ELSEVIER

Transportation Research Part A 38 (2004) 223–247

TRANSPORTATION
RESEARCH
PART A

www.elsevier.com/locate/tra

Factors influencing light-rail station boardings in the United States

Michael Kuby *, Anthony Barranda, Christopher Upchurch

Department of Geography, Arizona State University, Tempe, AZ 85287-0104, USA

Received 30 December 2002; received in revised form 2 October 2003; accepted 8 October 2003

Abstract

Many US cities have recently built or approved light-rail systems to combat congestion, sprawl, and pollution. Critics questions light rail's ability to generate ridership in low-density, automobile-oriented, polycentric US cities with smaller downtowns. Proponents counter that sufficient numbers of homes and workplaces have convenient access to stations via walking, park-and-ride, or bus to develop feasible corridors connecting major residential areas with suburban concentrations of employment and the CBD. With this in mind, we used multiple regression to determine factors that contribute to higher light-rail ridership. Cross-sectional data on average weekday boardings were collected for the year 2000 for 268 stations in nine US cities representing a variety of urban settings. The results showed the importance of land use and accessibility. Employment, population, and percent renters within walking distance, as well as bus lines, park-and-ride spaces, and centrality, were significant. Dummy variables for terminal and transfer stations and international borders were all positive and significant. Total degree-days were negative and significant, lowering expectations for cities with extreme climates. Notably, the stations in the CBD generate much higher boardings, but these are explainable by the same variables present in lesser combinations at non-CBD stations and account for their generally lesser boardings. Importantly, a dummy variable for CBD location was not significant. The resulting model may be useful as a first-cut, one-step approach for predicting demand for possible light-rail alignments.
© 2003 Elsevier Ltd. All rights reserved.

Keywords: Light rail; Transportation; Regression model; Demand; Land use; Station

* Corresponding author. Tel.: +1-480-965-6850; fax: +1-480-965-8313.
E-mail address: mikekuby@asu.edu (M. Kuby).

0965-8564/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tra.2003.10.006

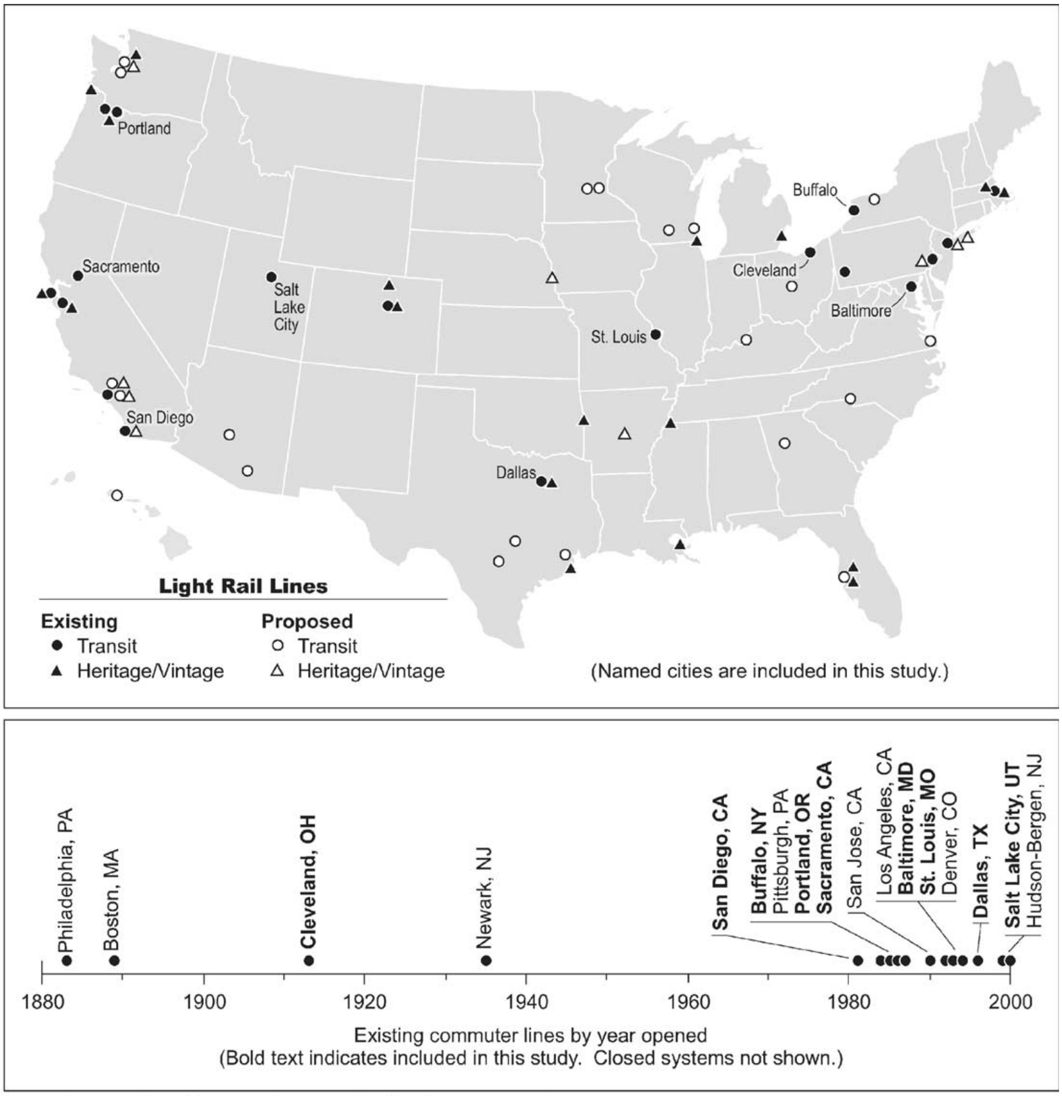
1. Introduction

Light rail transit is undergoing a resurgence, both in North America and overseas. Popularly known as streetcars or trolleys before the automotive heyday, light rail was the first mechanized form of urban mass transit. Its introduction in the 1880s fed the expansion of cities and the development of a star-shaped pattern focusing around the radial streetcar network (Adams, 1970). Light rail's popularity declined rapidly with the widespread adoption of the automobile beginning after World War I, and the dismantling of many streetcar systems in the 1930s and 1940s (Plane, 1995).

The resulting dependence on the automobile has led to congestion, sprawl, and air quality problems. To deal with these, many cities in the United States have returned to light-rail transit (LRT) with hopes of coaxing drivers out of their cars. Twelve cities have built light-rail systems aimed at commuters and residents in the last two decades, while 20 others are being planned (Fig. 1). There is also a wave of construction of heritage or vintage streetcar lines aimed at tourists or downtown workers. Cities are turning to light rail in more heavily traveled corridors because it is faster and more appealing than buses and yet cheaper to build than heavy rail. Revenue vehicle-miles for LRT grew by 77% from 1991 to 1999, compared with 8% for population and 30% for transit in general (Bureau of Transportation Statistics, 2000). According to the American Public Transportation Association (2002), just over 1 million unlinked light-rail trips were made nationally on an average weekday in 2000—up 77% since 1989—over an average length of 4.2 miles.

Yet questions remain regarding how many riders light-rail transit (LRT) can attract, especially in low-density, automobile-oriented, polycentric cities. The historical argument that streetcar lines shaped urban land use a century ago can be turned on its head to contend that light rail cannot attract enough riders in cities that have evolved around the automobile. Critics of LRT argue that today's cities have neither enough jobs downtown nor enough residents along their corridors to generate substantial ridership, especially in western cities like Phoenix, Salt Lake City, and Dallas whose urban forms crystallized in the auto era. Yet in today's polycentric metropolitan areas, light-rail networks can connect residential areas with large concentrations of employment both within and outside of the traditional CBD, while buses, bicycles, and park-and-ride can be coordinated with the rail stations.

Numerous examples from our study of nine existing light-rail systems in 2000 reinforce the polycentric and multipurpose nature of light rail today. While the 70 downtown stations average just under 10,000 employees within a half-mile walking distance, three non-CBD stations have over 10,000 employees nearby, and 35 have walking access to over 3000 employees. Directionally, during the 3-h evening peak period, the San Diego Trolley boasted a ratio of outbound to inbound passengers of only 4:3, indicating not all passengers are CBD commuters. The Portland LRT averaged 65,100 daily riders on weekdays, and a still healthy 49,500 on Saturdays for what can be presumed to be a wider variety of trip purposes. Non-work trips are also illustrated by St. Louis, where special trains averaged 9373 fans per game for 10 Rams football games, 7616 fans per game for 87 Cardinals baseball games, and 45,000 riders for the 3-day Fair St. Louis celebration. Bus service can augment non-CBD ridership, as seen in 12 non-CBD stations in our study with 10 or more bus lines and 1200–4100 daily boardings, despite limited park-and-ride capacities averaging only 250 spaces. As these varied facts show, an analysis of light-rail ridership should



Source: American Public Transportation Association (http://www.apta.com)

Fig. 1. Map and timeline of existing and proposed light-rail systems.

consider factors other than employment inside the CBD and population near the non-CBD stations.

Our study was designed to assess the factors driving LRT ridership in the United States in 2000, including but not limited to land use factors. Data on weekday boardings were collected for the year 2000 for 268 stations in nine US cities ranging from Buffalo to St. Louis to San Diego. For each city, five categories of independent variables were collected. Land use variables measured population, employment, college enrollment, airports, and international borders capable of generating traffic from within a half-mile walking distance of each station. Network structure variables dealt with hypotheses relating to station spacing, interline transfer points, centrality, and critical mass. Several citywide variables accounted for effects that could raise or lower ridership at all stations in a city, such as climate and total population. Intermodal access variables focused on bus, park-and-ride, and other rail systems. Finally, a socioeconomic variable represented the differential propensity of residents to travel by mass transit.

The goal of the study was to quantify the effects of these factors on the average weekday boardings of existing light-rail stations. A key improvement of this study over prior research is that the CBD stations were treated the same as those in every other part of the city, fully characterized by their nearby employment, population, accessibility, and other features. The approach taken was polycentric, multi-modal, and multi-activity, in recognition of the reality of the contemporary American city in the 21st century.

2. Prior research

This research lies squarely within the ongoing debate about what kinds of cities and corridors are conducive to high transit ridership. This is an important policy issue on which billions of dollars are riding. Scholars such as Morrill (1991) and Giuliano (1995), as well as countless politicians and newspaper writers, have questioned whether rail transit can succeed in today's dispersed cities. Fielding (1995) summarized the conventional wisdom that transit use is higher in older and most congested urban areas than in newer Sun Belt cities. The Boston–Washington corridor and the greater Chicago area account for 64% of transit use but only 15% of the country's population. Pickrell (1992) found that ridership in ten new rail transit systems was 15–75% below forecast levels in their feasibility studies. Morrill (1991), Pickrell (1992), and Kain and Liu (1999) all argued that buses can serve today's spatially dispersed demand more cost-effectively.

Density or compactness lies at the heart of these concerns (Pushkarev and Zupan, 1982; Smith, 1984; Nelson/Nygaard Consulting Associates, 1995). According to Rosenbloom and Clifton (1996), the evidence for a positive relationship between population density and transit ridership is "fairly well established." Density has proven to be a significant determinant of ridership at the metropolitan level of aggregation (Rosenbloom and Clifton, 1996) and at the station-area level (Parsons Brinckerhoff Quade & Douglas, Inc., 1996), in part because station-area residents are 5–7 times more likely to travel by rail than residents beyond the station area (Cervero, 1993).

Density, however, is not the only important spatial factor. Rosenbloom and Clifton (1996) found the percentage of workers using mass transit to be higher in low-density cities over 1 million than in smaller but denser cities, although their research did not control for the differing availability of transit service in cities of different sizes. Kain and Liu (1999) linked transit ridership gains with network expansion, fare reductions, and growth in employment and population. Parsons Brinckerhoff Quade & Douglas, Inc. (1996) found density and CBD size to be major influences on LRT traffic and found greater use of non-automotive modes in "traditional" pre-1950 neighborhoods than suburban areas.

Because many urban rail systems in the US radiate from the CBD, much research has focused on commuting to the CBD as the mainstay of rail transit. As recently as 1980, mass transit in cities such as Chicago, Boston, New York, and San Francisco captured over half of the work trips to the CBD (US Department of Transportation, 1984). At the opposite end of the spectrum, Cervero (1993) found negligible rail usage for suburban office parks beyond the station area and with free parking. Parsons Brinckerhoff Quade & Douglas, Inc. (1996) found that LRT ridership increased exponentially with total CBD employment and CBD employment density.

Other research, however, has focused on niches other than commuting to the CBD. Fielding (1995) notes that 14% of transit trips are school-related. Filion (2001) found that mixed-use suburban centers have been successful in attaining higher transit use than the typical suburban area and advocates creating high-density, transit-oriented corridors. Mason (1998) noted the high demand for transit to downtown entertainment.

Conversely, vehicle ownership and employment dispersion are major factors behind the decline of densities and transit use. From 1993 to 1996, 82% of US central cities lost job share to surrounding suburbs (Brennan and Hill, 1999), while during the 1980s, downtown jobs fell as a share of total jobs in every metropolitan area (Leinberger, 1993). Meanwhile, the percentage of US households owning zero or one vehicles shrank from 50% in 1977 to 40% in 1995 (Federal Highway Administration, 2000). Once the investment is made, the convenience of a car often leads to an immediate decline in transit use.

The speed advantage of cars over trains can disappear entirely as a result of congestion. Congestion delays have nearly tripled from 1982 to 1997 in 68 major US metropolitan areas, increasing from 16 to 45 h per year (Shrank and Lomax, 1999). Over time, the delays have been consistently worse for cities over 3 million, and consistently less for cities under a half million. The importance of city size for transit ridership found by Rosenbloom and Clifton (1996) may be due to congestion rather than size per se. Good transit management can also offset transit's disadvantages. Strategies target service, speed, reliability, intermodal coordination, weather protection, advertising, subsidy removal, and higher parking fees, among others (Fielding, 1995; Shoup and Willson, 1992).

Several studies have confirmed that ridership is higher among low-income households, which tend to locate in denser neighborhoods, thus placing more individuals without cars in walking distance of transit (e.g., Pushkarev and Zupan, 1977; Wachs, 1989). Rosenbloom and Clifton (1996), however, found that the percentage of people using commuter and subway/elevated rail modes goes up consistently with income—and not simply because incomes are higher in larger cities. This contradiction may be related to the location of different income groups within the metropolitan area and their subsequent average length of commute. Lower-income households take shorter trips and are therefore more likely to use buses, which cater more to shorter trips; likewise, wealthier workers take longer trips and are more likely to use rail (Wachs and Taylor, 1998; Wachs et al., 1993).

International comparisons raise questions about the cause of low transit usage in the United States. Bly et al. (1980) removed the effects of better service and fare subsidies, and attributed some of the remaining differences to urban density. Fielding (1995), however, pointed out that some Canadian cities with similar car ownership rates and urban compactness to US cities have substantially higher ridership due to governmental policies. Gordon and Willson (1984) found LRT ridership in US cities to be even lower than expected based on their low densities, and high incomes and auto ownership.

3. Methodology

The Parsons Brinckerhoff Quade & Douglas, Inc. (1996) study, on which our methodology is loosely based, used a cross-sectional regression analysis of station-level data to study land use characteristics that make a corridor conducive to light rail. They combined data from 261 stations on 19 lines in 11 metropolitan areas, and their dependent variable, daily station boardings, is the same as ours. Following their study, our models also "bypass the usual four-step travel demand modeling process with a simplified approach that estimates travel demand directly, incorporating trip demand, mode choice, trip distribution, and traffic assignment features" (p. E-2).

In addition to updating the data by a decade, we have tried to improve on their approach in other ways. We introduced an improved method of estimating the walking distance around stations. We added some independent variables that they did not consider, such as climate and international borders. We revised some of the variables they used, such as by counting the number of connecting bus lines instead of using a feeder-bus dummy variable. But the most important change that sets this work apart from theirs is that we have treated the CBD stations the same as all others. Their study, in fact, pointed the way for us to follow:

> In the future, as cities continue to evolve toward multiple centers, transit systems that link the CBD with subregional employment centers will be especially cost-effective, offering opportunities for two-direction flows at all times of the day. Further, within compact urban regions, transit service in corridors that contain a variety of residential and non-residential activities will prove especially attractive and competitive. (p. 2)

Despite this recommendation, their 1996 study excluded CBD stations from the analysis, looking only at ridership at non-CBD stations. It included CBD employment as an independent variable, but ignored employment around other stations. Their study computed distances from the stations to the CBD, but ignored a station's accessibility to other stations. Finally, it ignored residential population in the CBD. Steeped in the monocentric assumptions about cities and transit, their approach seems possibly biased against Sun Belt metropolitan areas that fit the polycentric "edge city" model better. Our approach, which uses the same ridership, employment, population, and intermodal access variables for both CBD and non-CBD stations, is perhaps better suited for assessing light-rail ridership in today's cities.

The data in this study were analyzed using multiple regression because of its ability to simultaneously evaluate the effects of a large number of factors. Multiple regression is flexible, widely used, and easily understood by a broad audience. Its ability to handle both numerical and dummy (0–1) variables was also important because of several 0–1 conditions being evaluated. Another advantage of using multiple regression is that the resulting model can be used, within certain limits, for predictive purposes in other cities (see Section 6).

Before and after running the regression, the data were checked for violations of the Gauss–Markov assumptions regarding multi-collinearity, non-linearity, spatial autocorrelation, and heteroscedasticity. Results of these tests of assumptions are discussed in the first part of Section 5, after the variables themselves are introduced in Section 4. Following that, the results are presented both with and without corrective steps taken to deal with violations.

4. Data

The study focused on the year 2000, so that recent boardings data would coincide with the 2000 census data. Because data were collected from national sources and local agencies, a major effort was made to maintain consistency among cities.

4.1. Definition of LRT and study areas

The distinction between light and heavy rail (subway, elevated, or commuter rail) is not perfect. The American Public Transportation Association (2002) defines LRT thusly:

> Light Rail is lightweight passenger rail cars operating singly (or in short, usually two-car, trains) on fixed rails in right-of-way that is not separated from other traffic for much of the way. Light rail vehicles are driven electrically with power being drawn from an overhead electric line via a trolley or a pantograph. Also known as "streetcar," "tramway," [or] "trolley car."

Because this study pooled stations from different systems, it was of critical importance to establish consistent criteria for which LRT systems to include. Of the list of existing LRT systems on the APTA website, the following systems were included:

1. Baltimore, MD	MTA (Maryland Transit Administration)
2. Buffalo, NY	Metro (Niagara Frontier Transit Metro System)
3. Cleveland, OH	RTA (Greater Cleveland Regional Transit Authority)
4. Dallas, TX	DART (Dallas Area Rapid Transit Authority)
5. Portland, OR	TriMet (Tri-County Metropolitan Transp. Dist. of Oregon)
6. Sacramento, CA	SRTD (Sacramento Regional Transit District)
7. Saint Louis, MO	Bi-State Transit (Bi-State Development Agency, BSDA)
8. Salt Lake City, UT	UTA (Utah Transit Authority)
9. San Diego, CA	SDT (San Diego Trolley)

A number of LRT systems were eliminated from consideration because of inconsistencies with the main group. First, Denver, San Jose, and Newark met all criteria but were eliminated due to inadequate data. Second, this study applied a stricter definition than APTA's regarding underground stations, which blur the distinction between a subway system and a light-rail system and thus differ with regards to weather, safety, atmosphere, and interaction with road traffic. Boston, Los Angeles, Philadelphia, and Pittsburgh were eliminated because they have numerous underground stations, but Buffalo was included because it has only one, which did not seem enough to alter the general perception of the system as a surface mode. Finally, eight systems were eliminated because they primarily serve downtown areas (e.g., the Portland Streetcar) and/or tourist traffic (e.g., New Orleans). APTA reclassified these eight systems in September 2002 as "heritage or vintage trolleys" (vehicles built before 1960 or modern replicas).

4.2. Definition of walking distance

A critical assumption for this study was how to define the area within walking distance of a station. This assumption affects a number of variables, such as employment, population, and college enrollments, as well as the socioeconomic characteristics, and the total covered employment. The assumption breaks down into two parts: how far is walking distance, and how to determine the area within such a distance.

A number of different values have been applied in the literature. Shaw (1991) assumed a 2000 foot walking distance for the Miami Tri-Rail system. For the Washington Metrorail, JHK and Associates (1987) found that, within one-third mile, 18–63% of apartment building residents, and 5–36% of office building employees, used transit. Simons and El Jaouhari (2001) found that homes in Cleveland located within 1000 ft of light rail were worth up to $12,000 more. Barber (1995) reported that median walking distances vary between 400 and 1200 ft. Untermann (1984) found a distance-decay relationship in which most people were willing to walk 500 ft, 40% would walk 1000 ft, but only 10% would walk a half mile. Parsons Brinckerhoff Quade & Douglas, Inc. (1996) tested half-mile and two-mile walking distances. To be consistent with the latter study, and to minimally encapsulate all of the zones found to be significant in these other studies, a round distance of one-half mile was selected.

Because of this study's focus on land uses that generate light-rail ridership, it required accurate estimation of mutually exclusive service areas within walking distance of each station. Furthermore, the relevant distance should be the shortest network path, not the Euclidean distance. A special raster-based approach, called the Linked On–Off Network method, was developed for this project to improve upon standard GIS buffering commands, which often do not allocate off-network areas on an exclusive or realistic basis (Upchurch et al., 2004). Realistic allocation of the areas between the streets was important to this project because the census data and some of the employment data were based on census blocks or traffic analysis zones. The method used here assumes that pedestrians will reach the street network at the closest point to their origins and from there follow the shortest path to reach their closest station (Fig. 2). Transit agencies were contacted to ascertain pedestrian access across freeways and other barriers.

4.3. Dependent variable: average weekday boardings

This paper aims to explain LRT average weekday boardings at the station level. The focus is on weekdays because ridership is higher on weekdays, some systems do not collect data on weekends, and reduction in weekday traffic congestion helps justify the investment in LRT. Local transit agencies provided boardings data for 2000. Each agency conducted their surveys in different ways, at different frequencies, and in different seasons, however, all agencies follow National Transit Database standards in counting "unlinked" trips. Passengers that transfer between light-rail lines are counted twice—at their original station and the transfer point. Weekday boardings per station averaged 1260 passengers, with a range from 16 at Cleveland's Attleboro station to 13,610 at San Diego's San Ysidro border station (Table 1).

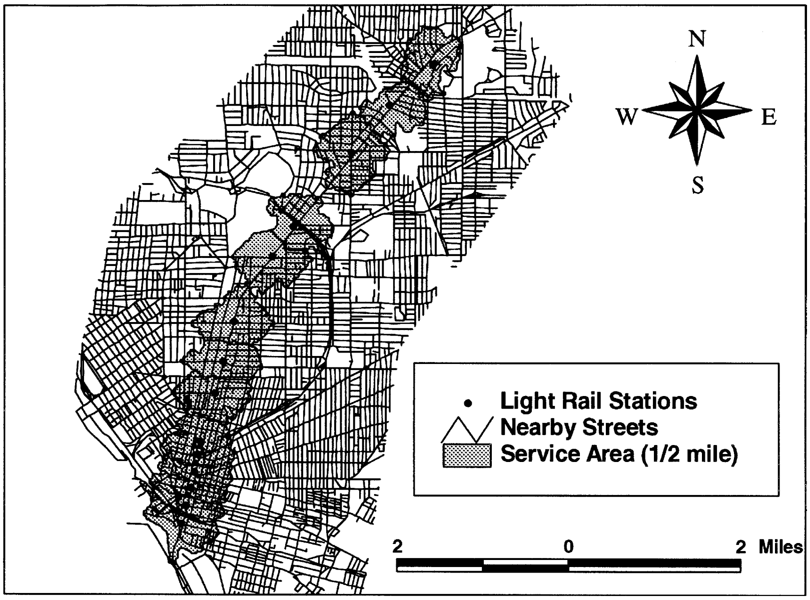


Fig. 2. Mutually exclusive, network-based service areas for the Buffalo, NY light-rail system. Buffers such as these were used to estimate employment and population within a half-mile walking distance.

4.4. Independent variables

The independent variables represent factors hypothesized to influence station boardings (Table 1). The variables fall into five categories: (1) traffic generation; (2) intermodal connection; (3) citywide; (4) network structure; and (5) socioeconomic.

4.5. Traffic generation/land use variables

4.5.1. Employment within walking distance

Employment within walking distance of each station was hypothesized to be the most important factor for work trips. Much effort went into an unsuccessful search for a national employment data source that was sufficiently accurate and spatially disaggregated. As non-local researchers with limited resources, we decided to rely on local metropolitan planning organizations (MPOs) to provide the best and most disaggregated employment data that they were able to produce for their own purposes. Their raw sources included state unemployment data, Dun and Bradstreet address-matched data that had been error-checked, and others. Each was asked to provide total public and private employment in 2000. Most available employment data were for 1999–2001, except for St. Louis and Cleveland, which provided projections to 2000 that they themselves use.

Data were requested by establishment, not by corporate headquarters, preferably in point form or as the most disaggregated areal data possible. Two of the MPOs provided point data, which were allocated to the proper service area using ArcView GIS. Five MPOs provided areal data, and ArcView calculated the employment within each station's service area assuming uniform spread of

Table 1
Summary of independent variables and hypotheses

Independent variable	Shorthand variable name	Hypothesis	Minimum value	Average value	Maximum value	Zero entries
Employment within walking distance	EMPLOYMENT	+	11	3899	51,096	0
Population within walking distance	POPULATION	+	0	1490	9361	4
Airport	AIRPORT	+	0	n.a.	1	266
International border	BORDER	+	0	n.a.	1	267
College enrollments	COLLEGE	+	0	678	35,614	248
CBD dummy variable	CBD	+	0	n.a.	1	197
Park-and-ride spaces	PARK&RIDE	+	0	143	1150	147
Bus connections	BUS	+	0	3.0	28	92
Other rail lines	OTHER_RR	+	0	n.a.	1	260
Heating and cooling degree-days	DEGREE_DAYS	–	152	404	591	0
PMSA population	PMSA_POP	+	1,170,111	2,282,025	3,519,176	0
Terminal station	TERMINAL	+	0	n.a.	1	246
Station spacing	SPACING	+	0	2044	17,270	136
Designated transfer station	TRANSFER	+	0	0.0075	1	266
Normalized accessibility	CENTRALITY	–	0.35	0.64	1.0	0
Percentage of PMSA employment covered by system	EMPLOY_COV	+	8%	18%	30%	0
Percent renters within walking distance	PCT_RENT	+	0	0.63	1	2

employment over each employment polygon. The final two MPOs calculated total employment within the service areas we provided using point data. Employment within a half-mile walking distance averaged 3899 employees per station, from a low of 11 at Salt Lake City's southern terminal station to a high of 51,096 at downtown Dallas' Akard station.

4.5.2. Population within walking distance

Higher residential population was hypothesized to be positively associated with boardings. Census block population was obtained from the 2000 census redistricting files. ArcView calculated the population within the half-mile service area, assuming uniform population distribution within each block. Population averaged 1490 people, from a low of 0 (many cases) to a high of 9361 at the 25th and Commercial station in San Diego.

4.5.3. Airport

It was hypothesized that airports generate LRT passenger traffic above and beyond the ridership of airport employees. A dummy variable equal to 1 is used for LRT stations that serve passenger terminals, and 0 otherwise. Because airport employment may be underestimated in cities with areal employment data (airport employment is not equally spread over the entire polygon), this variable may also capture increased employee ridership. St. Louis and Baltimore had airports served by light rail.

4.5.4. International border

The busiest station by far was San Diego's San Ysidro-Tijuana station, with 13,610 boardings per day. Rather than eliminate it as an outlier, an international border dummy variable was established with 0s at all stations except San Ysidro. The hypothesis is that an international border location generates substantial additional boardings.

4.5.5. College enrollments

Twenty of the 268 stations had colleges within walking distance, which were hypothesized to produce extra boardings above and beyond the employment and residence effects. Several special cases involved universities served by two stations, in which case the enrollment was divided equally between them. Data on college enrollment came from The 2002 World Almanac and Book of Facts (2002), university web sites, and phone calls.

4.5.6. CBD dummy variable

A CBD dummy variable was included to test whether downtown stations generated ridership above and beyond what would be expected given their other measurable attributes such as population and employment. It was hypothesized that some important aspects of the primary CBD might be overlooked, such as high parking costs, traffic congestion, a pedestrian-friendly environment, more employment within a shorter walking distance, and cultural/lifestyle effects. Because other studies such as Parsons Brinckerhoff Quade & Douglas, Inc. (1996) excluded CDB stations, this independent variable was one of the most important.

Defining and delimiting a CBD is a difficult task beyond the scope of this study. Therefore, we used the CBD definitions from Parsons Brinckerhoff Quade & Douglas, Inc. (1996), allowing direct comparison with their results. They delimited CBDs at the zip code level based on employment density, type of industry, and expert local knowledge. If a station was located within a CBD zip code, we coded that station as a 1; otherwise 0. Although the CBDs in our study could have grown or contracted slightly since the early 1990s, it seemed unlikely that this would change the definition at their rather coarse zip code level. The Parsons Brinckerhoff study defined a CBD for seven of our nine cities. [1] For the remaining two (Dallas and Salt Lake City) we analyzed employment density and known CBD locations to determine the core CBD zip codes. [2]

4.6. Intermodal access variables

4.6.1. Park-and-ride spaces

Many commuters, especially those not within walking distance, reach light-rail stations via park-and-ride. A strong positive association was hypothesized. These data were obtained mainly from transit agency web sites. Park-and-ride lots were available at 47% of stations, with an average of 319 spaces per lot, from a low of 5 to a high of 1150 at the Dallas Park Lane station.

4.6.2. Bus connections

The number of different bus lines intersecting with a station was hypothesized to be positively related to station boardings. Data were obtained mainly from transit agency web sites. All bus lines within a block of a station were counted, so as not to ignore buses on one-way streets. No attempt was made to distinguish among: bus lines parallel or perpendicular to the rail line; shuttles, regular buses, or express buses; or whether the bus intersects more than one rail station. The average number of bus connections at light-rail stations was 3, with a high of 28 bus connections at Portland's SW 4th St. Transit Mall station. A total of 92 stations had no bus connections at all.

4.6.3. Other rail lines

Stations connecting to other kinds of rail lines, such as subways, vintage streetcars, heavy-rail commuter lines, or intercity passenger rail, were hypothesized to gain traffic from intermodal connectivity. This variable was set up as a 0–1 dummy. The criteria imposed was that the stations had to be physically co-located.

4.7. Citywide variables

4.7.1. Heating and cooling degree-days

Weather could be a factor in discouraging LRT ridership, and while it is not something under the control of transit agencies (with the exception of building shelters), it is a potentially confounding variable for which one must control. Weather is operationalized as average monthly heating and cooling degree-days for all stations within each metropolitan area (NOAA, 2002). [3]

[1] For Cleveland, the 44113 zip code was omitted from their list. Considering that this zip code contains the main downtown transit hub of Tower City, we included this zip code in addition to 44114 and 44115.

[2] For Salt Lake City, these were 84101 and 84111; for Dallas, 75201 and 75202.

[3] Heating (cooling) degree-days are the total accumulated degrees below (above) a 65 °F base temperature.

The hypothesis is that hot or cold weather discourages transit use because it often requires walking and waiting outdoors. Heating and cooling days were added together because of multicollinearity between them. San Diego had the lowest monthly degree-days with 152 (i.e. the most even temperatures close to 65); Buffalo had the highest with 592.

4.7.2. PMSA population

Primary metropolitan statistical area (PMSA) population was used here as a convenient measure of city size. The PMSA was considered more relevant to LRT ridership than the consolidated MSA because light rail is a short-distance mode, averaging 4.2 miles nationwide. In metro areas with no PMSAs, the MSA population was used. Transit use is higher in larger cities (Rosenbloom and Clifton, 1996) for many reasons, some of which are partially measured by other variables (employment and population density, socioeconomic status, number of rail lines) and some of which were not (congestion). This variable was included to test generally whether larger cities have conditions more favorable to LRT ridership. Buffalo had the lowest population in the study, with 1.17 million, while Dallas had the highest with 3.52 million.

4.7.3. City-specific dummy variables

Although not used in the base model, dummy variables for each city were added in additional runs to account for spatial autocorrelation among observations within the same city. These variables test for otherwise unmeasured city-specific factors such as crime and congestion and system-specific factors such as fare structure.

4.8. Network structure variables

4.8.1. Terminal station

A dummy variable for all stations at the terminus of an LRT line was hypothesized to be positively associated with station boardings. The terminal station is the nearest station for residents of a large area beyond the end of the line. This variable was also significant in the Parsons Brinckerhoff Quade & Douglas, Inc. (1996) study.

4.8.2. Station spacing

Light rail planners are faced with a difficult decision when locating stations. If stations are spaced closely, there will be more potential passengers within walking distance of a station. However, close spacing increases travel times for through passengers. Parsons Brinckerhoff Quade & Douglas, Inc. (1996) studied distance to the nearest station (significant), and distance to the next inbound station. The variable used here essentially combines these two by measuring *extra* spacing beyond a station's walking-distance buffer on both sides. The hypothesis was that additional spacing around a station would draw additional riders to that station.

4.8.3. Designated transfer station

Given that the standard method for counting boardings is unlinked trips, it was expected that stations designated for interline transfers would have higher boardings. Only Cleveland and San Diego had multiple lines in 2000, not counting small spurs in Baltimore and Dallas. In Cleveland and

San Diego, although there were numerous stations where riders can change lines, Tower City in Cleveland and 12th and Imperial in San Diego were the only advertised transfer stations. These stations were given a dummy value of 1.

4.8.4. Normalized accessibility (centrality)

This variable represents the relative accessibility of each station to all other stations, as determined by average travel times. Average travel time (including transfer time) was computed weighting all stations equally, in recognition that in today's polycentric cities it cannot be assumed that the destination is the CBD. The average travel time for each station was normalized by dividing by the highest average travel time for that system. Otherwise, even the most central of stations in a large system would have higher average travel times than the most remote station in a small system. The resulting variable permits relative comparisons both within and between systems. The variable ranges from a low of 0.35 in central areas to a high of 1 at one of the terminal stations (Fig. 3). This variable can be contrasted with the distance-to-CBD variable used by Parsons Brinckerhoff Quade & Douglas, Inc. (1996).

4.8.5. Percentage of PMSA employment covered by system

It is often said of today's decentralized cities that employment is too dispersed to be served efficiently by a fixed-rail system. According to this hypothesis, if a LRT system can be laid out in such a way that a high percentage of a city's employment lies within walking distance of stations, overall ridership should increase. This variable, which ranges from a low of 7.8% in St. Louis to a high of 29.6% in Portland, can be considered a second *critical mass* hypothesis. It could also be classified with the other citywide factors, but it is included here under the network structure category because it is a function of the route of the corridor(s). Total MSA or PMSA employment data is from the 1997 Economic Census.

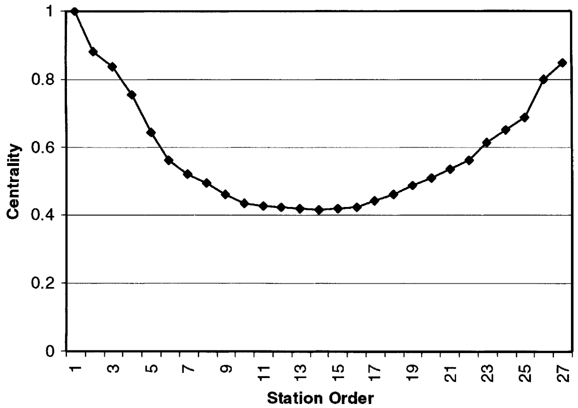


Fig. 3. Plot of centrality, or normalized accessibility, for the St. Louis light-rail system. College Station in East St. Louis, IL is station 1, the CBD is stations 11–16, and the airport is station 27.

4.9. Socioeconomic variables

4.9.1. Percent renters within walking distance

Many studies have found that transit ridership is correlated with socioeconomic status. Income and car ownership have direct effects on a person's alternative mode choices. At the time this analysis was performed, however, the only 2000 census data released at the block level was the redistricting file (US Census Bureau, 2002). The most relevant information available was percent of households that rent, which is hypothesized to be positively related to station boardings. Renters tend to be disproportionately poor, young, located in denser multifamily housing, which may lack parking. While this variable is imperfect as a measure of socioeconomic status, we considered it better to include a flawed measure of socioeconomic status than a 10-year-old measure or no measure at all. Percent renters was calculated by (1) assuming uniform values within each census block, (2) estimating the overlap of census blocks with station service areas, (3) calculating total renters and total households within the area, and (4) dividing renters by households.

5. Results

5.1. Tests for Gauss–Markov assumptions

To test for multi-collinearity among independent variables, Pearson correlation coefficients were generated for all pairwise combinations. The highest observed correlation is 0.531 between centrality and the terminal station dummy variable. The reason is obvious—terminal stations tend to be the least centrally located—yet these variables clearly represent different aspects of route geometry. Centrality is a function of travel time to all stations and varies continuously, whereas the terminal station variable represents their greater proximity to customers beyond the end of the light-rail corridor. The next highest correlation is 0.427 between percent renters and employment, which is explained by the fact that housing in commercial and industrial areas tends to be rental units. All four of these variables were kept in the model because the correlations were well below the danger level of 0.7 (Clark and Hosking, 1986), and each measures a unique and decidedly different factor.

The Parsons Brinckerhoff study found an exponential relationship between boardings and CBD employment, and used a log–log transformation of the dependent variable and the numeric independent variables. No logarithmic transformation was deemed necessary here, for two reasons. First, none of the independent variables have a non-linear relationship with the dependent variable. Second, the effects of the independent variables would not theoretically be expected to be multiplicative, as implied by a log–log transformation. If the model were dealing with employment in one location and population in another, the effects would be multiplicative, as in most gravity-type models. In this study, however, employment and population data are for the *same* station area. A station could generate substantial boardings with a large employment base and no population, or vice versa. In a multiplicative (log–log) model, near-zero values of any independent variable effectively reduce the predicted value to near-zero. As the last column of Table 1 shows, many of the numeric independent variables have many cases with zero values.

The initial results of the model showed signs of heteroscedasticity. This can be seen in a triangle-shaped scatterplot (Fig. 4). For this reason, this paper presents results using regular OLS

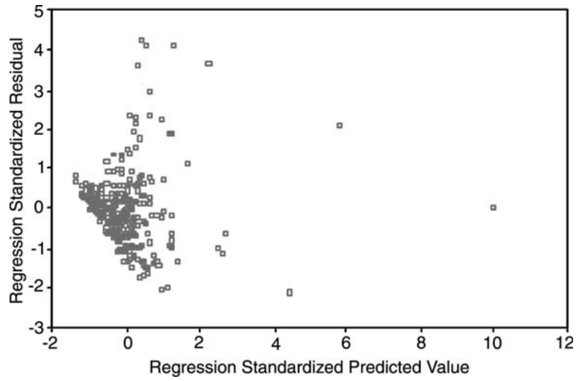


Fig. 4. Scatterplot of standardized residuals showing heteroscedasticity.

regression and also results from LIMDEP corrected for heteroscedasticity (Greene, 1998). LIMDEP uses a revised and robust OLS covariance matrix based on White (1978), who developed a consistent estimator of the standard errors under heteroscedasticity. The correction does not change the coefficient estimates, but it does change the standard errors and t-ratios.

5.2. Regression model results

A number of statistical analyses were run to test the 17 different hypotheses and develop the best model for explaining and predicting LRT boardings. Initially, models were run using OLS multiple regression, which helped identify the heteroscedasticity and narrow the list of worthwhile independent variables. The variables that were clearly not significant in explaining average weekday boardings were station spacing, total PMSA population, college enrollments within walking distance, and dummy variables for heavy intra-city rail connections and CBD locations. All of these had been expected to be positively associated with boardings, but none were significant at even the 0.25 level.

The remaining 12 independent variables—most of which were highly significant in both models—were analyzed further using the LIMDEP statistical program. The 12 significant and borderline significant independent variables were analyzed with LIMDEP with and without White's correction factor. Exclusion of the five highly non-significant independent variables was justified by the fact that their partial redundancy with the potentially significant variables would have altered the coefficients and significance of the more significant variables. The results are summarized in Table 2, and in the following regression equation:

$$\begin{aligned}\text{BOARDINGS} = {} & 1584 + 0.023 * \text{EMPLOYMENT} + 0.092 * \text{POPULATION} \\ & + 915 * \text{AIRPORT} + 12{,}055 * \text{BORDER} + 0.774 * \text{PARK\&RIDE} \\ & + 123 * \text{BUS} - 1.52 * \text{DEGREE\_DAYS} + 660 * \text{TERMINAL} \\ & + 5735 * \text{TRANSFER} - 1872 * \text{CENTRALITY} \\ & + 1301 * \text{EMPLOY\_COV} + 624 * \text{PCT\_RENT} + \text{error}\end{aligned}$$

Table 2
Multiple regression results with (and without, in parentheses) the correction for heteroscedasticity

Independent variable	b-Coefficient	t-Ratio	Significance
CONSTANT	1583.82	4.178	0.0000
		(4.577)	(0.0000)
EMPLOYMENT	0.02294	2.024	0.0440
		(2.436)	(0.0155)
POPULATION	0.09156	2.307	0.0219
		(2.414)	(0.0165)
AIRPORT	914.54	2.460	0.0145
		(1.560)	(0.1201)
BORDER	12,054.94	54.400	0.0000
		(14.854)	(0.0000)
PARK&RIDE	0.77415	3.519	0.0005
		(3.407)	(0.0008)
BUS	122.88	6.577	0.0000
		(9.819)	(0.0000)
DEGREE_DAYS	−1.51690	−3.518	0.0005
		(−4.220)	(0.0000)
TERMINAL	660.42	3.489	0.0006
		(2.969)	(0.0033)
TRANSFER	5734.83	4.622	0.0000
		10.244	(0.0000)
CENTRALITY	−1871.77	−5.256	0.0000
		(−5.011)	(0.0000)
EMPLOY_COV	1300.99	1.933	0.0543
		(1.728)	(0.0851)
PCT_RENT	623.87	3.990	0.0001
		(3.053)	(0.0025)
n	268		
df	255		
R^2	0.727		
Adjusted R^2	0.715		
F	56.69		

The model, with 255 degrees of freedom, has an R^2 value of 0.727 (adjusted R^2 of 0.715), and an F-statistic value of 56.69, significant at the 0.000 level. The model thus explains 73% of the variance in average weekday boardings over all stations. This compares with an R^2 value of 0.536 (adjusted R^2 of 0.523) in the Parsons Brinckerhoff study that analyzed only non-CBD stations.

5.3. Discussion

All of the variables have the hypothesized positive or negative coefficients, and all are significant at the 0.05 level except for two borderline variables. The first of these, the airport dummy variable, is significant when using the corrected OLS covariance matrix—the more consistent estimator—but not when using standard OLS. The b-coefficient implies that an additional 915 passengers board at airport stations. The second—proportion of PMSA employment covered—is significant at the 0.054 and 0.085 levels in the two models, indicating a possible relationship. The *b*-coefficient of 1301 indicates ridership at all stations rises by 13 passengers per day for each percentage of PMSA employment within walking distance of any station. Though not quite significant at the 0.05 level, the coefficient appears reasonable and we would leave this variable in when using the model for predictive purposes.

Employment within the station area is statistically significant at the 0.05 level. Its *b*-coefficient implies an increase of 2.3 passengers for every 100 employees within walking distance. This coefficient is reasonable, being higher than the percentage of LRT riders in the population as a whole, but believable for those with an LRT station near their jobs. The significance of the employment variable is especially impressive given the variety of ways the data were estimated by the different cities. That this variable is significant for a study of both downtown and outlying stations is an important finding of this study. A job is a job regardless of location, and a certain amount of employees take light rail to their workplace even if it is located in a suburban shopping center, a local commercial street, or an office or industrial park.

This finding is reinforced by the fact that the CBD dummy variable is decidedly not significant. To confirm this conclusion, the model was run again with the 12 primary variables and the CBD variable, which remained not significant ($t = 1.02$, $p = 0.31$ in the standard model and $t = 0.98$, $p = 0.33$ in the corrected model). The non-significance of the CBD dummy variable is another important result. This indicates that the main CBD no longer plays a special, privileged role in light-rail transit. Downtown stations get more boardings, but this can be attributed to their measurable characteristics such as high employment density, many bus routes, high renter percentage, interline transfer stations, and greater centrality. Other places with similar characteristics such as edge cities or urban villages can expect to support ridership accordingly.

The conclusion that there is no remaining unexplained variation in ridership at CBD stations must, however, be qualified somewhat. Most of the large positive residuals in the model (where boardings are higher than predicted) are located in the CBD. This initially led us to believe that our model was missing an important independent variable that accounted for extra-high CBD boardings. However, most of the large negative residuals are also CBD stations. The standardized residuals also show that the model fits CBD stations less well than non-CBD stations. CBD stations account for nine of the ten largest negative standardized residuals and seven of the ten largest positive standardized residuals. The CBD remains an area of extremes for light-rail transit, but it is not consistently under or over predicted.

Another reason to qualify our conclusions is that the cities included in this sample are not the most extreme in terms of road congestion and parking cost and availability, which provide added incentive for transit use. This was not by design, but neither is it by coincidence. In order to compare equivalent systems, the sample excluded LRT systems with numerous underground stations, such as Boston's and Philadelphia's. The same forces of high land values, dominant CBDs, and high densities that prompted the adoption of underground LRT stations, however, also limit parking and create congestion. Therefore, it is possible that the CBD variable could be more significant in a sample that includes systems with underground stations.

The centrality variable is highly significant at the 0.0001 level, with a negative *b*-coefficient. For each 1% increase in relative travel time to all other stations, a station loses 18.72 boardings. It could be argued that this variable is the missing signal of the CBD's special power to attract transit ridership, but we contend otherwise. Geographers and transportation analysts have long recognized the strength of distance decay and the importance of accessibility for all forms of transportation. The centrality variable exhibits a gradual U-shaped curve when graphed by location (Fig. 3). In contrast, the CBD dummy variable has a reverse-U shape that has been squared off, with no gradual transition—a station is either in the CBD or not. If centrality were simply a stand-in for CBD location, the CBD dummy variable, with its all-or-nothing shape, would have been a better statistical fit. The high level of significance of the centrality variable indicates that its effect on boardings matches the gradual curve of the variable itself.

It is also important to note that the CBD is not always the most accessible, centrally located area of a light-rail system. On three of the nine systems studied—Salt Lake City, Buffalo, and Cleveland (excluding the heavy-rail Red Line)—the CBD is located on a peripheral part of the light-rail network with the least accessibility. This fact makes it even less likely that the centrality variable is simply a stand-in for CBD location.

The significance of the centrality variable carries an important message to LRT planners: travel time matters. Planners cannot expect to put large numbers of stations between origins and destinations without slowing down travel and diminishing ridership. Given that the station spacing variable is not significant but the population variable is, a complex tradeoff emerges. Wider spacing of stations does not seem to gain stations higher ridership from the commuters beyond the walking distance. Closer station spacing would place more residents within walking distance of stations, raising system ridership, but the additional stops would be a disincentive to the more peripheral riders. No simple conclusion can be reached here, other than suggesting that the regression model be used predictively to test different proposed alignments and stations.

The next variable, population within walking distance, is highly significant, and indicates an additional 9.2 boardings per 100 nearby residents. It is not surprising that this coefficient is higher than that for employment, because employment is more concentrated around stations (employment averages 3900 per station, while population averages 1500). On average, 2.1% of PMSA population is near stations, compared with 16.9% of employment. This is consistent with Cervero's (1993) finding for San Francisco, Sacramento, and San Diego that the rate of transit use among station-area residents is almost twice as high as that of station-area workers.

The percentage of households that rent instead of own is used as an indicator of socioeconomic status. Despite the fact that renters are not synonymous with lower income and lower car ownership, this variable was significant at the 0.0001 level. Each increase of 1% point leads to an additional 6.24 boardings. Renter percentage is correlated with other variables in the study, such as population and employment, but not excessively. The fact of the matter is that the renter percentage is independent: it can be high in CBD areas of high employment and no single-family homes, in suburban apartment complexes with little employment, or in inner-city slums.

Degree-days carries a negative coefficient of −1.52, meaning that more extreme temperatures discourage LRT ridership. Cities blessed with temperate climates, such as San Diego, can expect up to 300 more boardings per station than the average. Cities like Buffalo, Cleveland and Salt Lake City, which experience more extreme temperatures, would see boardings at each station reduced by a similar amount. The magnitude of this relationship might be dismissed as exaggerated were it not for the extremely high significance. The implication for cities planning LRT systems is large. Cities with extreme climates such as Phoenix and Minneapolis should not expect to match the ridership of more mild cities, where walking and waiting for transit is more comfortable. Furthermore, they should build covered waiting areas if they hope to mitigate this effect.

The two intermodal variables—number of bus connections and number of park-and-ride spaces—are both highly significant at the 0.001 level or better. Both have positive *b*-coefficients that make a good deal of sense. Each additional bus route intersecting with a station yields 123 weekday boardings, while each additional park-and-ride space nets 0.77 boardings. Given that some spaces will go empty while others may contain a car that brought more than one passenger, the 0.77 increment per space seems reasonable. These figures could possibly be used along with infrastructure cost data to compare the two intermodal strategies for cost effectiveness. However, as Parsons Brinckerhoff Quade & Douglas, Inc. (1996) pointed out, park-and-ride lots are a "double-edged sword," because while they provide intermodal access, they preclude in-close housing and employment.

The most significant variable and the one with the highest coefficient (+12,055 boardings) is the dummy variable for the international border station. This dummy variable was created to account for the extreme value of San Ysidro (Tijuana) station at the southern terminus of San Diego's Blue Line. Without adding a dummy variable, this station would have had an extreme positive residual. The choice of whether to eliminate this case or add a dummy variable to explain it came down to whether San Ysidro's incredibly high boardings might be duplicated elsewhere. Twin cities are sprouting all along the US–Mexico border. If El Paso were to build an LRT connecting to Ciudad Juàrez, it seems conceivable, even likely, that it would experience a similarly large contingent of day-workers, tourists, relatives, and shoppers who either do not own cars or do not want to bring a car through customs. To make sure the results were not overly influenced by this dummy variable that applies to only one case, the model was run without the border variable and the San Ysidro case. In models corrected and uncorrected for heteroscedasticity, the coefficients and significance of the remaining independent variables remains virtually identical. This makes sense, because when there is only one dummy case, the Y-intercept shift is exactly what is needed to perfectly predict the lone case with a residual of 0. Because the border variable does not affect the estimation of the other variables, we left it in the model to provide some numerical guidance for the benefit of other border cities contemplating LRT. Leaving it in does, however, raise the R^2 value. Removing the border variable and the San Ysidro case lowers the R^2 from 0.727 to 0.626.

The second largest coefficient in absolute terms is +5735 for the transfer stations in San Diego and Cleveland. These are the only cities in the study with more than one major light-rail line, though Baltimore and Dallas have spurs and Salt Lake and Portland are constructing second lines. This variable is partly an artifact of the counting of unlinked trips.

The final significant dummy variable is for terminal stations. Highly significant at 0.01, being at the end of a line adds 660 boardings per day (Parsons Brinckerhoff estimated a coefficient of 982). This variable is important for planning purposes, because bus and park-and-ride opportunities at these stations should be maximized. This variable has a larger number of cases (28) than the other dummy variables, making its coefficient estimate more reliable.

5.4. City dummy variables

The regression analysis treated all 268 stations as statistically independent observations. This assumption, however, is not entirely valid because several unmeasured factors, such as the system's age, fare structure, advertising, general condition, frequency of service, and the methods used by cities to measure employment and boardings, are constant across the stations within each city but different between cities. Other factors such as crime rates and congestion also vary between cities, although they are not constant across all of a city's stations. The analysis did include several systemwide variables such as degree-days, PMSA population, and employment coverage, however, these were numerical variables targeting specific factors. To investigate the possible effect of this type of spatial autocorrelation, dummy variables for each city were added to the analysis one at a time and in groups.

In general, these statistical analyses with the city dummy variables did not substantially alter the results. The only city dummy variable that was consistently significant was Cleveland, with a negative coefficient. Baltimore (negative) and Dallas (positive) were sometimes significant as well. Depending on which combination of these three dummy variables was included, the airport and employment coverage variables sometimes slipped below the 5% significance level. Similarly, when airport and employment coverage were removed, the Baltimore and Dallas dummy variables were no longer significant. Also, when the Dallas dummy variable was included, the significance of the station-area employment variable slipped to between 0.06 and 0.09. Given that several stations in the Dallas CBD contained the highest employment levels in the entire sample, this is not unexpected. Overall, these analyses confirm that airport and employment coverage are marginally significant variables, and confirm the robustness of the others. We would group station-area employment among the robust relationships given the error introduced by the wide variety of ways in which the nine cities provided employment data.

6. Conclusions

The recent resurgence of light rail after its almost universal demise in the first half of the 20th century is one of the most remarkable turnarounds in transportation history. It is an idea whose time appears to have come, gone, and come again. Much doubt, however, swirls around the second coming of light rail. Critics offer dire predictions of empty, subsidized trains eschewed by habitual drivers whose origin–destination patterns resemble an everywhere-to-everywhere scramble. Proponents hope light rail will gain moderate ridership, marginally reduce congestion and air pollution, promote infill development, and provide an alternative mode with higher capacity than buses along busy corridors.

In the midst of the current LRT explosion, it is important to replace suppositions with facts and analysis. This paper has analyzed a cross-section of 268 actual stations in nine different cities representing a variety of different conditions. Baltimore, Buffalo, Cleveland, and St. Louis were built up during the first streetcar era, whereas San Diego, Sacramento, Salt Lake City, and Dallas are exemplars of postmodern western urban places. Portland, the ninth city, stands somewhat alone as a progressive example of anti-sprawl planning.

The study found 10–12 significant variables for explaining ridership (depending on the statistical model and the cutoff for significance), and five decidedly non-significant factors. The large number of significant variables means there are many factors at work in generating ridership. In every case, the size and direction of the b-coefficients was intuitively reasonable. At least one independent variable from each of the five major categories was significant. From the land use/traffic generator category, employment, population, international borders, and airports were relevant. From the network structure category, terminal and transfer stations, and centrality were significant, while employment coverage was nearly so. From the citywide variables, extreme temperatures were a significant detriment with a sizable impact. Buses and park-and-ride were shown to be significant determinants of ridership from the intermodal category, as was renters, the only socioeconomic factor studied.

By performing the analysis at the station level using highly disaggregated employment and population data and accurate network-based buffers generated by GIS, this study has been able to assess the influence of land use on light-rail transit. This analysis tells us that a station need not be downtown to generate substantial ridership. All else being equal, each 100 jobs leads to 2.3 boardings; each 100 residents to 9.3 boardings; each 100 park-and-ride spaces to 77 boardings; each bus to 123 boardings; and an airport to 913 boardings. Stations outside the CBD with substantial amounts of these factors can generate boardings over 2000 per day—as 25 non-CBD stations actually did. A well-designed alignment with intermodal access passing through major employment nodes and fairly dense neighborhoods should be able to generate substantial ridership even in the absence of a mega-downtown. There is no guarantee that such an alignment is possible in every American city, but if so, there is reason to believe that light rail can succeed.

Not only can LRT boardings be explained to a high degree by these factors, but the same factors explain both CBD and non-CBD boardings. There was no consistent unexplained variation that could be swept up into a significant CBD dummy variable. The CBD stations generate higher boardings, but that is explainable by the same employment, population, bus, parking, and accessibility factors that are present in lesser combinations at non-CBD stations and account for their generally lesser boardings. Still, the fact that most of the extreme positive and negative standardized residuals are for CBD stations strongly suggests that other factors affecting CBD boardings have been left out, such as: parking costs and congestion; venues for sports, arts, retail, and entertainment; and greater densities very close to the station.

As Parsons Brinckerhoff argued, a travel demand model such as this combines features from all four steps of the traditional UTMS model. It is clearly not as comprehensive and systematic as the UTMS process, lacking as it does explicit consideration of the overall origin–destination (O–D) flow pattern or competing modes. Yet its transparency and high R^2 may make it useful for experimenting with different alignments of new LRT corridors, for which the values of all independent variables could be easily calculated. The most difficult problem would be dealing with temporal changes. The independent variables will capture some of the urban changes over time, such as dispersion or intensification, but the changing basis of competition with other modes would seem to give this model limited shelf-life for predictive purposes.

Any data-intensive study such as this—covering a number of cities with locally generated data—is bound to have some weaknesses. Consistent employment data and more disaggregated population data are the greatest needs, but even the dependent variable is collected in different ways at different times of year by different cities. Also, would a more sophisticated distance assumption than all-or-nothing buffers work even better? Sports arenas can be major traffic generators for light rail, which can greatly reduce game-day congestion. However, they could not be included in this study due to inconsistencies in how, when, and where sports trips were counted. Other land uses that generate non-employee ridership are schools, stores, zoos, museums, parks, restaurants, and theaters, but such data were beyond the scope of this study.

At the conclusion of this study, a host of other questions remain unanswered. Would heavy rail or vintage light rail be explained by the same set of factors? Second, the interaction of park-and-ride spaces and bus connections with station spacing is worthy of further research. One would expect that there is more opportunity for intermodal connections to raise ridership when the spacing between stations is greater. Third, what explains the extreme lack of significance of college enrollments, which logically should generate additional ridership? Denver's light-rail system, which was not included because of data problems, nevertheless reported that their system ridership drops by 3000–4000 during the Auraria campus holidays. Fourth, do heating degree-days and cooling degree-days have the same effect on boardings?

International comparisons on a station by station basis would also help answer some difficult questions about confounding influences. If ridership is higher outside the US because the cities are denser and car ownership is lower, then that might be adequately captured by the existing set of variables (with car ownership replacing percent renters). If, the model still underpredicted boardings at non-US stations (or a US dummy variable were significant), it might point to cultural or political factors being involved. Gordon and Willson (1984) constructed a similar model of light-rail demand at the city level of aggregation and found causal factors such as density, income, and spacing to be stable across boundaries. Having modeled CBD and non-CBD stations with a universal model, a logical next step might be to develop a model that applies equally to US and non-US transit, as well as light and heavy rail.

Acknowledgements

This data-intensive project would have been unthinkable without the cooperation of the transit agency and metropolitan planning organization in each city studied. Many individuals too numerous to mention by name selflessly prepared data and answered numerous questions. Although two of the original eleven light-rail systems had to be excluded due to data problems, we are still very appreciative of the effort made by those agencies. Some of the ideas for this study originated in a meeting with Mary O'Connor, Robert Yabes, Jyme Sue Olson, and Gregg Kent of the Tempe Transit Department. Arizona State University provided student funding under the Graduate College Support Program. Finally, we would like to thank C. Scott Smith and Michael Zoldak of the ASU GIS Lab and cartographer Barbara Trapido-Lurie.

References

Adams, J.S., 1970. Residential structure of Midwestern cities. Annals of the Association of American Geographers 60, 48–63.

American Public Transportation Association, 2002. Available from <http://www.apta.com/sites/transus/lightrail.htm>.

Barber, G., 1995. Aggregate characteristics of urban travel. In: Hanson, S. (Ed.), The Geography of Urban Transportation, second ed. Guilford Press, New York, pp. 81–99.

Bly, P.H., Webster, F.V., Pounds, S., 1980. The effect of subsidies on urban public transport. Transportation 9, 311–331.

Brennan, J., Hill, E.W., 1999. Where are the Jobs: Cities, Suburbs, and the Competition for Employment. The Brookings Institution, Washington, DC.

Bureau of Transportation Statistics, 2000. Transportation Statistics Annual Report 2000. US Department of Transportation, Washington, DC. Available from <http://www.bts.gov/publications/tsar/2000/pdf/entire.pdf>.

Cervero, R., 1993. Ridership Impacts of Transit Focused Development in California. University of California, Berkeley.

Clark, W.A.V., Hosking, P.L., 1986. Statistical Methods for Geographers. Wiley, New York.

Federal Highway Administration, 2000. Summary of travel trends: 1995 nationwide personal transportation survey. US Department of Transportation, Washington, DC.

Fielding, G.J., 1995. Transit in American cities. In: Hanson, S. (Ed.), The Geography of Urban Transportation, second ed. Guilford Press, New York, pp. 287–304.

Filion, P., 2001. Suburban mixed-use centres and urban dispersion: what difference do they make? Environment and Planning A 33, 141–160.

Gordon, P., Willson, R., 1984. The determinants of light-rail transit demand—an international cross-sectional comparison. Transportation Research A 18, 135–140.

Giuliano, G., 1995. Land use impacts of transportation investments: highway and transit. In: Hanson, S. (Ed.), The Geography of Urban Transportation, second ed. Guilford Press, New York, pp. 305–340.

Greene, W.H., 1998. LIMDEP Version 7.0 User's Manual. Econometric Software, Plainview, NY.

JHK and Associates, 1987. Development-Related Ridership Survey. Washington Metropolitan Area Transit Authority, Washington, DC.

Kain, J.F., Liu, Z., 1999. Secrets of success: assessing the large increases in transit ridership achieved by Houston and San Diego transit providers. Transportation Research A 33, 601–624.

Leinberger, C., 1993. Suburbia Land Use in Transition: Emerging Forces and Issues Shaping the Real Estate Environment. Urban Land Institute, Washington, DC.

Mason, J.W., 1998. The buses don't stop here anymore: sick transit and how to fix it. American Prospect 37, 56–62.

Morrill, R.L., 1991. Myths about metropolis. In: Hart, J.F. (Ed.), Our Changing Cities. Johns Hopkins Press, Baltimore, pp. 1–11.

National Oceanic and Atmospheric Administration (NOAA), 2002. Climate Prediction Center (2002) Degree Days Statistics. Available from <http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/cdus/degree_days/>.

Nelson/Nygaard Consulting Associates, 1995. Land use and transit demand: the transit orientation index. Primary Transit Network Study (Draft), Chapter 3. Tri-Met, Portland, OR.

Parsons Brinckerhoff Quade & Douglas, Inc., 1996. Transit and Urban Form. National Academy Press, Washington, DC.

Pickrell, D., 1992. A desire named streetcar: fantasy and fact in rail transit planning. Journal of the American Planning Association 58, 158–176.

Plane, D., 1995. Urban Transportation: Policy Alternatives. In: Hanson, S. (Ed.), The Geography of Urban Transportation, second ed. Guilford Press, New York, pp. 435–469.

Pushkarev, B.S., Zupan, J.M., 1977. Public Transportation and Land Use Policy. Indiana University Press, Bloomington, IN.

Pushkarev, B.S., Zupan, J.M., 1982. Where transit works: urban densities for public transportation. In: Levinson, H.S., Weant, R.A. (Eds.), Urban Transportation Perspectives and Prospects. Eno Foundation, Westport, CT, p. 341.

Rosenbloom, S., Clifton, K., 1996. The puzzle of income, race, and density: preliminary evidence on transit use from the 1991 American Housing Survey. Journal of Public Transportation 1, 87–102.

Shaw, S.L., 1991. Urban transit accessibility analysis using a GIS: a case study of Florida's Tri-Rail System. Southeastern Geographer 31, 15–30.

Shoup, D., Willson, R., 1992. Employer-paid parking: the problem and proposed solutions. Transportation Quarterly 46, 169–192.

Shrank, D., Lomax, T., 1999. The 1999 annual mobility report: information for urban America. Texas Transportation Institute, College Station, TX.

Simons, R.A., El Jaouhari, A., 2001. The Effect of Railroad Tracks and Train Activity on Residential Property Values in Cuyahoga County, Ohio. American Real Estate Society, Coeur D'Alene, ID.

Smith, W., 1984. Mass transit for high-rise, high-density living. Journal of Transportation Engineering 110, 521–535.

The 2002 World Almanac and Book of Facts, 2002. World Almanac Books, New York.

Untermann, R., 1984. Accommodating the Pedestrian: Adapting Neighborhoods for Walking and Bicycling. Van Nostrand Reinhold, New York.

Upchurch, C., Kuby, M., Zoldak, M., Barranda, A., 2004. Using GIS to generate mutually exclusive service areas linking travel on and off a network. Journal of Transport Geography 12, 23–33.

US Census Bureau, 2002. Redistricting Data. Available from: <http://www.census.gov/clo/www/redistricting.html>.

US Department of Transportation, 1984. The Status of the Nation's Local Public Transportation. US Government Printing Office, Washington, DC.

Wachs, M., 1989. US transit subsidy policy: in need of reform. Science 244, 1545–1549.

Wachs, M., Taylor, B.D., 1998. Can transportation strategies help meet the welfare challenge? Journal of the American Planning Association 64, 15–19.

Wachs, M., Taylor, B.D., Levine, N., Ong, P., 1993. The changing commute: a case-study of the jobs-housing relationship over time. Urban Studies 30, 1711–1729.

White, H., 1978. A heteroscedasticity consistent covariance matrix and a direct test for heteroscedasticity. Econometrica 46, 817–838.