy value close to one indicates heterogeneous land use compositions whereas a value close to zero indicates virtually all land is devoted to a single use.

Pricing Variables

According to economic theory, the demand for transit should also be influenced by the price of its chief competitor, the private automobile. No price variables for competitive modes were used, however, since the differentials between these prices and those charged for transit trips would be similar across all stations (assuming a standard cost per automobile mile of travel was used). Additionally, research shows that most automobile travelers do not weigh full costs when making automobile trips, but rather are most cognizant of direct, out-of-the-pocket costs. In general, including price variables for competitive modes is only necessary when a longitudinal analysis is being conducted and the influences of factors like changing fuel prices over time need to be considered.

Economic theory holds that, all else being equal, demand declines as prices increase. Two price variables were included in the analysis. One measures, for each of the 34 BART stations, the average adult cash fare to the remaining 33 stations that was charged in the spring of 1990. Because there was not significant variation across the stations, a second price variable was included — the weekday adult cash fare from each station to the downtown San Francisco Montgomery Street station, one of the busiest and most central stations on the system. It should be noted that since BART employs distance-based fares, this variable also serves as a proxy for the distance of a station from downtown San Francisco, the Bay Area's dominant CBD.

Bus Services

Another factor influencing transit demand is service levels. During any time of the day, most BART stations receive similar service levels, so including this as a predictor variable in a cross-sectional model is problematic. In a longitudinal analysis, adjustments would need to be made for changing service intensity over time, however the lack of significant variation in a cross-sectional study precludes including measures of direct service level.

The transit service variable that does vary significantly across stations is the route miles of feeder bus service. Feeder buses complement rail services and thus should contribute positively to rail ridership. A bus route was considered to be a feeder service if one of its stops was within 200 feet of a station entrance.

In addition, some bus services compete with rather than complement rail services. These are generally services that operate parallel to rail lines, often on major arterials one or more blocks away. As shown in Table 1, the route miles of parallel bus services were included as a predictor variable to reflect this competitive influence.

Transportation Supply Variables

Other features of transportation supply that are thought to shape transit demand are parking supplies, whether a station is a terminal or not, and the proximity of the nearest freeway. Nearly all suburban BART stations are surrounded by large parking lots, meaning they are able to draw passengers from a larger catchment area. Terminal or near-terminal stations, moreover, also tend to have large catchments since they are the first stations reached by those living beyond rail corridors.

Demographic Variable

The only demographic variable used in these analyses was the annual 1990 household income for households within station catchments.

2.2 Anticipated Relationships

Table 2 postulates the expected statistical relationship between each of the ridership variables and predictor variables considered in this analysis. Density and land area are expected to positively influence ridership levels and rates. The effects of mixed land uses, however, is unclear. Mixed uses are thought to exert their greatest influence on access trips to rail stations — e.g., they can encourage walking from reasonably close by residences to stations. At the station-to-station, or corridor, level, however, their influences are likely fairly weak. Whether they would encourage rail riding is unclear. On the one hand, having mixed uses in station areas might induce some to opt for rail transit for convenience reasons. On the other hand, to the degree that mixed uses obviate the need for rail travel in the first place, they could be negatively associated with ridership.

Price can clearly be expected to exert a negative influence on ridership. One must invoke the *ceteris paribus* assumption, however. Transit ridership rates, for example, might very well be high for distant stations where high average fares are paid; however, controlling for distance and other features of these trips, the marginal contributions of higher fares should be to reduce ridership rates.

As noted, feeder buses complement rail services and thus should be positively associated with ridership. The partial correlation of ridership rates and parallel route miles, however, should be negative.

Parking supply and terminal locations (which often have large parking supplies and frequent feeder bus services) should be positively correlated with ridership rates. The effects of freeway proximity, however, are difficult to foretell. On the one hand, being close to a freeway could drain away transit ridership by providing an alternative line-haul route. On the other hand, being close to a freeway means that those who are park-and-riding have easier access to stations, particularly suburban stations.

Table 2. Anticipated Sign of Partial Correlation between Ridership and Independent Variables

	Anticipated Sign
Land Use Variables	
EMPDENS	+
POPDENS	+
AREA	+
ENTROPY	?
Pricing Variables	
AVGFARE	-
CBDFARE	-
Complementary and Competing Bus Service Variables	
FEEDER	+
PARALLEL	-
Transportation Supply Variables	
PARKING	+
TERM_NT	+
FWYPX	+/-
Demographic Variable	
HHINCOME	+/-

Lastly, the likely influence of household incomes on ridership is also difficult to postulate. Transit is often thought of as an inferior good, thus ridership should decline with income. However, this relationship holds less for large metropolitan areas with serious traffic congestion problems; in these places, the San Francisco Bay Area included, rail transit often becomes a legitimate and convenient means of travel, especially to major downtowns. Additionally, since most heavy rail systems are radial, designed to funnel suburbanites to downtown, and since many downtown San Francisco workers have well-paid office jobs, the association between income and ridership might be positive. Indeed, ridership surveys reveal that BART's peak hour customers average higher incomes than the Bay Area average.

3.0 Models of BART Ridership

3.1 Total Weekday Ridership

Table 3 reveals that, controlling for the land area of the catchment zone as well as other factors, employment density increases total passenger trips (turnstile entries and exits) for BART stations. Holding all other factors constant, an increase of 10 workers per acre for a one to two mile radius from a station is associated with 1,430 more daily rail trips. Interestingly, mixed land uses were negatively associated with rail patronage, controlling for other factors. This suggests that any benefits of mixed land use are unlikely to be registered on the line haul segment of trips; instead, most benefits should be associated with access trips, such as revealed in the Topic 3 report.

Consistent with expectations, feeder bus services increase rail ridership and higher fares lower ridership. Every ten route miles of service within a station's catchment is, on average, associated with 1,662 more daily trips to that station, removing the influences of all other factors.

Overall, a fairly good predictive model was derived for total ridership. Over 92 percent of the variation in total ridership was predicted by the five variables shown in Table 3. All of the predictor variables were statistically significant at the .05 probability level.

3.2. Total Weekday Ridership, Log-Log Form

Fairly similar relationships were found when the model was expressed in log-log form, as shown in Table 4. Although the R-squared statistic declined slightly, the ability to produce coefficients in elasticity form is a significant advantage of this model. Based on the elasticities, ridership appears to be only moderately sensitive to changes in density and feeder bus services. Demand is significantly more sensitive to fare levels — on average, around three times as much as it is to employment densities or feeder bus service levels.

Two additional variables entered the log-log model of total weekday ridership — household income and the terminal/near-terminal dummy variable — and one of the variables from Table 3, the entropy land-use mixture index, failed to enter the model. Household income was positively associated with total ridership, meaning once factors like size of catchment zone and fare levels are controlled for, the tendency is for residents of higher income households to patronize BART more. This no doubt reflects the high share of downtown San Francisco office workers who commute via BART each workday. And as expected, being a terminal or near-terminal station induces transit ridership once factors like catchment zone size are controlled for. (Terminal stations average relatively large catchment zones, so the controls are important here.) This suggests that the convenience of terminal stations to many suburban, exurban, and rural households that patronize BART to reach the core of the region leads to relatively high ridership levels.

Table 3. Regression Model Predicting Total Weekday Passenger Trips (Entries and Exits) at BART Stations, 1990

Dependent Variable: Total Weekday Passenger Trips (Entries and Exits) at Station, 1990

Variable	Coefficient	Standard Error	T Statistic	Probability
EMPLOYMENT DENSITY: Employees per acre within one mile radius of downtown stations and two mile radius of all other stations, 1990	143.476	17.435	8.229	.000
LAND AREA: Catchment area accounting for 90% of access and egress trips to station, in square miles, 1990	7.558	3.006	2.551	.016
LAND-USE MIXTURE: Entropy index of land-use mixture within one mile radius of downtown stations and two mile radius of all other stations, 1990[1]	-10261.721	4696.646	-2.185	.037
FEEDER BUS SERVICE: Total route miles of feeder bus services (within 200 feet of station) for station catchment area, 1990	166.198	32.085	5.180	.000
FARE TO SAN FRANCISCO CBD: One-way adult cash fare from station to Montgomery Street Station, 1990	-4375.408	1244.743	-3.515	.001
CONSTANT	17255.308	3365.033	5.128	.000

SUMMARY STATISTICS:

R^2 = .924
F = 67.953, prob. = .000
No. of cases = 34

Note:
[1] Entropy = $\{;_i[p_i{}^{*}\ln(p_i)]\}/\ln(k)$ where p_i = proportion of land area in land-use category i , and k = number of land-use categories. Ranges between 0 and 1, where 0 signifies land devoted to a single use and i signifies land area evenly spread among all uses.

Overall, the key land-use variable that was found to influence total ridership levels on BART was employment density immediate to the station. BART's busiest stations — Montgomery and Embarcadero stations — average 26,000-28,000 customers per day; these stations also have, by far, the highest employment densities, over 200 employees per acre.

The effects of employment density on total ridership levels were plotted for three different fare scenarios — $1, $2, and $3 fares to downtown San Francisco. These scenarios are shown in Figure 1. (All other predictor variables from Table 4 are set at their mean or median values — annual household income of $30,000, catchment area of 80 square miles, and 30 feeder bus route miles.) The plot clearly reveals that ridership rises with employment densities and falls with fares. Assuming a $2 fare, the model predicts around 12,000 weekday passenger trips in settings with 20 employees per acre (e.g., Walnut Creek BART station). At 150 employees per acre (e.g., Embarcadero station) and a $2 fare, the model estimates nearly 20,000 rail trips would be generated by a station.

Ridership levels vary even more with fare levels (as suggested by the elasticity estimates). At 80 workers per acre, a fare of $3 could be expected to produce around 11,500 daily passenger trips; decreasing the fare to $1, however, increases this figure to 26,000.

3.3 Passenger Trips per 1,000 Population

Indexing passenger volumes (entries and exits) by the population of the catchment area provides a more intuitive and transferable output. Table 5 reveals that, consistent with expectations, ridership per capita rises with both population and employment density as well as feeder bus service intensities.

The model reveals that, holding other factors constant, an increase of 10 workers per acre for a radius of one to two miles of a BART station increases the weekday passenger trips by 6.5 per 1,000 population within the ridership catchment. And according to the model, an increase in population density of 1,000 inhabitants per square miles adds an average of 8 more rail trips per 1,000 residents.

Table 4. Regression Model Predicting the Natural Log of Total Weekday Passenger Trips (Entries and Exits) at BART Stations, 1990

Dependent Variable: Natural Log of Total Weekday Passenger Trips (Entries and Exits) at Station, 1990

Variable	Coefficient	Standard Error	T Statistic	Probability
LOG of EMPLOYMENT DENSITY: Natural Log of employees per acre within one mile radius of downtown stations and two mile radius of all other stations, 1990	.2128	.0415	5.121	.000
LOG of LAND AREA: Natural Log of catchment area accounting for 90% of access and egress trips to station, in square miles, 1990	.1604	.0528	3.040	.005
LOG of FEEDER BUS SERVICE: Natural log of total route miles of feeder bus services (within 200 feet of station) for station catchment area, 1990	.2381	.0585	4.067	.000
LOG of FARE TO SAN FRANCISCO CBD: Natural log of one-way adult cash fare from station to Montgomery Street Station, 1990	-.6488	.1109	-5.850	.000
LOG of HOUSEHOLD INCOME: Natural log of annual household income for station catchment 1990	.2287	.1010	2.264	.032
TERMINAL: Terminal or near-terminal station (0=no, 1=yes)[1]	.2680	.1038	2.582	.016
CONSTANT	6.9979	.3300	21.206	.000

SUMMARY STATISTICS:
R^2 = .901
F = 40.637, prob. = .000
No. of cases = 34

Note:
[1] Near-terminal represents stations toward the end of a line that function like terminals because they are closer to freeways than actual terminals and thus serve a larger catchment area. BART's near-terminal stations, El Cerrito del Norte and Pleasant Hill, have larger supplies of parking than terminal stations since they are easier to reach by freeway.

Figure 1. BART Daily Trips by Employment Density, Sensitivity Test of Three Fares Levels for Trips to San Francisco CBD

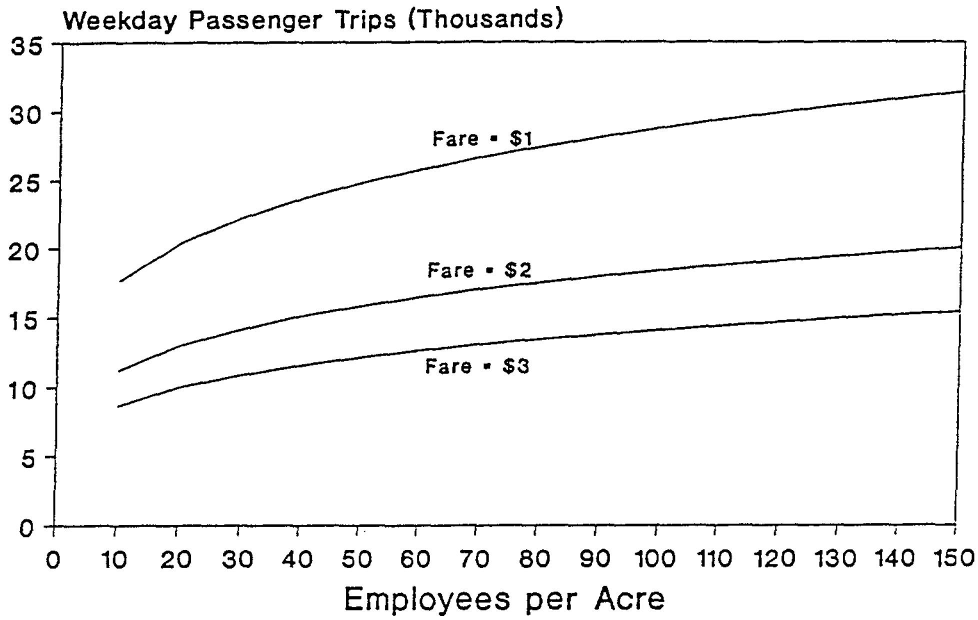


Table 5. Regression Model Predicting Weekday Passenger Trips (Entries and Exits) per 1,000 Population, BART Stations, 1990

Dependent Variable: Weekday Passenger Trips (Entries and Exits) at Station per 1,000 Population in Catchment Area, 1990

Variable	Coefficient	Standard Error	T Statistic	Probability
EMPLOYMENT DENSITY: Employees per acre within one mile radius of downtown stations and two mile radius of all other stations, 1990	.6545	.1814	3.608	.001
POPULATION DENSITY: Population per square mile within station catchment area, 1990	.0080	.0031	2.563	.016
FEEDER BUS SERVICE per 1,000 POPULATION: Total route miles of feeder bus services (within 200 feet of station) per 1,000 population of station catchment area, 1990	1.9849	.4358	4.555	.000
CONSTANT	5.799	13.060	0.444	.660

SUMMARY STATISTICS:

R^2 = .805
F = 41.354, prob. = .000
No. of cases = 34

3.4. Passenger Trips per 1,000 Population, Log-Log Model

Expressing relationships in log-log form produces Table 6. The model again shows that rail trips per capita rises significantly with employment density and feeder bus service intensity. In log-log form, two other significant predictors entered the model: catchment land area and household income. According to the model, the larger the catchment area, the lower the per capita transit-trip making. This likely reflects the influences of population density (which was a significant predictor in Table 5) — larger catchment areas generally average lower population densities, which are in turn associated with lower per capita ridership rates. Consistent with earlier findings, income appears to raise per capita rail trips, controlling for factors like employment density and feeder service intensity.

Interpreting regression coefficients as elasticities, Table 6 reveals that per capita rail travel was most sensitive to changes in income levels, followed by feeder bus service intensities, land area, and lastly, employment densities. All of these relationships are fairly inelastic. Moreover, all are statistically significant at around the .01 probability level. Additionally, the log-log model provided a slightly better fit in predicting per capita rail travel than the linear model.

The sensitivity of rail trips per capita to employment density is revealed in Figure 2. Three levels of feeder service intensity are shown in this graph — 0.25, 0.5, and 1.0 route miles of bus service per 1,000 population within the catchment. (The average value for the 34 BART station catchments was 0.25.) In generating these estimates, the remaining variables in Table 6 were set at the following mean values: catchment land area = 80 square miles and annual household income = $30,000.

The graph underscores the moderate sensitivity of ridership rates to employment density and the somewhat stronger sensitivity to changes in feeder bus service intensity. Assuming an average of one route mile per 1,000 catchment area population, Figure 2 projects that a station area with 120 workers per acre will produce 175 rail trips per 1,000 population, compared to 145 trips per 1,000 population at 20 workers per acre. With service levels at just 0.25 route miles per 1,000 inhabitants, ridership falls to 90 rail trips per 1,000 residents at densities of 20 workers per acre.

3.5. Passenger Trips per Square Mile

The last set of models predict ridership per square mile of station catchment area. Table 7 shows that the predictor variables that entered the model were identical to those in the model for per capita ridership levels (Table 6). Expressing ridership on a square mile basis, however, increased the predictive powers of the model considerably — raising the R-squared statistic to .911.

Table 6. Regression Model Predicting the Natural Log of Weekday Passenger Trips (Entries and Exits) per 1,000 Population, BART Stations, 1990

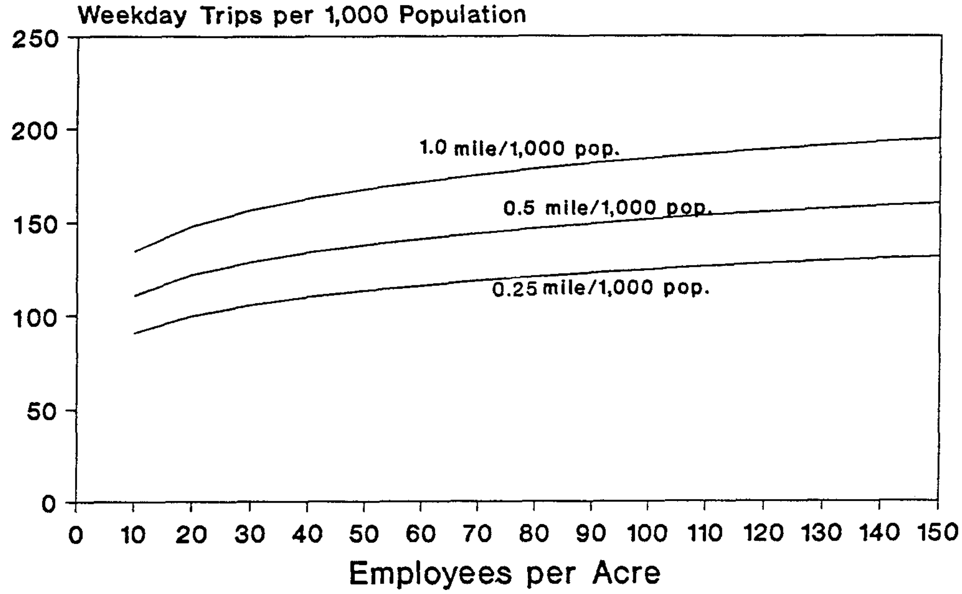
Dependent Variable: Natural Log of Weekday Passenger Trips (Entries and Exits) at Station per 1,000 Population in Catchment Area, 1990

Variable	Coefficient	Standard Error	T Statistic	Probability
LOG of EMPLOYMENT DENSITY: Natural Log of employees per acre within one mile radius of downtown stations and two mile radius of all other stations, 1990	.1351	.0498	2.717	.011
LOG of LAND AREA: Natural Log of catchment area accounting for 90% of access and egress trips to station, in square miles, 1990	-.2413	.0651	-3.710	.001
LOG of FEEDER BUS SERVICE per 1,000 POPULATION: Natural log of total route miles of feeder bus services (within 200 feet of station) per 1,000 population of the station catchment area, 1990	.2831	.0785	3.606	.001
LOG of HOUSEHOLD INCOME: Natural log of annual household income for station catchment, 1990	.3637	.1405	2.717	.011
CONSTANT	4.4128	.4421	10.465	.000

SUMMARY STATISTICS:

R^2 = .833
F = 36.170, prob. = .000
No. of cases = 34

Figure 2. BART Weekday Trips per 1,000 Catchment Area Population, by Employment Density and Bus Route Service Intensity



*Feeder Route Miles per 1,000 Population

Table 7 indicates that increasing employment density by 10 workers per acre raises the number of rail passenger trips in a catchment zone by 21.8 per square mile, holding other factors constant. Every 1,000 person increase in population density, moreover, adds around 53 rail trips per square mile of catchment zone.

Table 7. Regression Model Predicting Weekday Passenger Trips (Entries and Exits) per Square Mile, BART Stations, 1990

Dependent Variable: Weekday Passenger Trips (Entries and Exits) at Station per Square Mile of Catchment Area, 1990

Variable	Coefficient	Standard Error	T Statistic	Probability
EMPLOYMENT DENSITY: Employees per acre within one mile radius of downtown stations and two mile radius of all other stations, 1990	2.184	1.276	1.712	.097
POPULATION DENSITY: Population per square mile within station catchment area, 1990	.0529	.0214	2.476	.019
FEEDER BUS SERVICE per SQUARE MILE: Total route miles of feeder bus services (within 200 feet of station) per square mile of station catchment area, 1990	3.804	.5522	6.888	.000
CONSTANT	-133.134	64.111	-2.077	.047

SUMMARY STATISTICS:

R^2 = .911
F = 10.1699, prob. = .000
No. of cases = 34

8.0 Predictive Model of Passenger Trips per Square Mile, Log-Log Model

The final model predicts passenger trips per square mile in a multiplicative, or log-log, form. This produces a model with explanatory variables similar to those produced from other models, as shown in Table 8. The log-log estimate of trips per square mile was the best-fitting equation (R-squared = .944) and contained a balance of predictor variables related to employment and population density, feeder bus service intensity, and transit fares.

Table 8 reveals greater elasticity between rail ridership rates and density. Rail trips per square mile increased most rapidly with increases in catchment population density — every 10 percent increase in population density was associated with around a 5 percent increase in trips per square mile, holding all other factors constant. In fact, ridership levels expressed on a square mile basis were around twice as sensitive to population densities as they were to employment densities. There appeared to be a moderate degree of sensitivity to feeder bus service intensities and fare levels.

Figure 3 presents a sensitivity analysis for rail trips per square mile as a function of employment density, at three hypothetical fare levels. (All other predictor variables from Table 8 were set at their mean values.) These curves are steeper than those shown for Figure 2, indicating a greater sensitivity of ridership rates to employment densities when ridership is expressed on a square mile basis versus a per capita basis. From Figure 3, we see that at a fare of $2 per trip, the model estimates around 240 rail trips per catchment zone square mile at 100 workers per acre, compared to 160 trips per square mile at just 20 workers per acre. There are relatively greater shifts in ridership rates between the three hypothetical fares, however, underscoring the finding that fare elasticities are higher than employment density elasticities.

The final plot, Figure 4, highlights what was found in the log-log model — that ridership per square mile is most sensitive to catchment area population density. Visually, this is best seen by the relatively steep curves in the graph, much steeper than those generated by any of the previous graphs. Again assuming a fare of $2 (and setting all other predictor variables at their mean values), Figure 4 estimates that there would be nearly 200 trips per square mile for a station with a catchment zone averaging 4,000 residents per square mile; this compares to just 135 trips per square mile for a catchment zone with 2,000 inhabitants per square mile. And at 2,000 persons per square mile, if the average fare were to fall to $1, and all else remained equal, Figure 4 indicates ridership levels would rise again to around 190 rail trips per square mile of catchment area. The most fortuitous scenario for transit would be an average population density of 5,000 residents per square mile and an average fare to downtown San Francisco of $1 — a combination that, the model estimates, would produce over 300 weekday trips per square mile of catchment zone.

Table 8. Regression Model Predicting the Natural Log of Weekday Passenger Trips (Entries and Exits) per Square Mile of Catchment Area, BART Stations, 1990

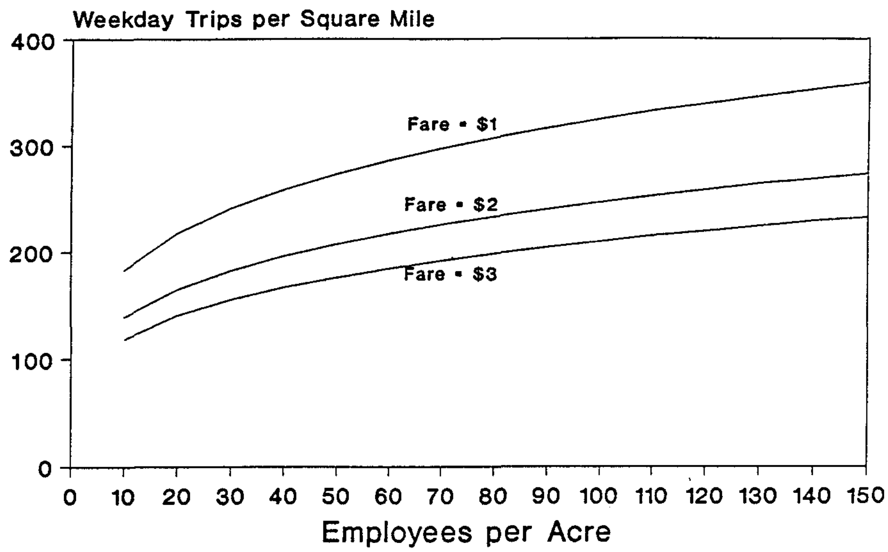
Dependent Variable: Natural Log of Weekday Passenger Trips (Entries and Exits) at Station per Square Mile of Catchment Area, 1990

Variable	Coefficient	Standard Error	T Statistic	Probability
LOG of EMPLOYMENT DENSITY: Natural Log of employees per acre within one mile radius of downtown stations and two mile radius of all other stations, 1990	.2489	.0701	3.549	.001
LOG of POPULATION DENSITY: Natural Log of population per square mile within station catchment area, 1990	.5154	.1695	3.041	.005
LOG of FEEDER BUS SERVICE per SQUARE MILE OF CATCHMENT: Natural log of total route miles of feeder bus services (within 200 feet of station) for station catchment area per square mile of catchment area, 1990	.3503	.1060	3.306	.003
LOG of FARE TO SAN FRANCISCO CBD: Natural log of one-way adult cash fare from station to Montgomery Street Station, 1990	-.3983	.2403	-1.658	.108
CONSTANT	1.169	1.486	0.787	.437

SUMMARY STATISTICS:

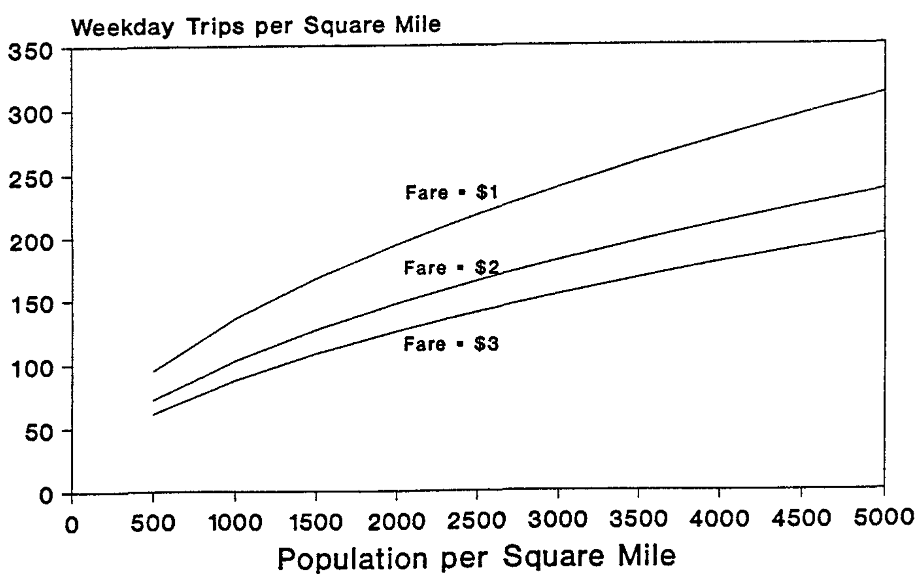
R^2 = .944
F = 121.592, prob. = .000
No. of cases = 34

Figure 3. BART Weekday Rail Trips per Square Mile of Catchment Zone, by Employment Density and Fare to San Francisco CBD



*Fare to San Francisco CBD

Figure 4. BART Weekday Rail Trips per Square Mile of Catchment Zone, by Population Density and Fare to San Francisco CBD



*Fare to San Francisco CBD

9.0 Conclusion

The demand models estimated for the San Francisco BART heavy rail system produced intuitive and consistent results that support accepted economic theories on transit ridership. Both population and employment densities were associated with higher ridership levels and rates. Ridership generally increased even more strongly as a function of feeder bus service intensities. High transit fares, on the other hand, significantly depressed ridership levels. Other, somewhat weaker, contributors to ridership increases were household incomes and the presence of terminal or near-terminal station. Mixed land uses, somewhat surprisingly, were found to lower rail ridership levels, though the relationship here was fairly weak and quite likely reflecting the influences of other, omitted variables.

In general, models presented in multiplicative, log-log form provided the best fit of data, and produce results in the most interpretative form. Additionally, expressing ridership on a per capita or per square mile basis provided output that has most transferability to other rail transit properties in the U.S. In combination with the model results of the main section of this report should provide useful benchmarks for analyzing the sensitivity of rail ridership to the built environment and other factors influencing transit demand.

APPENDIX B

This page left intentionally blank.

Comparisons of the National and BART Models With Chicago's Heavy and Commuter Rail

1.0 Introduction

This appendix analyzes the ridership demand for Chicago's CTA rapid rail and Metra commuter rail using the United States models developed in the main report and the BART models developed in Appendix A. Along with the San Francisco Bay Area, the Chicago area is being used for conducting in-depth analyses for several of the research topics. This comparison validates that general demand models work for specific systems and also shows some differences between these specific systems and the general models and BART. Some of these differences are expected since CTA is a rapid rail system and the demand models are only for light rail and commuter rail.

The two Chicago rail systems are

- Metra commuter rail which operates between the suburbs and the downtown Loop
- CTA rapid rail which operates within the city limits.

Full information on the Chicago data sources can be found in the Topic 3 and 4 reports. In order to compare the Chicago systems with the BART model for this analysis, catchment area sizes had to be estimated. Two of the Chicago data sources provide ways to estimate these catchment areas. Using land-use information (residential and employment densities and percent of land used for various purposes), the stations were classified into six clusters for each system (see Topic 3 report). Travel diary information provided the median home-end access distance traveled to each type of station. The median catchment area equals pi times the square of the median distance traveled to stations of the cluster type. Multiplying the catchment area by the population density (two mile density for commuter rail and half mile density for rapid rail) at the station gives the median catchment area population. The above catchment areas are for the median distance (50th percentile) as opposed to the 90th percentile of distance traveled used by Cervero. However, assuming there is a constant relationship between the radii of the two areas will yield equations which are equivalent.

A second possible way to determine catchment areas is from a park and ride survey performed by Metra. This information is only available for commuter rail stations with lots and only relates to automobile travel, but it provides an accurate look at the distance driven to the individual stations. The median distance is used again and the relatively few stations with no park and ride lot are assigned a catchment area of three-quarters of a mile. This provides the median catchment area and median catchment population.

The variables used in the analysis are listed in Table 1.

Table 1: Variable List

Variable	Definition	Log of Variable	
TOTON	Total station boardings	(LOGTOTON)	
ONSPERTP	Boardings per 1,000 people in cluster catchment	(LOGONPTP)	
ONSPERSM	Boardings per square mile in cluster catchment	(LOGONPSM)	
ONSPPRTP	Boardings per 1,000 people in p&r survey catchment	(LOGPRPTP)	Metra only
ONSPPRSM	Boardings per square mile in p&r survey catchment	(LOGPRPSM)	Metra only
RESIDEN2	Persons per acre within two miles of station	(LOGREDE2)	Metra only
HHLDDEN2	Households per acre within two miles of station	(LOGHHDE2)	Metra only
HHLDINC2	Median household income within two miles of stn.	(LOGHINC2)	Metra only
EMPLDENS	Employees per acre within half-mile of station	(LOGEMDEN)	
ENTROPY	Measure of mix between the land uses		
PARKING	Presence of park and ride station		
TERMINAL	Indicator if station is a terminal		
SOMEBUS	Presence of feeder bus service		Metra only
BUSMILES	Mileage of street served by bus within half-mile	(LOGBUSMI)	
DISTCBD	Distance in miles to CBD	(LOGDCBD)	
CBDLDCBD	DISTCBD*LOGDCBD (interacts well with LOGDCBD)		
DISTNEAR	Distance in miles to the next nearest station	(LOGDNEAR)	
CLCATCHA	Catchment area using station cluster	(LOGCLCAT)	
PRCATCHA	Catchment area using P&R survey	(LOGPRCAT)	Metra only

2.0 Comparison with National Model

This section of the report compares Metra to the national commuter rail equation and CTA to the national light rail equation. Metra was one of the systems used in developing that national commuter rail and provided about one-third of the data points. The models should therefore be similar. CTA is not a light rail, but a comparison shows whether ridership on CTA is based on the same or different factors.

The same variables are used here as in the national models except that the national models used employment figures for each CBD. This cannot be done for Metra and CTA because there is only one CBD making this a constant. Instead, the station level employment density is used in each equation. The employment measures are therefore not comparable. Also, the CTA estimation does not include an indicator of feeder bus service since bus service is available near all of the CTA stations. The national models were developed without the downtown stations; the same procedure is followed here.

Metra, No CBD, National CR List of Variables

DEP VAR:Log of Daily Boardings N: 204 MULTIPLE R: 0.558 SQUARED MULTIPLE R: 0.311
ADJUSTED SQUARED MULTIPLE R: .286 STANDARD ERROR OF ESTIMATE: 0.925

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-3.993	2.521	0.000	.	-1.584	0.115
Parking	0.855	0.257	0.222	0.789	3.327	0.001
Feeder bus	0.140	0.168	0.054	0.827	0.836	0.404
Log of HH income	0.681	0.242	0.239	0.488	2.821	0.005
Log of pop density	-0.091	0.158	-0.073	0.221	-0.576	0.565
Log of miles to CBD	0.772	0.315	0.402	0.131	2.452	0.015
Nonlinear distance	-0.005	0.004	-0.183	0.140	-1.158	0.248
LOG of emp. density	0.244	0.093	0.243	0.410	2.628	0.009

The signs of the significant coefficients are compared below.

	Parking	Feeder Bus	Household Income	Population Density	Distance to CBD	Distance to CBD (non-linear)	Employment
Metra	+		+		+		+
Nat'l CR	+	+	+	+	+	-	+

For Metra, as for commuter rail nationally, park-and-ride lots, higher household incomes, greater distances to the CBD, and higher employment near stations means more riders. Higher bus service levels and population densities do not increase Metra ridership although they do for the national commuter rail set. A non-linear measure of distance to the CBD is not significant for Metra in this equation although it was in other Metra models.

CTA, No CBD, National Light Rail List of Variables

DEP VAR:Log of Daily Boardings N: 127 MULTIPLE R: 0.508 SQUARED MULTIPLE R: 0.258
ADJUSTED SQUARED MULTIPLE R: .221 STANDARD ERROR OF ESTIMATE: 0.711

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	5.483	0.526	0.000	.	10.430	0.000
Terminal	0.804	0.265	0.293	0.663	3.038	0.003
Parking	-0.152	0.243	-0.059	0.686	-0.625	0.533
Log miles to nearest station	0.648	0.151	0.415	0.660	4.291	0.000
Log of miles to CBD	0.275	0.134	0.242	0.447	2.058	0.042
Log of pop density	0.238	0.095	0.231	0.723	2.499	0.014
Log of emp density	0.495	0.104	0.573	0.426	4.757	0.000

The signs of the significant coefficients are compared below:

	Terminal	Parking	Bus Service	Distance nearest station	Distance CBD	Residential Density	Employment
CTA	+		N/A	+	+	+	+
Nat'l LR	+	+	+	+	-	+	+

As with light rail, CTA ridership is greater at stations with higher residential and employment density. In addition stations at the end of the line or with wider spacings have higher ridership in both models. Parking is not significant for CTA, but it is strongly correlated with terminal stations as they are virtually the only stations with parking. Distance to the CBD has opposite signs in the two models. Distance to the CBD is positively related to ridership for CTA unlike in the national light rail model, where stations closer to the CBD have higher ridership after controlling for parking and terminal stations.

This analysis allows us to compare the elasticities of key land use variables from the national models with those for the Chicago systems as shown in Table 2. All of the elasticities are significantly less than one (i.e., a one percent change in the independent variable, say residential density, produces less than a one percent change in rail ridership). Of the density elasticities, CBD employment densities for the commuter rail systems have the most impact. A doubling of CBD employment density (100 percent increase) yields a 71.5 percent increase in commuter rail ridership. The employment density elasticity for the Metra system is much smaller, but it is measuring something quite different. The

Metra elasticities indicate that increasing employment densities near stations outside the CBD will increase ridership somewhat (a doubling of station area employment densities yields almost a 25 percent increase in ridership.) The elasticities for residential densities are higher for light rail systems than they are for commuter rail systems or for the CTA heavy rail system. For commuter rail in general and Metra commuter rail in particular, higher average household income does more to increase ridership than either measure of density. Household income however, was not a significant variable for light rail and, therefore, was not included in CTA analysis.

Table 2. Comparison of Elasticities of Key Land Use Varibles for United States and Chicago System Models

Variable	United States Commuter Rail	United States Light Rail	Metra Commuter Rail	CTA Heavy Rail
Residential Density	0.249	0.592	not significant	0.238
Employment Density*	0.715	0.400	0.244	0.495
Average Household Income	0.877	not significant	0.681	not included

* Employment density for the United States models is for CBD employment. Employment density for the Metra and CTA stations is for station area employment with CBD stations excluded from the analysis.

3.0 Comparison with BART Models

This section of the report compares ridership demand on the CTA and Metra rail systems with that on BART using the BART models developed in Appendix A. There are differences in the way some variable are defined and measured. In the BART analysis, catchment areas were based on the census tracts that produced 90 percent of a station's riders. This type of information was not available for the Chicago riders. Catchment areas are therefore estimated as explained earlier. Secondly, residential density for the Chicago stations is for a area roughly within 2 miles of the stations, while residential density for the BART analysis is for the entire catchment area which is generally much larger. The variable for bus service here only measures the number of miles of nearby service without distinguishing whether it is feeder or competitive service as the BART analysis did.

CBD stations are included in this analysis because they were included in the BART modeling. Some issues related to this are discussed in the conclusions.

Because many of the variables are defined and measured somewhat differently, the following comparisons should be made with caution. They are presented mainly to understand whether the systems are fundamentally the same or different, not to compare the differences in elasticities or other more precise measures. Only the signs of the coefficients are compared although the full regression results are shown as an attachment the differences in elasticities or other more precise measures. Only the signs of the coefficients are compared although the full regression results are shown as an attachment to this appendix. Tables compare the signs of the coefficients of significant variables for BART, Metra, and CTA. Variables significant at the 0.05 level appear as two symbols (i.e. + + or —), while those only significant at the 0.10 level appear as one symbol (i.e. + or -).

Ridership (Linear Model)

	Employment Density	Land Area	Entropy	Bus Service	Distance to CBD
BART	++	++	--	++	--
Metra	++				
CTA	++				

This model does a poor job of predicting Metra or CTA ridership. Only station area employment density is significant.

Ridership (Log - Log Model)

	Employment Density	Land Area	Household. Income	Bus Service	Distance to CBD	Terminal
BART	++	++	++	++	--	++
Metra	++	++	++	--	--	++
CTA		-	++			++

Results for BART and Metra are similar with this model and CTA partially follows the model. The Metra commuter rail and BART have the same set of significant variables for this equation; the only difference is a change in sign for bus service. This may be due to the Metra variable not differentiating between competing and complementary bus service or to differences in feeder bus services between the systems. For all three systems, ridership increases with household income. More riders get on at terminal stations in all three systems as well. Only three of the six variables—land area, household income, and terminal station—are significant in explaining CTA ridership. The same factors apparently do not explain ridership on the older, urban CTA heavy rail system and the newer BART system that extends out into the suburbs.

Riders per Thousand People (Linear Model)

	Employment Density	Population Density	Bus Service
BART	++	++	++
Metra	++	--	++
CTA	++	-	-

All three of the variables are significant for Metra, but population density has the opposite sign of BART. Curiously, bus service near Metra stations agrees in sign with the BART model in this case. This model apparently does not work well for CTA. It says that CTA ridership is inversely related to population density, the opposite result of most analysis for the CTA system.

Log of Riders per Thousand People (Log - Log Model)

	Employment Density	Land Area	Bus Service	Household. Income
BART	++	--	++	++
Metra	++	--	--	++
CTA	++	--		++

Again, the Metra system resembles BART. All of the BART variables are significant for Metra although bus service again has the opposite sign from BART. CTA has a reasonable match with BART as well, agreeing in significance and sign with all variables but bus service.

Riders per Square Mile (Linear Model)

	Employment Density	Population Density	Bus Service
BART	+	++	++
Metra	++	-	
CTA	++	++	

Metra again shows the opposite effect of population density from the BART, and this time CTA has the expected positive sign on population density. Bus service levels are not significant for Metra or CTA.

Log of Riders per Square Mile (Log - Log Model)

	Employment Density	Population Density	Bus Service	Distance to CBD
BART	++	++	++	--
Metra	++		-	--
CTA	++	++		

As usual, employment densities are positively related to ridership for all three systems. In this model, Metra ridership is not influenced by population densities, while CTA ridership has the expected positive sign. Both BART and Metra have declining ridership with distance from the CBD (because CBD stations are included in the analysis), but distance does not affect CTA ridership levels.

4.0 Conclusions for BART, Metra, and CTA Comparisons

As in the BART analysis in Appendix A, the multiplicative (log-log) models provide better results. The linear models sometimes produce results that are the opposite of what was expected or inconsistent with other analysis done with the Chicago data.

This comparison indicates that BART is more like the Metra commuter rail system than the urban CTA rail system. All of the variables from the BART model are significant and all but bus service have the same sign for Metra in the log-log models of ridership and riders per thousand population. In both of these models, there are fewer significant explanatory models for CTA.

Each of the analyses in Topic 1 measures employment density in a different way, yet in every case, employment density does matter. In the national models of the main report, employees per acre are only measured for the CBD. In the previous analysis of Metra and CTA system using the national models, employment density is measured for a 0.5 mile circle around each station—excluding CBD stations. In the BART analysis (and the Metra and CTA analysis using the same models)—employment density is measured around all stations whether inside or outside the CBD. The basic conclusion is that more jobs per acre in CBDs, near stations within CBDs, and in station areas outside CBDs all boost ridership on rail systems.

Overall, residential and employment densities do influence rail systems boardings, but other factors such as household income and characteristics of the rail system also are influential.

Attachment to Appendix B

Regression Results for Metra and CTA Using BART Models

Metra, CBD In, BART List (Boardings)

DEP VAR: TOTON N: 213 MULTIPLE R: 0.857 SQUARED MULTIPLE R: 0.734
ADJUSTED SQUARED MULTIPLE R: .727 STANDARD ERROR OF ESTIMATE: 2265.262

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	736.709	773.350	0.000	.	0.953	0.342
EMPLDENS	130.127	6.848	0.887	0.590	19.004	0.000
PRCATCHA	0.078	4.786	0.001	0.948	0.016	0.987
ENTROPY	-1150.379	1095.123	-0.039	0.956	-1.050	0.295
BUSMILES	-8.966	16.600	-0.032	0.364	-0.540	0.590
DISTCBD	15.168	19.166	0.038	0.549	0.791	0.430

CTA, CBD In, BART List (Boardings)

DEP VAR: TOTON N: 144 MULTIPLE R: 0.420 SQUARED MULTIPLE R: 0.176
ADJUSTED SQUARED MULTIPLE R: .146 STANDARD ERROR OF ESTIMATE: 2805.287

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	2346.473	2642.986	0.000	.	0.888	0.376
EMPLDENS	9.176	4.156	0.402	0.180	2.208	0.029
CLCATCHA	59.321	149.421	0.037	0.678	0.397	0.692
ENTROPY	-2067.642	2474.556	-0.072	0.797	-0.836	0.405
BUSMILES	9.974	28.905	0.088	0.093	0.345	0.731
DISTCBD	99.253	125.299	0.132	0.217	0.792	0.430

Metra, CBD In, BART List (Boardings)

DEP VAR:LOGTOTON N: 213 MULTIPLE R: 0.622 SQUARED MULTIPLE R: 0.387
ADJUSTED SQUARED MULTIPLE R: .369 STANDARD ERROR OF ESTIMATE: 0.992

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-7.555	2.269	0.000	.	-3.329	0.001
LOGEMDEN	0.344	0.098	0.347	0.305	3.511	0.001
LOGPRART	0.366	0.061	0.371	0.784	6.025	0.000
LOGBUSMI	-0.290	0.129	-0.207	0.350	-2.238	0.026
LOGDCBD	-0.540	0.139	-0.358	0.352	-3.896	0.000
LOGHINC2	1.393	0.211	0.437	0.682	6.614	0.000
TERMINAL	0.756	0.324	0.134	0.898	2.330	0.021

CTA, CBD In, BART List (Log Boardings)

DEP VAR:LOGTOTON N: 144 MULTIPLE R: 0.573 SQUARED MULTIPLE R: 0.328
ADJUSTED SQUARED MULTIPLE R: .299 STANDARD ERROR OF ESTIMATE: 0.716

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-0.013	1.811	0.000	.	-0.007	0.994
LOGEMDEN	-0.006	0.122	-0.011	0.108	-0.051	0.960
LOGCLCAT	-0.197	0.115	-0.167	0.521	-1.718	0.088
LOGBUSMI	0.041	0.135	0.032	0.445	0.305	0.761
LOGDCBD	-0.134	0.151	-0.176	0.124	-0.884	0.378
LOGHINC2	0.765	0.183	0.418	0.490	4.175	0.000
TERMINAL	0.721	0.249	0.234	0.752	2.895	0.004

Metra, CBD In, BART List (Boardings / Thousand People)

DEP VAR:ONSPPRTP N: 213 MULTIPLE R: 0.712 SQUARED MULTIPLE R: 0.507
ADJUSTED SQUARED MULTIPLE R: .500 STANDARD ERROR OF ESTIMATE: 404.633

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	7.734	51.739	0.000	.	0.149	0.881
EMPLDENS	11.729	1.297	0.607	0.525	9.044	0.000
RESIDEN2	-14.509	4.979	-0.180	0.619	-2.914	0.004
BUSMILES	6.124	2.893	0.166	0.383	2.117	0.035

CTA, CBD In, BART List (Boardings / Thousand People)

DEP VAR:ONSPERTP N: 144 MULTIPLE R: 0.562 SQUARED MULTIPLE R: 0.316
ADJUSTED SQUARED MULTIPLE R: .301 STANDARD ERROR OF ESTIMATE: 479.940

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	507.744	149.578	0.000	.	3.395	0.001
EMPLDENS	3.126	0.649	0.725	0.215	4.815	0.000
RESIDEN2	-5.605	2.906	-0.144	0.871	-1.929	0.056
BUSMILES	-6.111	3.170	-0.284	0.226	-1.928	0.056

Metra, CBD In, BART List (Log Boardings / Thousand People)

DEP VAR:LOGPRPTP N: 213 MULTIPLE R: 0.625 SQUARED MULTIPLE R: 0.391
ADJUSTED SQUARED MULTIPLE R: .379 STANDARD ERROR OF ESTIMATE: 1.243

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-10.744	2.799	0.000	.	-3.839	0.000
LOGEMDEN	0.404	0.105	0.323	0.414	3.839	0.000
LOGPRARt	-0.595	0.072	-0.478	0.866	-8.214	0.000
LOGBUSMI	-0.715	0.161	-0.403	0.353	-4.428	0.000
LOGHINC2	1.520	0.249	0.378	0.764	6.100	0.000

CTA, CBD In, BART List (Log Boardings / Thousand People)

DEP VAR:LOGONPTP N: 144 MULTIPLE R: 0.759 SQUARED MULTIPLE R: 0.576
ADJUSTED SQUARED MULTIPLE R: .564 STANDARD ERROR OF ESTIMATE: 0.922

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-4.008	2.330	0.000	.	-1.720	0.088
LOGEMDEN	0.383	0.092	0.411	0.315	4.174	0.000
LOGCLCAT	-0.619	0.140	-0.321	0.581	-4.434	0.000
LOGBUSMI	-0.060	0.159	-0.028	0.533	-0.375	0.708
LOGHINC2	0.794	0.220	0.266	0.565	3.618	0.000

Metra, CBD In, BART List (Boardings / Square Mile)

DEP VAR:ONSPPRSM N: 213 MULTIPLE R: 0.867 SQUARED MULTIPLE R: 0.752
ADJUSTED SQUARED MULTIPLE R: .749 STANDARD ERROR OF ESTIMATE: 1239.466

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-8.981	158.487	0.000	.	-0.057	0.955
EMPLDENS	70.760	3.973	0.847	0.525	17.812	0.000
RESIDEN2	-28.336	15.251	-0.081	0.619	-1.858	0.065
BUSMILES	6.128	8.862	0.038	0.383	0.691	0.490

CTA, CBD In, BART List (Boardings / Square Mile)

DEP VAR:ONSPERSM N: 144 MULTIPLE R: 0.653 SQUARED MULTIPLE R: 0.427
ADJUSTED SQUARED MULTIPLE R: .414 STANDARD ERROR OF ESTIMATE: 2802.288

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-1276.103	873.357	0.000	.	-1.461	0.146
EMPLDENS	15.062	3.791	0.548	0.215	3.973	0.000
RESIDEN2	66.696	16.966	0.270	0.871	3.931	0.000
BUSMILES	22.479	18.508	0.164	0.226	1.215	0.227

Metra, CBD In, BART List (Logs Boardings / Square Mile)

DEP VAR:LOGPRPSM N: 213 MULTIPLE R: 0.512 SQUARED MULTIPLE R: 0.263
ADJUSTED SQUARED MULTIPLE R: .248 STANDARD ERROR OF ESTIMATE: 1.290

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	6.595	0.696	0.000	.	9.472	0.000
LOGEMDEN	0.431	0.122	0.366	0.332	3.539	0.000
LOGREDE2	-0.149	0.154	-0.086	0.446	-0.967	0.335
LOGBUSMI	-0.365	0.183	-0.218	0.295	-1.991	0.048
LOGDCBD	-0.687	0.162	-0.383	0.437	-4.249	0.000

CTA, CBD In, BART List (Log Boardings / Square Mile)

DEP VAR:LOGONPSM N: 144 MULTIPLE R: 0.660 SQUARED MULTIPLE R: 0.436
ADJUSTED SQUARED MULTIPLE R: .420 STANDARD ERROR OF ESTIMATE: 0.925

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	5.314	0.985	0.000	.	5.393	0.000
LOGEMDEN	0.497	0.119	0.612	0.188	4.170	0.000
LOGREDE2	0.345	0.108	0.218	0.870	3.200	0.002
LOGBUSMI	-0.118	0.171	-0.064	0.465	-0.689	0.492
LOGDCBD	-0.137	0.182	-0.127	0.143	-0.754	0.452

APPENDIX C

This page left intentionally blank.

List of Light Rail and Commuter Rail Lines

Commuter Rail

city	line[1]	# of stations[2]	length[3]	boardings[4]
Boston	Attleboro	8	35.2	9388
Boston	Fitchburg	17	46.1	3811
Boston	Framingham	9	13.3	4682
Boston	Franklin	12	21.3	5603
Boston	Haverhill	13	28.4	3594
Boston	Ipswich	3	6.9	755
Boston	Lowell	7	20.0	4057
Boston	Needham Hts	9	8.7	3539
Boston	Readville	4	7.1	1028
Boston	Rockport	13	28.4	4630
Boston	Stoughton	2	3.1	1530
Chicago	BN	24	30.5	24867
Chicago	C&NW-N	24	45.0	13969
Chicago	C&NW-NW	20	56.1	19530
Chicago	C&NW-W	15	27.0	13880
Chicago	Elec-Blue Island	7	3.3	1130
Chicago	Elec-Main Line	23	22.2	16092
Chicago	Elec-S Chicago	8	3.9	3039
Chicago	Heritage	5	24.8	677
Chicago	Milwaukee-N	17	43.1	10425
Chicago	Milwaukee-W	21	33.9	11076
Chicago	RI-Main Line	13	30.4	9066
Chicago	RI-Beverly	12	5.8	6113
Chicago	South Shore	2	4.5	1752
Chicago	SWS	8	13.3	2869
Los Angeles	Orange County	8	77.9	1407
Los Angeles	Riverside	4	31.9	1224
Los Angeles	San Bernardino	10	43.7	2578
Los Angeles	Santa Clarita	8	71.9	2723
Los Angeles	Ventura	9	60.2	1639
Miami[5]	Tri-Rail	13	64.0	7279
Philadelphia	SEPTA-1235	5	4.9	2857
Philadelphia	SEPTA-2N	7	7.1	1483
Philadelphia	SEPTA-2S	16	20.7	3451
Philadelphia	SEPTA-3N	13	20.5	2921
Philadelphia	SEPTA-3S	14	11.8	4020
Philadelphia	SEPTA-5N	17	21.2	3824
Philadelphia	SEPTA-5W	23	38.8	12256
Philadelphia	SEPTA-6S	3	1.2	206

city	line[1]	# of stations[2]	length[3]	boardings[4]
Philadelphia	SEPTA-6W	12	14.4	1892
Philadelphia	SEPTA-7N	11	23.4	3738
Philadelphia	SEPTA-7W	10	5.1	2310
Philadelphia	SEPTA-8N	5	3.8	1928
Philadelphia	SEPTA-8W	10	5.8	1650
San Francisco	Caltrain	31	74.9	15073
Toronto[6]	6501E	9	20.0	13655
Toronto[6]	6501W	11	33.0	15878
Toronto[6]	6521	7	18.0	6621
Toronto[6]	6531	7	23.0	3943
Toronto[6]	6561	4	7.0	2561
Toronto[6]	6565	5	20.0	857
Toronto[6]	6571	7	19.0	1239
Washington DC	Brunswick	16	66.5	3155
Washington DC	Camden	10	30.5	1765
Washington DC	Penn	12	67.6	5840
Total		613	1470.1	307075

Light Rail

city	line[1]	# of stations[2]	length[3]	boardings[4]
Baltimore	MTA-N	7	8.1	3826
Baltimore	MTA-S	9	7.5	6207
Boston	Green-B	22	4.0	32979
Boston	Green-C	13	2.3	12727
Boston	Green-D	13	9.1	18124
Boston	Green-E	7	1.8	13451
Buffalo	Metro Rail	8	4.6	14440
Calgary[6]	NE	7	5.1	16180
Calgary[6]	NW	6	3.3	15140
Calgary[6]	S	7	6.4	18610
Cleveland	Blue	11	3.2	1440
Cleveland	Blue/Green	6	4.2	2550
Cleveland	Green	11	3.5	1350
Edmonton[6]	101	4	3.4	12274
Los Angeles	Blue	18	19.0	28360
Philadelphia	Media	33	8.3	3578
Philadelphia	Sharon Hill	16	3.5	1251
Pittsburgh	42	12	9.0	22818
Pittsburgh	47	8	10.9	5263
Portland	MAX	19	13.6	14460
Sacramento	RT-E	8	5.2	6542
Sacramento	RT-N	8	5.7	5328
San Diego	East	12	14.8	11982
San Diego	South	11	12.6	18554
St. Louis	MetroLink	9	8.6	11024
Total		285	177.7	298458

[1] Some lines have been broken into segments.
[2] The number of stations within each line segment
[3] The distance in miles from the furthest station in the segment to the nearest station in the segment.
[4] Total daily boardings of stations within the segment
[5] Data from Miami's rail system was not included in the demand analysis because it resembles an intercity line more than a CBD-oriented commuter rail.
[6] Data from Canadian rail system was not included in the demand analysis because employment and demographic data could not be obtained.

APPENDIX D

This page left intentionally blank.

Description of Data Collection and Processing for Demand Analysis

Five types of data were assembled with each station used as an observation:

1) Station identification information
2) Station ridership information
3) Transportation service characteristics
4) Population characteristics near station
5) CBD employment information

Each of the variables considered are identified in the following pages. Variables in italics are indicator variables, where a positive response is coded as a "1" and a negative response as a "0". Many other variables are included in the database, but the most significant are provided here to conserve space. Variables are also converted to the their natural logarithms. Variable names which appear in **bold** have logged equivalents which are used as well. The name given to the logged value appears in parentheses at the end of the variable description.

Information on the precise location of stations was determined from maps; travel times and service frequencies were determined from schedules. Follow-up calls to the transit operators generally filled in missing data.

1) Station identification information

The following variables help to uniquely identify each station, and also include some general classifying variables. This information is generally available from transit maps provided by the contacted agencies.

TYPE$	Coded 'LR' for light rail and 'CR' for commuter rail, to allow for different analysis between the two modes.
CITY$	Name of the central city on the line.
LINE$	Line that the station is located on. This is sometimes expanded to distinguish two different corridors. For example, Baltimore has two 'lines', one coded 'MTA-N' and one coded 'MTA-S'.
STATION$	Station name. Sometimes different sources give different names for the same station, but consistent names were adopted.
CANCODE	Is station in Canada? Allows comparisons between the countries.
INCBD	Is station very close to or inside the CBD? Stations where this is true are not included in the analysis. These are typically the CBD station for commuter rail systems and stations within a 'ride-free zone' for light rail systems. (If the classification is not straight-forward, light rail stations within one mile of the CBD and commuter rail stations within three miles are considered to be 'incbd'.)
DUPE	Has this stations already appeared earlier in the database? The database originally coded a station twice if two different lines pass through it. This variable identifies these stations so that analysis can also be done on unique stations.

2) Station ridership information

Boarding and alighting information was asked for by time of day and direction of travel. Only about half of the systems provided that level of detail. To use as many stations as possible, total boardings at each station are employed, regardless of time of day and direction traveled were used. Some of the data dated from the late 1980s while others are as recent as mid-1994. Most are the result of one-day counts. Some cases required slight computations to put the numbers into the proper form (i.e., divide total monthly boardings at a station by the appropriate factor).

INON	If known, daily inbound boardings; otherwise total daily boardings.
INOFF	If known, daily inbound alightings; otherwise, if known, total daily alightings.
OUTON	If known, daily outbound boardings.
OUTOFF	If known, daily outbound alightings.
TOTON	Total daily boardings. **(LOGTOTON)**

3) Transportation service characteristics

Since the means of reaching stations affects the ability to draw from a wider commutershed, data was collected for the amount and availability of parking at the stations and the amount and availability of bus services that can feed the stations from a broader area. Because the terminal station might be expected to draw from a larger area than other stations, this too is accounted for. Competition comes from other stations, either on the same line (next nearest station) or competing lines (competing rail lines), so a variable to reflect the presence of a nearby rail transit line is noted.

TIMECBD	Time in minutes to CBD. Some interpolation needed for the cases where the schedule only includes a few time points. (**LOGTCBD**)
DISTCBD	Distance in track-miles to CBD. (**LOGDCBD**)
DISTIN	Distance in track-miles to next inbound station. (**LOGDIN**)
DISTNEAR	Distance in track-miles to nearest adjacent station. (**LOGDNEAR**)
PEAKNUM	Number of inbound trains run through the station in the AM peak hour (7AM - 8AM). (**LOGPEAKN**)
DAYNUM	Total daily number of inbound trains run through the station. (**LOGDAYN**)
TERMINAL	Is this station the final one on the line? Yes = 1, no = 0
PARKING	Does this station have park-and-ride facilities? Yes = 1, no = 0
SPACES	Number of park-and-ride spaces available. (**LOGSPACE** - set to zero if no parking)
PARKRATE	Utilization rate of park-and-ride spaces.
SOMECOMP	Is there any competing service from nearby rail lines? Yes = 1, no = 0.
BUSSERV	Coded "0" for no feeder bus service, "1" for some service, and "2" if the station is a major transfer point.
SOMEBUS	Is there any feeder bus service? Yes = 1, no = 0

4) Population characteristics near station

Data was gathered from the US Census Bureau's 1990 Summary Tape File 3A (STF3A) which includes basic socio-economic data at a census tract level. Census tract maps were used to locate stations and two types of catchment areas were created around the station. The first is an oblong shape that extended two miles on either side of the station and two miles away from the CBD, but extended one mile towards the CBD. This area is believed to generally include about 70% of the trips to the stations and takes into account that people tend to travel toward the CBD to get to a station. The second catchment area is a simple half-mile radius around the station. Census tracts were then used to approximate this shape. However, because of the sometimes irregular shape of census tracts, it was often difficult to arrive at a good approximation of this half-mile ring. In rural areas, the tracts are quite large and the area included sometimes extends more than two miles from the station. Some judgment was exercised to decide whether to include tracts at the boundaries of these areas.

In downloading the STF3A files from a central database, some small gaps in the data were found. These gaps were typically filled with information from other sources or from typical averages. Canadian census information could not be obtained electronically. From a Canadian government depository, census tract maps were copied and the residential population and area of each noted. No additional socio-economic information was obtained at the Canadian tract level.

The relationship among population density, income and distance from the core of a metropolitan area bears upon the matter of transit ridership. While it is postulated that higher population density results in higher transit use, the lower incomes associated with higher population densities may also have a bearing, as might the fact that higher density areas tend to be closer to the core. In the case of commuter rail the high cost of this high amenity service suggests that an income variable be tested. The Census data's median household income for the two-mile commutershed is used.

POPTWO	Population over STF3A tracts within two-mile oblong area.
AREATWO	Acreage over STF3A tracts within two-mile oblong area. In theory this number should be approximately 6,500 acres, but the true acreage varies considerably.
DENSTWO	Population density over STF3A tracts within two-mile oblong area. (**LOGDENTW**)
POPHALF	Population over STF3A tracts within half-mile of station.
AREAHALF	Acreage over STF3A tracts within half-mile of station. In theory this number should be approximately 500 acres, but the true acreage varies considerably.
DENSHALF	Population density over tracts in half-mile of station. (**LOGDENHA**)
AVGCAR	Average cars per household of tracts in two-mile oblong area. (**LOGAVCAR**)
PCT01CAR	Percent of households in two-mile oblong area which own 0 or 1 car.
AVGHHINC	Average of median household incomes of tracts in two-mile oblong area. (**LOGAVINC**)

5) CBD employment information

Employment information for Central Business Districts is difficult to obtain with accuracy. Published sources of information are available only at a county level. Other sources which are reported on a small area basis, exclude workers not covered by social security, including government workers and the self-employed. A data vendor provided the "covered" employment at a ZIP code level. To address the matter of missing employees, state employment agencies were contacted to determine the percent of government workers in the Metropolitan Statistical Area, and CBD employment was adjusted upward to account for them. Typical percentages were in the 15 percent range, higher for state capitals.

The definition of the CBD proved to be a substantial problem, given the need to rely on ZIP code areas as the building blocks. To define the CBD, a combination of employment density ("covered" employees divided by gross area), the distribution of employees by industry type, and knowledge of the CBD by the researchers was used. The employment and area data for ZIPs was obtained for a rather expansive definition of ZIPs based on local knowledge. Then ZIP "candidates" were eliminated by examining the mix of industries, and the fall off of employment densities from the highest density ZIP. A ZIP was generally excluded if it increased CBD employment by less than half the percentage that it increased acreage if its inclusion resulted in the percent of the land area of the CBD. For example, at the margin, if a ZIP added 40 percent to the land area of the CBD, but only 15 percent to its employment, then the ZIP was not included. However, if the ZIP contained more than 25 percent of the CBD employment if it were included, it was generally retained. In borderline cases the industry mix was referred to; ZIPs with relatively high shares in the financial sector, usually associated with CBDs, suggested inclusion for that ZIP in the CBD definition.

Generally, ZIPs that were removed represented large land areas on the fringe of the CBD with a high share of heavy industry uses and/or residential areas. ZIPs in larger cities tend to be smaller, allowing the CBDs to be more precisely defined. The ZIPs included in each of the CBDs as defined for this study are listed on the next page.

Once the CBD was defined, both employment size (number of employees) and the employment density was calculated. Employment was factored up to account for the county-wide ratio of covered to total employment. Canadian employment information could not be located, making it impossible to include the three Canadian cities in the analysis.

TOTEMPS	Employment using CBD data factored for government employees. (**LOGTOTEM**)
EMPACRES	CBD ZIP code acreage. (**LOGEMPAC**)
EMPDENS	CBD employment density, in employees per gross acre. (**LOGEMPDE**)

List of Zip Codes Included in CBDs

Baltimore
- 21201
- 21202

Boston
- 02108
- 02109
- 02110
- 02111
- 02114
- 02116

Buffalo
- 14202
- 14203

Chicago
- 60601
- 60602
- 60603
- 60604
- 60606
- 60611

Cleveland
- 44114
- 44115

Los Angeles
- 90013
- 90014
- 90015
- 90017
- 90071

Pittsburgh
- 15219
- 15222

Philadelphia
- 19102
- 19103
- 19106
- 19107

Portland
- 97204
- 97205
- 97209

Sacramento
- 95814
- 95816

St. Louis
- 63101
- 63102
- 63103

San Diego
- 92101

San Francisco
- 94102
- 94103
- 94104
- 94105
- 94108
- 94111

Washington D.C.
- 20001
- 20004
- 20005
- 20006
- 20036
- 20500