SKETCH MODEL TO FORECAST HEAVY-RAIL RIDERSHIP

Len Usvyat
Parsons Brinckerhoff
1600 JFK Blvd.
Suite 700
Philadelphia, PA 19103
Phone: 215-209-1239
Email: lenusvyat@gmail.com

Linda Meckel
Parsons Brinckerhoff
1600 JFK Blvd.
Suite 700
Philadelphia, PA 19103
Phone: 215-209-1245
Email: Meckel@pbworld.com

Mary DiCarlantonio
Jeffrey Parker & Associates
1602 Kater Street
Philadelphia, PA 19146
Phone: 215-779-0747
Email: marydica@gmail.com

Clayton Lane
Email: claytonlane@gmail.com

1. EXECUTIVE SUMMARY

Ridership potential is a critical attribute in the feasibility stage of a transit project development. While general guidance exists to determine the transit mode most appropriate for various employment and residential densities, no simple sketch-level tool is available to predict the range of ridership values for a heavy-rail alignment. Sketch-level light and commuter rail ridership models occur in Transportation Research Record, No. 1986, titled "Sketch Models to Forecast Commuter and Light Rail Ridership, Update to TCRP Report 16." Current publication provides the third rail-based mode ridership estimation tool.

This analysis develops a sketch-level ridership forecasting tool for heavy-rail for medium and smaller size cities (those with population less than three million). It uses current ridership, demographic and transportation-system data from ten U.S. cities, including 32 heavy-rail corridors, and 474 stations, while it tests 186 possible explanatory variables. The efforts yield a linear multivariate regression equation that show close relationships between actual and predicted values with an adjusted R-squared value of 0.612.

The new sketch model for heavy rail can be used in place of a full-blown four-step modeling approach and requires only ArcGIS and Microsoft Excel. The data required for this model is readily available from Metropolitan Planning Organizations and/or the U.S. Census Bureau. Combining this heavy-rail model with the existing light rail and commuter rail models opens a wealth of opportunities for inexpensive methods to predict the ridership potential for rail-based development as well as the impacts on station ridership by new residential and commercial development close to stations.

2. INTRODUCTION

Ridership potential is an essential component of the transit project development process. The ability to estimate the ridership from a proposed transit development can either lead toward project implementation or eliminate the need for further analysis. Ridership projections also drive mode choice, frequency, vehicle type, and station spacing, which ultimately inform the process to derive capital and operating and maintenance costs.

However, ridership potential remains the least understood and often the most questioned aspects of project feasibility planning. The traditional four-step travel demand models tend to be costly, time-intensive and complex to understand and validate. Sketch-level approaches seem appropriate but may also be unavailable, particularly for heavy-rail corridor study or existing line extension.

A "Catch-22" exists: agencies do not typically have the money to undergo a full blown travel demand four-step model to check whether the potential ridership base exists for rail transit; yet, no simple model exists to guide this process. A study from the Transit Cooperative Research Program in 1990 entitled "Transit and Urban Form: Commuter and Light Rail Transit Corridors" (1) resulted in development of a multivariate linear regression model to predict light rail and commuter rail boardings. The update to this analysis for light rail and commuter rail modes appears in Transportation Research Record, No. 1986, titled "Sketch Models to Forecast Commuter and Light Rail Ridership, Update to TCRP Report 16" (2).

For heavy-rail, the seminal work by Pushkarev and Zupan (3) attempted to set minimum densities that support various types of transit. Pushkarev and Zupan took into account density of development at

both ends of the trip – residential densities at the start of the trip and downtown commercial densities at the end of the trip. Their work is the inspiration for sketch-level transit ridership potential models and is often used as a threshold to identify the feasibility of an area to support new transit modes.

While the Pushkarev and Zupan work is the starting point for an analysis for predicting heavy-rail patronage, no specific sketch model exists to date to predict the degree of heavy-rail ridership potential. In this publication, the third in a series of sketch models using current U.S. Census Bureau and ridership data to predict rail-based ridership—this time for heavy-rail.

The three resulting models from these two publications can help agencies or independent consultants develop an inexpensive ridership estimation tool to determine the number of boardings that can be achieved from within a certain corridor where light rail, commuter rail, or heavy-rail are alternatives. While Pushkarev's diagram can be the source for providing guidance for what mode to utilize, these three new tools can help show the potential ridership that can actually be achieved. These models are designed to be used in medium and smaller size cities and generally not recommended to be used in the extremely large American cities such as New York and Chicago.

3. LITERATURE REVIEW

The Federal Transit Administration's Section 5309 New Starts program mandates the development of ridership forecasts in order to qualify for federal funding for new transit projects. To obtain these forecasts, agencies have utilized models employing various indicators and at differing scales of applicability (4-17). Currie et al. (18) predicted transit ridership nationally based on different world events and gas prices, comparing the ridership in the US with that of other countries. Currie et al. indicated that trip frequencies are primarily a function of socioeconomic characteristics; trip frequencies are also a function of the built environment. Other authors have explored various other indicators including land use, neighborhood type and neighborhood category. Ewing et al. (19) engaged in an extensive literature review of 50 empirical studies analyzing the effect and elasticities of trip frequencies based on different outcome variables, types of neighborhood and activity center designs, land use patterns, transportation networks, and urban design features. Characteristically different built environments, such as auto-oriented vs. transit-oriented, are analyzed, with many studies occurring with a small sample size and at the neighborhood or county geographic scale. None of the literature reviewed in this study was nationally comprehensive.

Few nationally applicable ridership estimating tools have been developed. Most sketch ridership estimating tools have been applied to specific corridors and regions for which they were uniquely designed (2). Further, Cervero (20) explains the current and common use of a four-step travel model as not robust at the neighborhood scale, making the application of a four-step model to a neighborhood-sized area an unreliable method of estimating ridership because fine-grained features, such as certain design and land-use existing at the neighborhood scale which influences ridership, are often overlooked. Cervero further illuminates the strength of using sketch-planning tools as they "may do a better job of picking up some of the nuanced relationships between smart growth and travel demand" (20). These models have produced statistically better relationships between indicators and more accurately explain influences on ridership (2, 20).

Previously, for light rail and commuter rail modes, predicted ridership showed close relationships with actual values when adjusted R-squared values were applied (2). Prior to this study on light rail and commuter rail in 2006, two other nationally applicable models saw strong relationships when using demographic and transportation variables to explain observed ridership at different stations (1, 3).

In 1982 Pushkarev et al. used indicators such as population density, geographic population distribution, auto ownership, transit orientation, downtown size, radial line length, auto speed, and mode of access in a model that was used to make the case for a national policy to expand urban passenger rail systems in many U.S. cities (3). In the cities that did develop new systems, retrospective research validates that their assessments was generally correct. Pushkarev et al. further examined the differences between transit systems, such as light rail, and commuter rail, identifying the characteristics of cities in which each system worked best. Demery and Setty (21) also demonstrated that the research completed by Pushkarev et al was generally correct and that using sketch ridership models is a reliable forecasting tool for transit feasibility planning.

The rise of reverse commuting based on changes in settlement patterns and travel patterns since 1980 was not included in Pushkarev et al. or the research completed by Demery and Setty (2), therefore updating these studies to reflect changing demographics is desirable. The 1996 *TCRP Report 16* utilized demographic and transportation data from the early 1990s and incorporated station specific variables into the model (PB – etc 1996). This model was successfully applied to feasibility studies in Philadelphia (22, 23), Baltimore (24), Phoenix, Charlotte (25), Wilmington (26), Richmond (27), and Dayton, as well as numerous corridor-specific models that utilized similar variables (4, 5, 12, 13). However, each area specific model has limited application outside the project for which each was developed, highlighting the need for an updated nationally relevant model.

The calibration data derived from the *TCRP Report 16* was improved on by Lane et al. in 2006. This analysis improved technically upon the national model yielding higher R-squared values for commuter rail (from 0.64 to 0.84) and light rail (from 0.15 to 0.47) modes of transit (2). Application of the model to heavy-rail has not yet been explored. Lane et al. further identified variable and issues to be improved upon, these remain:

- Central business district (CBD) employment,
- CBD density,
- Household income,
- Population,
- Reverse commuting behavior,
- Special transportation centers,
- Fares, and
- Changing travel relationships since the early 1990s.

Further obstacles observed from Lane et al. include the difficulty of subjectively defining the downtown as well as how to approach station catchment areas that sometimes double or triple count the population. These two issues were addressed by Lane et al. The Central Business District is defined by contiguous zones with employment densities within two standard deviations of the mean being included and adjacent zones employment densities within 1.5 standard deviations. Exclusive and nonexclusive station catchment areas were defined by dividing data for overlapping areas equally between stations using a spatial mapping program (2).

The Florida Department of Transportation has developed a bus sketch planning tool called TBEST that predicts patronage by stop. TBEST now is required for use by all transit systems in Florida. The Los Angeles MTA just has adopted it as well. A rail version of such tool is not yet available but a potential of using the data in this publication for development of TBEST for Rail maybe appropriate.

Extrapolating this study to heavy-rail will attempt to illuminate unique indicators for this mode of transit. In general, previous models have tended to underestimate ridership at stations adjacent to special generators such as airports, hospitals and malls. The purpose of this paper is to develop a heavy-rail specific model that addresses station-area specific variables.

4. METHODOLOGY

A multivariable linear regression model was created for U.S. heavy-rail modes to predict ridership using information from public sources such as the U.S. Census Bureau, MPO-specific demographic data, and transit-related system characteristics data. The primary analytical tools included ArcGIS and Maptitude to portray geographic information, SPSS statistical software to analyze the data using linear and non-linear regression models, and Microsoft Excel to portray the results.

4.1 Data

Current ridership data (2004-2006) was collected for all 474 U.S. heavy-rail transit stations. In addition, demographic information for both the areas surrounding the stations areas as well as for the entire metropolitan area at large was collected. The transit service characteristics for each transit line evaluated were also examined. Table 1 shows full details of the systems and stations studied.

Table 1. Heavy-rail lines used in the analysis

Region	Agency	Heavy-Rail Line	Stations	Non-CBD Stations
Baltimore	MTA	Metro	17	12
Boston	MBTA	Blue Line	51	45
		Orange Line		
		Red Line		
Chicago	CTA	Blue	143	97
		Brown		
		Green		
		Orange		
		Pink		
		Purple		
		Red		
		Yellow		
Cleveland	GCRTA	Red Line	18	17
Los Angeles	MTA	Red Line	16	14
		Purple Line		
Miami	County Transit	Metrorail	23	20
New York	PATH	Hoboken-33rd Hoboken WTC	13	5
		Journal Sq. - 33rd Newark-WTC		
		Newark-WTC		
Philadelphia	SEPTA	Broad Street	51	47
		Broad Street & Spur		
		Market Frankford El		
	DRPA	PATCO	13	9
San Francisco	BART	Dublin-SFO	43	39
		Fremont-Richmond		
		Pittsburgh-Daly City		
		Richmond Daly City		
Washington, DC	WMATA	Blue	86	76

	Green Orange Red Yellow

The primary source for all demographic data was regionally-approved socioeconomic data adopted by each respective Metropolitan Planning Organization. Where needed, this data was supplemented with the information from the U.S. Census Bureau's Census Transportation Planning Package (CTPP) for the year 2000. However, MPO data was preferred as most MPOs have future population and employment projections and also conduct analysis of existing and future transit patronage.

Most heavy-rail systems exist in large urban areas where socioeconomic and transportation system usage data is available and frequently updated. Data was collected from ten cities from a total of 32 heavy-rail lines and 381 non-CBD stations. Cities included were Baltimore, Boston, Chicago, Cleveland, Los Angeles, Miami, New York (PATH train), Philadelphia, San Francisco, and Washington, DC. Atlanta's MARTA system was omitted given it's been around in its current alignment less than 15 years. Also, New York's subway system was excluded from the analysis due to its unique transit ridership patterns.

The rail systems studied provided diverse enough data for use for this project. Because heavy-rail systems generally exist in cities with a metropolitan population of at least one million, the model's applicability is best directed at larger metropolitan regions. Other guidance on applicability is discussed later in Section 6.2

4.2 Tested Variables

186 variables were collected and tested, most of which were directly obtained and some derived based on the collected data. The variables used in the model were result of authors knowledge of driving factors for ridership. The independent variables consisted of information related to:

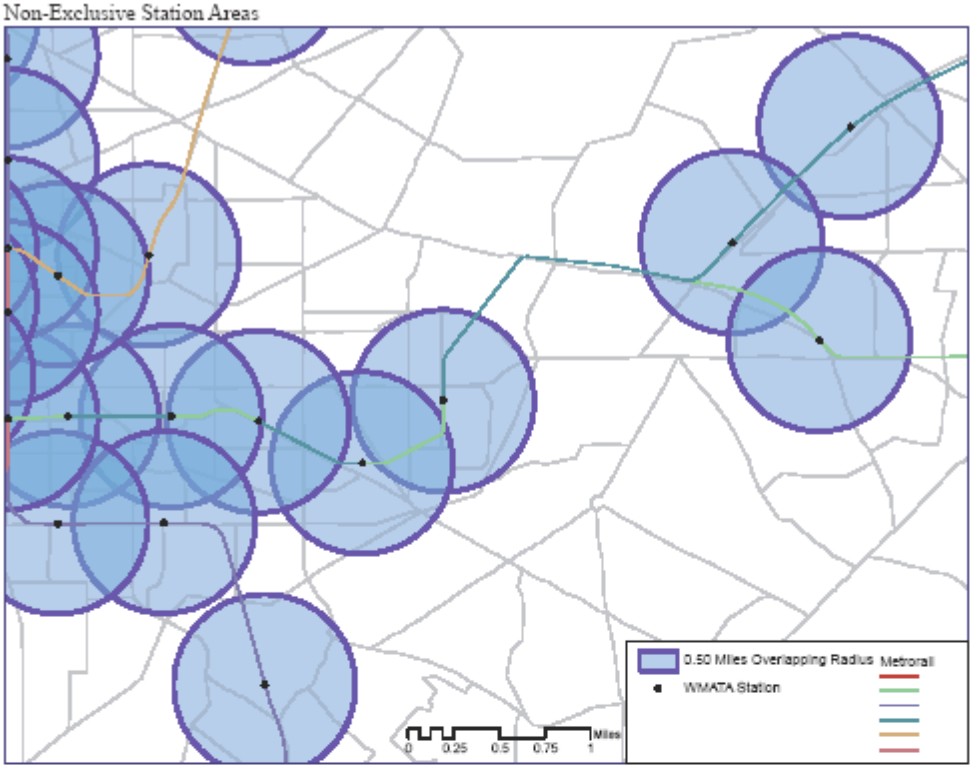
- demographic data around station areas
- demographic data on the CBD served by the heavy-rail system
- demographic data on the metropolitan region
- demographic data relating to the characteristics around the entire heavy-rail system
- general information about the heavy-rail system overall
- other station specific information
- service characteristics of the heavy-rail system in question

Station Area Demographics

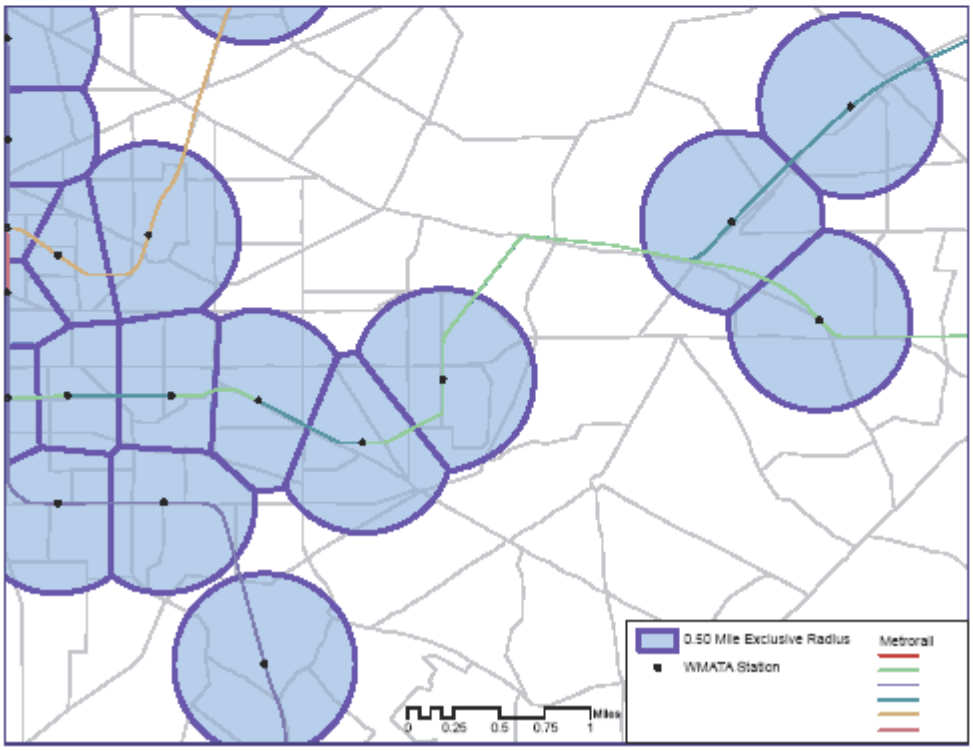
Similar to the previous publication on light rail and commuter rail ridership, station-area demographics were measured for radii of ¼, ½, 1, and 2 miles with exclusive geographies. The previous publication specifically demonstrated the difference between exclusive and non-exclusive demographic regions (2). It was determined that exclusive regions perform much better when predicting station boardings because this method avoids double counting of population and employment around station areas. Figure 1 demonstrates the example of exclusive versus non-exclusive station areas.

Figure 1. Exclusive and Non-Exclusive Station Areas

Non-Exclusive Station Areas



Exclusive Station Areas



The variable categories follow:

- Employment (all jobs)
- Population
- Population divided by employment

- Population 16+ years old
- Average household size
- Median household income (1999 dollars)
- Average number of vehicles per household
- Zero-car households
- Households with cars
- Zero-car households divided by households with cars
- One-car households
- Two-car households
- Three-car plus households

In addition to demographic data in various radii from the station, buffers of the same demographics were computed—for example, the population between 0.25 to 0.5 miles, population between 0.5 and 1 miles, etc. The above variables were used in computation of demographic data for the following buffers:

- 0.25 – 0.5 miles
- 0.5 – 1 miles
- 1 – 2 miles

Station-Specific Transportation Attributes

- Whether a station is a secondary downtown (yes or no)
- Availability of a feeder bus (yes or no)
- Number of buses connecting to the station (i.e. number of individual bus route numbers within direct proximity to the station)
- Connection to other rail system (yes or no)
- Number of rail lines connected to this station
- Availability of parking (yes or no)
- Whether it is a terminal station (yes or no)
- Indication if the station is a special transit attractor, such as airport, major university, major hospital, or transit center. This delineates any transit station that would derive its ridership from any source not identified in population or employment around the station (yes or no)
- Cash fare to downtown (yes or no)
- Commuter fare to downtown
- AM peak headway (in minutes)
- PM peak headway (in minutes)
- Midday headway (in minutes)
- Weekend headway (in minutes)
- Hours transit system is operated per 24-hour period
- Distance to primary downtown (miles)
- Time to primary downtown (minutes)
- Speed to primary downtown
- Distance to secondary downtown (miles)
- Time to secondary downtown (minutes)
- Speed to secondary downtown
- Distance to nearest station (miles)

While the primary downtown is defined by Lane et al. in 2006 (2), the secondary downtown is defined more generally. Knowledge of the urban geographies was used to define what is considered a

secondary downtown. In Philadelphia, for instance, the City Hall station area is considered a "downtown" station while other stations on Market Street are considered secondary downtown stations.

Corridor demographic characteristics

All corridor demographic variables are measured within 2-mile radii from all station on the heavy-rail network

- Zero-car households divided by households with cars
- Households with cars along the corridor
- Zero-car households along the corridor
- Total employment along corridor
- Total population along corridor
- Total employment divided by total population along the corridor

Metro area demographic characteristics and transportation attributes

- Median household income in the metro area (1999 dollars)
- Vehicles per household in the metro area
- Central Business District (CBD), as defined in by Lane et al. in 2006 (2)
    - CBD area (in square miles)
    - CBD area divided by total metro area
    - CBD density (employees per square mile)
    - CBD employment (all types of jobs)
    - CBD employment divided by total metro area employment
    - CBD employment density (jobs per square mile)
- Metropolitan area (as defined by boundaries of the MPO)
    - Metro area (square miles)
    - Metro employment
    - Metro population

5. ANALYSIS

The correlation of each of the independent variables with the dependent variable was tested. Correlations were done only for variables where a logical relationship can be found. For each dependent variable, both the average daily boarding itself and the natural logarithm of average daily boardings were correlated. In addition to independent variables, the natural logarithms of the independent variable were tested with average daily boardings and natural logarithm of average daily boardings. Using the natural logarithm is an effective method for addressing variables not normally distributed or those with highly extreme values.

Strong correlation between explanatory variables was undesirable while correlation with the dependent variables was certainly desirable. The Person Product Moment was used to test possible linear correlations and Spearman's nonparametric coefficient was used to test possible non-linear correlations.

Mostly, stronger correlations occurred with the dependent variable when it was expressed as the actual value, not as a natural logarithm of average daily boardings. This is in contrast to the findings with the light rail and commuter rail boardings. 77% of the 13 variables with the highest correlation coefficient

with the dependent variable show that the actual boardings result in better correlation with independent variables than the natural logarithm of actual boardings. Because we applied this model to medium and smaller size cities, it is logical that actual value of boardings is better correlated with independent variable as opposed to normalized, natural logarithm, value.

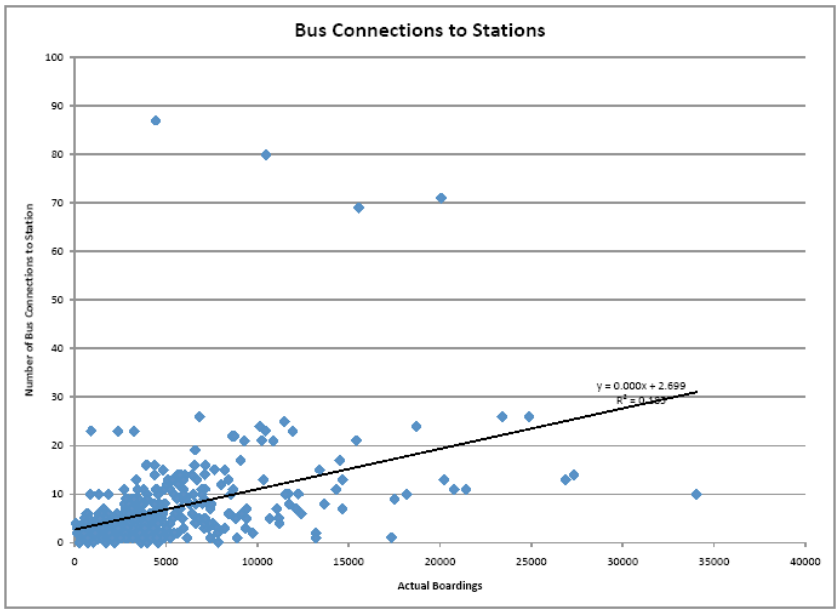
Similar to the prior publication, the correlation analysis also revealed that assigning geography exclusively to a single nearest station, rather than to all nearby stations, yields a considerably better fit with the dependent variable. Non-overlapping areas not only eliminate double counting of the population and employment but also result in more precise definition of the station area catchments. Stations located closer to each other may result in much smaller catchment areas and hence fewer boardings.

Linear multivariate models were used in the analysis because we expect that boardings are a direct result of the demographic data. We expect that an increase in number of people drives the increase in boardings in a certain 1-to-1 or 1-to-0.5 relationship and we expect that this relationship is linear. It could be argued that this is not the case and the relationship is curvilinear. While we tested basic logarithmic dependency structures, we did not test other potential non-linear relationships.

Generally, employment ended up being the best predictor of the actual boardings while transit service characteristics were the best predictors of the natural logarithm of boardings. In the analysis of light rail and commuter rail, however, service characteristics were consistently better predictors than actual demographic information.

Employment around the station area is by far the best predictor of heavy-rail boardings. The number of buses connecting to the station is also an important predictor of boardings. This is expected, as the number of buses is often a proxy for a major transit hub. As the model in the following section suggests, every extra bus connecting to a station can bring on as many as 70 extra boardings per station. Figure 2 shows a scatter plot of the actual boardings and the number of buses connecting to each station.

Figure 2. Number of Connecting Buses and Actual Boardings at Stations



The ratio of employment-to-population around the entire transit line is also an important predictor—higher employment to population ratio within two miles of the transit line suggests that more people would be using the transit line.

Midday headway is also an important predictor where stations with longer headways have fewer boardings. The strong correlation of headway with boardings raised some concern about the direction of causality. Headways can both affect and be affected by demand since schedulers generally calibrate peak-period service headways to match demand. Because heavy-rail headways are much more likely to be guided by agency policy, headway was left as part of the model.

Distance-to-downtown had a strong positive relationship to boardings as well. This suggests that heavy-rail is most often used as a substitute for automobile for work-related trips. Consequently, it is likely that more boardings would occur further away from downtown where fewer modes of reaching the downtown exist. This is also why parking is an influential variable in predicting boardings.

It is observed that existing research demonstrates differences in demographic characteristics in areas directly adjacent to heavy-rail stations versus those located within 0.5 or one mile away. Often, area in direct vicinity to the train station is less desirable because of the noise or heavy retail-oriented land uses. It can be noted that differences do in fact exist, particularly between population within 0.25 and population from 0.25 to 0.5 miles of the station (or the buffer region). Figure 4 demonstrates the geographic information for Washington Metropolitan Area Transit Authority (WMATA) while Figure 3 actually shows the relationship of the two population groups. It is apparent that a stronger positive relationship exists between boardings and the population within the buffer than the relationship between boardings and the population within 0.25 miles of the station.

Figure 3. WMATA Stations Showing 0.25, 0.5-mile Radii With a Buffer of 0.25 to 0.5 Miles

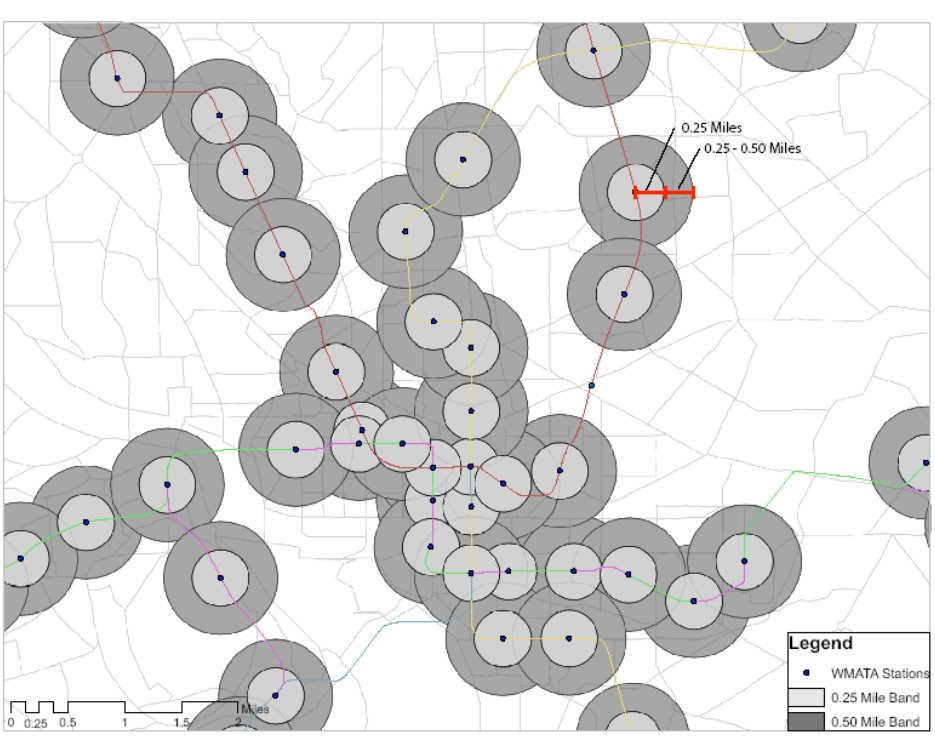
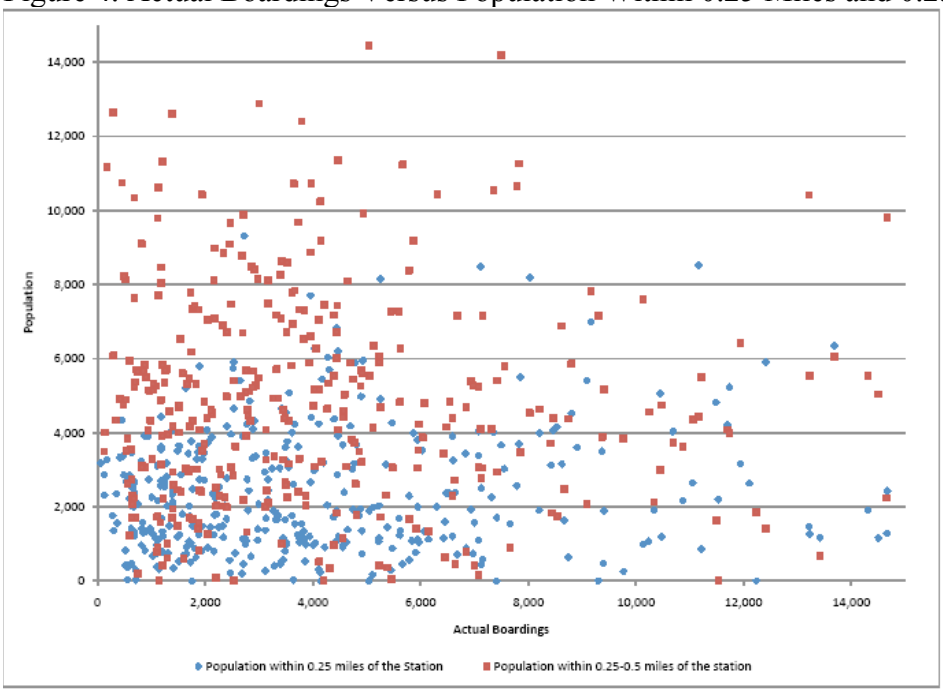


Figure 4. Actual Boardings Versus Population Within 0.25 Miles and 0.25-0.5 Miles of the Station



Numerous combinations can be tested to determine possible relationships. In addition, hundreds of iterations of the model were run to come up with the most reasonable, best fit, and logical predictive model. CBD stations were eliminated from the analysis as their boardings are so heavily driven by employment only comparing them in the same model with residential stations would produce inapplicable results.

Many variables are collinear and were removed from the model. Typical examples of collinearities are cases where one variable is a subset of another. For instance, population over 16 years old and total population is collinear with each other. Similarly, percent of households with cars and average number of cars per household are collinear as well. Income and number of cars per household is a typical example of collinear variables. To address these issues, we looked at variables that "don't overlap." For instance, we look at total population and percent of population over 16 years old.

When testing various models, those with fewer variables were more desirable because less effort is required to apply the final product. Thus if a variable improved the model's fit very modestly, it was not considered for of inclusion. Likewise, variables that were easier to measure generally were favored.

6. RESULTS

6.1 Sketch Models

The final model for heavy-rail system is shown below. Average daily boardings are predicted at 381 heavy-rail stations.

Heavy-rail station boardings = -971.7
+ 1625.1 * [If this is a terminal station, 0 if not]
+ 1345.9 * [If this station is a secondary downtown, 0 if not]
+ 1709.7 * [if this is a special transit attractor station, 0 if not]
+ 69.5 * [number of buses connecting to this station]
+ 883.6 * [if there is parking available, 0 if not]
+ 2270.9 * [if there is connection to other rail modes, 0 if not]
+ 115.4 * [distance to downtown, in miles]
- 2791.8 * [ln(midday headway in minutes)]
+ 0.024 * [CBD density, in employees per square mile]
+ 0.224 * [employment within 0.25 miles of the station]
+ 0.133 * [employment within 0.25 to 0.5 miles of the station]
+ 0.219 8 [population within 0.25 to 0.5 miles of the station]
+ 5938.1 * [employment within 2 miles of entire line div by population]

Table 2 shows the statistical performance of the model. The heavy-rail data yields an model with an adjusted R-square of 0.619. Coefficients for all variables are statistically significant with $p<0.05$ and minimal collinearities. All variables also have coefficients that are intuitively correct in their direction.

Table 2. Model Output

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta	B	Std. Error
(Constant)	-971.670	1583.272		-.614	.540
If this is a terminal station, 0 if not	1625.102	537.285	.107	3.025	.003
If this station is a secondary downtown, 0 if not	1345.904	525.708	.094	2.560	.011
if this is a special transit attractor station, 0 if not	1709.663	723.027	.077	2.365	.019
number of buses connecting to this station	69.492	19.627	.134	3.541	.000
if there is parking available, 0 if not	883.643	410.786	.090	2.151	.032

if there is connection to other rail modes, 0 if not	2270.988	498.359	.150	4.557	.000
distance to downtown, in miles	115.447	41.213	.118	2.801	.005
[ln(midday headway in minutes)	-2791.773	532.421	-.206	-5.244	.000
CBD density, in employees per square mile	.024	.004	.195	5.439	.000
employment within 0.25 miles of the station	.224	.028	.320	8.074	.000
employment within 0.25 to 0.5 miles of the station	.133	.022	.207	5.979	.000
population within 0.25 to 0.5 miles of the station	.219	.042	.176	5.180	.000
employment within 2 miles of entire line div by population	5938.143	787.853	.259	7.537	.000

Figure 5 presents the results of predicted versus actual boardings by station for all 381 stations. The line "Actual" represents where actual boardings are equal to predicted boardings, i.e. it shows a 1:1 relationship between actual and predicted boardings. It is clear that the relationship of the actual and predicted is strong and positive. The primary intention of the model is to look at multiple stations as the model is designed to evaluate proposed lines or heavy-rail extensions. When the results are aggregated to the route level or route section, the models perform appropriately as errors at the station level are mitigated, producing an R-squared of 0.702 as shown in Figure 6. On a city-wide basis, the model performs even better with an R-squared of 0.814 as shown in Figure 7.

Figure 5. Predicted Versus Actual Boardings by Station

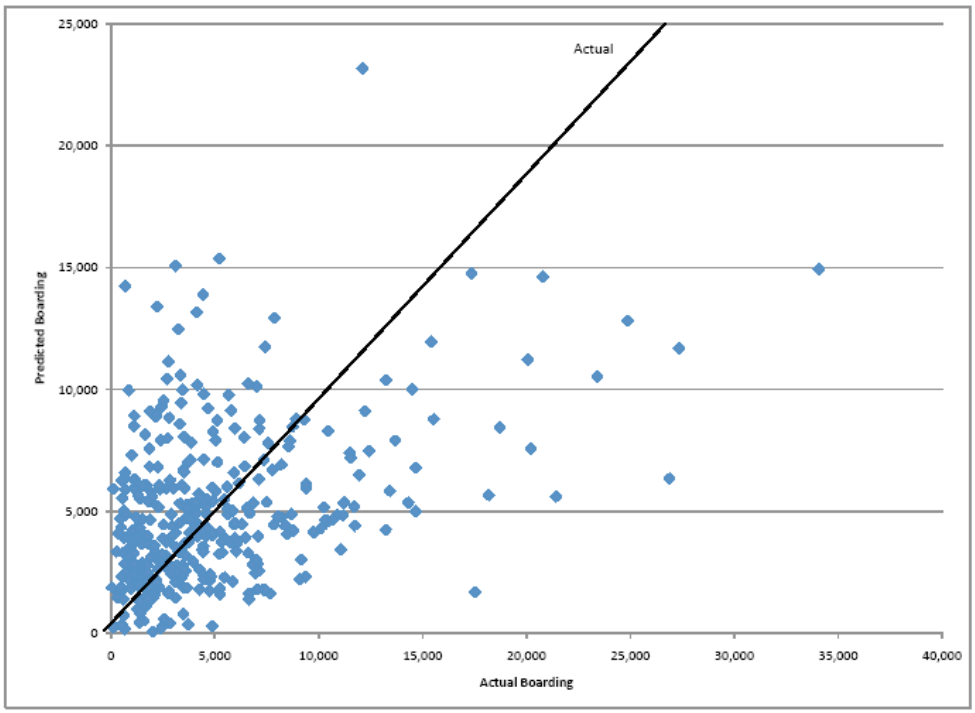


Figure 6. Predicted Versus Actual Boardings by Line

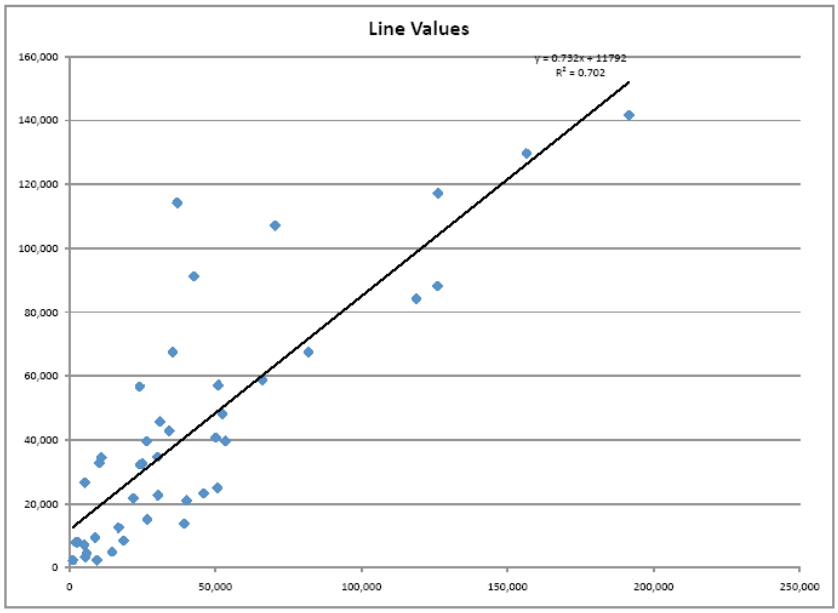
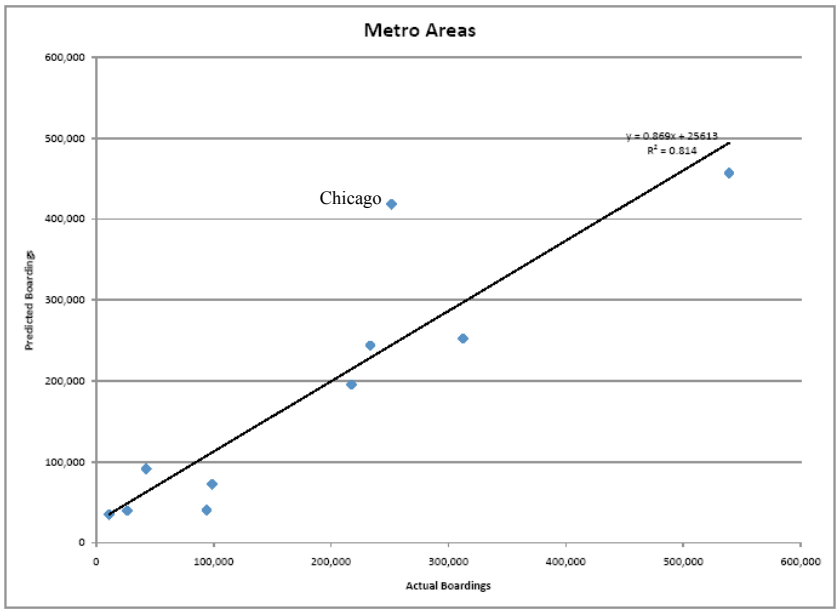


Figure 7. Predicted Versus Actual Boardings by City



Because this is a linear model, a change in the independent variable affects the dependent variable in the amount equal to its coefficient. For example, as noted in Table 4, a one mile increase in distance from downtown would bring an additional 115 people boarding at a given station, all else held equal. Similarly, 10 additional residents within 0.25 to 0.5 miles of the station would bring 2.2 additional people boarding at the station (0.219*10). A terminal station on a line results in additional 1,625 people boarding.

6.2 Application Guidance

It is important that the model is applied in cities with similar characteristics as those used for the development of the model. To do this, Table 3 presents the mean and standard deviation of each variable utilized in the model (for binary variables, the mean represents "percent of stations with binary value of 1"). For example, if a proposed extension is going to occur on a line where the employment within 0.25 miles of the station is above 17,000 (and it is not considered a CBD station), a planner should be careful using the results, since it is more than two standard deviations above the mean (17,000 is the average of 3,837 plus two standard deviations of 6,716).

Table 3. Mean Values and Standard Deviations for Variables in the Model

	Mean	**Standard Deviation**
If this is a terminal station	0.11	0.31
If this station is a secondary downtown	0.12	0.33
if this is a special transit attractor station	0.05	0.21
distance to downtown, in miles	5.89	4.78
number of buses connecting to this station	6.69	9.07
if there is parking available	0.36	0.48
if there is connection to other rail modes	0.11	0.31
CBD density, in employees per square mile	132,013.90	38,743.29
employment within 0.25 miles of the station	3,837.22	6,715.76
population within 0.25 to 0.5 miles of the station	5,035.60	3,778.04
employment within 0.25 to 0.5 miles of the station	4,088.24	7,295.42
[in(midday headway in minutes)	2.29	0.35
employment within 2 miles of entire line div by population	0.74	0.20

The model is also created for non-CBD stations only and it is not recommended that the model is used to predict boardings on CBD stations. Future analysis may illuminate better methods for predicting ridership for CBD stations. Generally speaking CBD stations are driven only by boardings due to employment around the station—and not population such as in residential areas.

Overall, the urban-area culture toward transit may vary how much ridership can be expected at each station and on each line. New York's subway system was specifically excluded because the ridership and the density of heavy-rail stations is so much higher than any other urban area, resulting in a different transit use pattern than that in other cities. Figure 7 shows the model over predicts the boardings in Chicago. This is most likely a factor of the density of the subway system. To resolve these issues, potential improvements to the model will seek to account for issues such as density of stations per number of residents/employees. As noted earlier, the model is really designed for medium and smaller size cities.

Ultimately, because transit usage and boardings can be very city specific, it is recommended that the outcomes of the model are calibrated with the actual boardings in the existing rail system, if an existing heavy-rail system exists. If the system does not exist, additional attention should be given to geographies with existing but newer heavy-rail systems, such as Miami or Washington, DC.

Finally, because this is a linear model, the outcome from the model may produce negative results. When this happens, the planner should be cautious of why that is the case. With linear models, it is possible that other factors contributing to boardings may be excluded from the model. As such, a "special transit station" may actually result in many more boardings than what the model assumes as

the average for all "special" stations. Improvements to the model can be accomplished by looking at non-linear models as well.

7. CONCLUSION

This research develops a relevant, sketch-level ridership forecasting tool for heavy-rail. Combined with the previous publications and sketch-level tools for light rail and heavy-rail, these tools provide a set of inexpensive methods to predict potential ridership along proposed transit corridors.

While this research looks into what may seem as an exhaustive list of variables, more variables could be added into the model to enhance its accuracy and applicability. With the use of readily-available data and tools such as ArcGIS and Microsoft Excel, these models can be applied before a full four-step travel demand modeling effort is undertaken.

8. REFERENCES:

1. Parsons Brinckerhoff Quade & Douglas, Inc., R. Cervero, Howard/Stein Hudson Associates, Inc., and J. Zupan. *TCRP Report 16: Commuter and Light Rail Transit Corridors: The Land Use Connection.* TRB, National Research Council, Washington, D.C., 1996.

2. Lane, Clayton, Mary DiCarlantonio, and Len Usvyat. Sketch Models to Forecast Commuter and Light Rail Ridership: Update to TCRP Report 16. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1986.* Transportation Research Board of the National Academies, Washington, D.C., 2006, pp 198-210.

3. Pushkarev, B., J. M. Zupan, and R. S. Cumella. *Urban Rail in America: An Exploration of Criteria for Fixed-Guideway Transit.* Indiana University Press, Bloomington, 1982.

4. Davidson, W. *Urban Rail Sketch Planning Ridership Forecasting Model.* Maricopa Association of Governments, Phoenix, Ariz., Sept. 1999.

5. R. L. Banks & Associates, Inc. *Kansas City Region Commuter Rail Study: Initial Corridor Screening.* Mid-America Regional Council, March 20, 2001.

6. Politics: Light Headed? Alternatives for Getting from Here to There. *Austin Chronicle,* Vol. 18, No. 47. www.austinchronicle.com/issues/vol18/issue47/pols.capmetro.side1.html. Accessed Nov. 10, 2005.

7. Blash, L., M. Duncan, R. Hranac, and B. Khazai. *AC Transit TelegraphMIS Project: Integrating GIS and Travel Model Evaluation Methodologies.* www-laep.ced.berkeley.edu/classes/cp255.people/cp255-bx/Project/final.html. Accessed Nov. 10, 2005.

8. LS Transit Systems, Inc. *Around the Bay Rail Study.* www.kko.com/sprc/research/baystudy/. Accessed Nov. 10, 2005.

9. Rice, W. E. A Model of Selected Ridership Forecasting Techniques with Limited Database—A Case Study: The Central Coast Diegan. *Proc., International Conference on Urban Transport and the Environment for the 21st Century. Computational Mechanics,* 1998, pp. 3–12.

10. Savage, J. P., Jr. Simplified Approaches to Ferry Travel Demand Forecasting. In *Transportation Research Record 1608,* TRB, National Research Council, Washington, D.C., 1997, pp. 17–29.

11. *Northern Virginia 2020 Transportation Plan: Technical Report.* Transportation Coordinating Council. www.virginiadot.org/projects/nova/nv2020/techdocs/docsects/tblocont.htm. Accessed Nov. 10, 2005.

12. Warner, M. *Travel Demand Forecasting.* Warner Transportation Consulting. www.warnertransportation.com/TDF.htm. Accessed June 14, 2005.

13. Saur, G., and R. Lee. Direct Ridership Forecasting. Presented at ITE Conference, Fehr and Peers, Inc., 2004

14. Yoder, S. L., S. E. Weseman, and J. DeLaurentiis. Capital Costs and Ridership Estimates of Personal Rapid Transit. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1704,* TRB, National Research Council, Washington, D.C., 2000, pp. 58–67.

15. Benjamin, J., S. Kurauchi, T. Morikawa, A. Polydoropoulou, K. Sasaki, and M. Ben-Akiva. Forecasting Paratransit Ridership Using Discrete Choice Models with Explicit Consideration of Availability. In *Transportation Research Record 1618,* TRB, National Research Council, Washington, D.C., 1998, pp. 60–65.

16. Benjamin, J. M., and R. Sakano. Equilibrium Model Based on a Microfoundation for Forecasting Dial-a-Ride Ridership. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1760,* TRB, National Research Council, Washington, D.C., 2001, pp. 135–142.

17. Outwater, M. L., S. Castleberry, Y. Shiftan, M. Ben-Akiva, Y. Zhou, and A. Kuppam. Attitudinal Market Segmentation Approach to Mode Choice and Ridership Forecasting: Structural Equation Modeling. In *Transportation Research Record: Journal of the Transportation*210 Transportation Research Record 1986 *Research Board, No. 1854,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 32–42.

18. Currie, Graham, and Justin Phung. Transit Ridership, Auto Gas Prices, and World Events: New Drivers of Change? In *Transportation Research Record: Journal of the Transportation Research Board, No. 1992*, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 3-10.

19. Ewing, Reid, and Robert Cervero. Travel and the Built Environment. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1780*, paper no. 01-3515, Transportation Research Board of the National Academies, Washington, D.C., 2001, pp. 87-114.

20. Cervero, Robert. Alternative Approaches to Modeling the Travel-Demand Impacts of Smart Growth. *Journal of the American Planning Association*, Vol. 72, No. 3, 2006, pp. 285-295.

21. Demery, L. W., Jr., and M. D. Setty. *The Recommendations and Forecasts of Pushkarev, Zupan, and Cumella: A Retrospective.* Oct. 2004. www.carquinezassociates.com/ptlibrary/trafficdensity/Pushkarev2.pdf. Accessed July 31, 2005.

22. Parsons Brinckerhoff Quade & Douglas, Inc. *Schuylkill Valley Rail Assessment Study.* Norfolk Southern Corporation, Philadelphia, Pa., Jan. 2005.

23. Parsons Brinckerhoff Quade & Douglas, Inc. *Quakertown-Stony Creek Rail Restoration Study.* Bucks County Planning Commission, Doylestown, Pa., Dec. 2000.

24. Parsons Brinckerhoff Quade & Douglas, Inc. *Track A Extension Feasibility Study, Phase II.* Cecil County Office of Planning, Zoning and Parks, Elkton, Md., 2004.

25. Parsons Brinckerhoff Quade & Douglas, Inc. *Gastonia Rapid Transit Alternatives Study: Corridor and Modal Options, Task 5: Evaluation.* City of Gastonia and Gaston Urban Area Metropolitan Planning Organization, Gastonia, N.C., June 2005.

26. Parsons Brinckerhoff Quade & Douglas, Inc. *2025 Long-Range Plan.* Delaware Transit Corporation, Wilmington, Del., Dec. 2000.

27. Parsons Brinckerhoff Quade & Douglas, Inc. *Richmond Rail Transit Feasibility Study.* Richmond Regional Planning District Commission, Richmond, Va., May 2003.