



On the social desirability of urban rail transit systems

Clifford Winston^{a,*}, Vikram Maheshri^b

^a *Brookings Institution, 1775 Massachusetts Avenue, NW, Washington, DC 20036, USA*

^b *U.C. Berkeley, Department of Economics, Berkeley, CA 94720, USA*

Received 22 February 2006; revised 21 July 2006

Abstract

Despite a decline in its mode share, investment to build new urban rail transit systems and extend old ones continues. We estimate the contribution of each U.S. urban rail operation to social welfare based on the demand for and cost of its service. We find that with the exception of BART in the San Francisco Bay area, *every* system actually reduces welfare and is unable to become socially desirable even with optimal pricing or physical restructuring of its network. We conclude rail's social cost is unlikely to abate because it enjoys powerful political support from planners, civic boosters, and policymakers.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Public transit; Network variables; Automobile congestion

1. Introduction

The evolution of urban rail transit in the United States over the past twenty years has been marked by three inescapable facts that signal an inefficient allocation of transit resources. Rail's share of urban travelers is declining during a period when there has been little investment in new roads; its deficits are rising sharply; and yet investment to build new systems and extend old ones continues.

In 1980, two million Americans got to work by rail transit. Today, in spite of an increase in urban jobs and transit coverage, fewer than one million U.S. workers commute by rail, causing its share of work trips to drop from 5 percent to 1 percent.¹ Although rail transit's farebox

* Corresponding author.

E-mail address: cwinston@brookings.edu (C. Winston).

¹ These figures are from the National Transit Database and the U.S. Census and do not include commuter rail passengers. Note that these figures refer to *workers* not unlinked trips. For example, a worker who makes a round trip by rail

revenues have consistently failed to cover its operating and capital costs since World War II, governmental aid to cover transit deficits has been increasingly available. Since 1980, annual operating subsidies have climbed from \$6 billion to more than \$15 billion today (APTA Transit Fact Books, figures in 2001 dollars). Capital subsidies have also increased as transit agencies struggle to maintain and provide new facilities, track, and rolling stock.

These worrisome trends, however, have not curbed U.S. cities' appetite for rail transit service. During the 1990s, Cleveland, Washington, Santa Clara, Sacramento and other cities expanded their systems, while Los Angeles, Denver, Dallas, and St. Louis built new ones. Recently, Houston and Minneapolis opened new light rail lines while small, sparsely populated cities such as Sioux City, Harrisburg, and Staunton, Virginia suggested that they want federal funds to help build their systems. And although county residents repeatedly nixed a referendum to build a \$4 billion extension of Washington's Metro out to Dulles airport, planners nevertheless circumvented popular will and diverted increased toll revenue from the Dulles toll road to finance a portion of the ultimate extension.

Any private firm that was losing market share and reporting increasing losses would be hard pressed to attract funds to expand. Almost certainly, it would try to determine the most efficient way to contract. Of course, a transit agency does not seek to maximize profits, but its public financing is justified only if it is raising social welfare, where social welfare can be measured as the difference between net benefits to consumers and the agency's budget deficit, also taking into account relevant externalities (for instance, the reduction in roadway congestion attributable to rail).

Although the costs and benefits of public rail transit operations have been debated in the policy community (see, for example, Litman [1]), we are not aware of a recent comprehensive empirical assessment of rail's social desirability.² The purpose of this paper is to estimate the contribution of each U.S. urban rail operation to social welfare based on the demand for and cost of its service. We find that with the single exception of BART in the San Francisco Bay area, *every* U.S. transit system actually reduces social welfare. Worse, we cannot identify an optimal pricing policy or physical restructuring of the rail network that would enhance any system's social desirability without effectively eliminating its service.

Rail transit's fundamental problem is its failure to attract sufficient patronage to reduce its high (and increasing) average costs. This problem has been complicated enormously by new patterns of urban development. Rail operations, unfortunately, are best suited for yesterday's concentrated central city residential developments and employment opportunities; they are decidedly not suited for today's geographically dispersed residences and jobs. At best, urban rail service may be socially desirable in a few large U.S. cities if its operations can be adjusted to mirror successful privatization experiments conducted abroad. Ironically, however, rail transit enjoys powerful political support from planners, civic boosters, and policymakers, making it highly unlikely that rail's social cost will abate.

that involves a single connection makes four unlinked trips. Overall rail transit ridership has not declined as sharply as its decline in commuter patronage. However, rail's ability to reduce roadway congestion decreases as its traffic mix includes a greater share of non-commuters because such travelers are much more likely than commuters to travel off peak.

² Richmond [2] provides an overview of U.S. rail transit systems including their ridership and financial performance. Winston and Shirley [3] estimate the net benefits of urban rail systems as of 1990. Other researchers such as Savage [4], Savage and Schupp [5], Kain [6], and Viton [7] have assessed the welfare properties of and alternative policies to improve individual transit systems.

2. An empirical framework for estimating rail transit's social benefits

Urban rail transit operators do not set prices to cover operating and capital costs. In fact, under Federal Transit Administration Section 5307, transit fares (including bus, rail, and paratransit) have to cover only some 17 percent of operating costs for most agencies to qualify for federal funds. State and municipal funding thresholds vary, but none requires even half of operating costs to be covered at the farebox. On average, the nation's rail transit systems cover only about 40 percent of operating costs, to say nothing of their substantial capital costs (National Transit database). Like any good or service, rail's net benefits to users are simply given by consumer surplus, but to justify continued operation, rail's surplus must offset the difference between farebox revenues and costs.³

We develop rail transit demand and cost models to estimate users' benefits and agencies' budget deficits. We also account for rail's effect on the cost of roadway congestion. A novel feature of the models is that they include variables that characterize a transit system's network configuration and stations, enabling us to explore whether rail could enhance its social benefits by expanding or contracting its facilities in an efficient manner.⁴

Demand

Our empirical analysis is conducted on a panel of U.S. urban rail transit systems. We specify travelers' demand (in passenger miles) for rail transportation, Q_{it}^D , on system i during year t as:

$$Q_{it}^D = D(p_{it}^D, Z_{it}, X_{it}; u_{it}), \quad (1)$$

where p_{it}^D is the average fare, Z_{it} contains exogenous network variables, X_{it} contains exogenous city characteristics, and u_{it} is an error term.⁵

Depending on the system, transit fares are determined by a transit agency, metropolitan planning organization, or city council. It is therefore reasonable to treat the average fare as exogenous because it is primarily determined through a regulatory process rather than market forces. In addition, the policy bodies that set fares have little incentive to adjust them to changing market conditions because as noted they are not subject to stringent financial performance goals to qualify for government funds. In any case, we test the assumption that fares are exogenous.

Rail demand is also influenced by the configuration of a transit system's network. Travelers are more likely to use rail if it provides more comprehensive coverage of a given area, offers more conveniently located stations, offers more connectivity, and accommodates travel in both north–south and east–west directions (i.e., its network is non-linear). The following components of a transit network enable us to derive four metrics from graph theory to capture the effect of a rail system's entire network on demand⁶:

³ Rail transit also directly generates a small portion of its revenues from advertisements, parking lot fees, and other auxiliary services. We explicitly exclude these revenue streams and their associated production costs when we perform our analysis.

⁴ An alternative (indirect) measure of rail transit's benefits is rail's impact on housing prices in the surrounding residential area near stations. Diaz [8] summarizes studies of the effect that some systems have had on residential property values. Baum-Snow and Kahn [9] provide a recent study.

⁵ We will interchange the term city with urbanized area and metropolitan (statistical) area (MSA). Urbanized areas and MSAs are determined by U.S. Census demographic criteria; nonetheless, they are often associated with a distinct city. Data on rail transit systems tend to pertain to a metropolitan statistical area.

⁶ Hagget and Chorley [10] provide a complete discussion of the measures.

- n number of stations served (nodes);
- e number of station to station links (edges);
- $l_{p,q}$ length of a particular link from station p to station q ;
- d length of the shortest route between the farthest two stations (diameter);
- A total land area served by the system.

Based on these components, the relevant network variables are:

$D = \frac{1}{A} \sum_{p \neq q} l_{p,q}$, indicating the density of the network. A network has greater density if it provides more track length in its service area.

$\eta = \frac{1}{e} \sum_{p \neq q} l_{p,q}$, indicating the average length of all links in the network. Smaller values of η indicate that stations tend to be closer together, which can be convenient for travelers.

$\Gamma = \frac{e}{3n-6}$, a measure of connectivity that is equal to the ratio of the actual number of links and the maximal number of links, given a set of stations. This is a standard, efficient measure of the evolution of a network over time. Greater connectivity can improve access to different points on the network.

$\pi = \frac{1}{d} \sum_{p \neq q} l_{p,q}$, indicating the (non-)“linearity” of the network. The measure takes on a minimum value of one, characterizing a perfectly linear network. Larger values characterize less linearity and thus broader coverage of a given geographical area.

We also control for rail’s service with the ratio of its peak service frequency to base service frequency.⁷ Because rail and bus systems are often designed to complement each other, we also control for bus transit’s peak-to-base ratio. Larger rail and bus frequency peak-to-base ratios should increase demand.

Turning to the characteristics of a city that may influence rail demand, we control for regional gasoline prices, the average number of days below 32 degrees Fahrenheit, and the average commute time. Higher gasoline prices may induce commuters to switch from driving, which would increase transit use, but they may also depress overall economic activity, which would tend to decrease transit use. Thus, the a priori effect of gasoline price is ambiguous. Cold temperatures tend to increase transit ridership because they discourage commuters, in particular, from walking and biking. A high average commute time in a city is likely to be caused by lengthy work-trip distances and factors that contribute to road congestion. The former may reduce travelers’ accessibility to rail and decrease the demand for transit, while the latter may increase the demand for rail transit as travelers try to avoid congestion. Thus, the a priori effect of average commute time is also ambiguous.⁸ Finally, we include the metropolitan area population and resident households’ median annual income. Demand for rail transit should be positively related to an area’s

⁷ Frequencies are reported only as a ratio, not separately. The peak service frequency is defined as the maximum average frequency of the 7–10 am morning and 4–7 pm evening rush hours. The base service frequency is defined as the average frequency of the 10 am – 4 pm off-peak period. It is reasonable to treat the frequency ratio as exogenous because the timing and duration of the peak are fixed and not influenced by demand, while peak and off-peak schedules are largely determined by labor contracts. We also tried to capture service frequency by specifying a system’s vehicle miles (seat miles are not available), but the peak-to-base frequency ratio produced more plausible and reliable estimates.

⁸ It is possible that average commute times may be higher in cities that have a measurable share of rail transit users, which would suggest that average commute time may be endogenous. We consider this possibility in our estimations as well as explore an alternative specification using average auto time. We also tried using average commute distance

population. Although transit is often regarded as an inferior good, the average income of travelers who use rail is much higher than the average income of travelers who use bus (Winston and Shirley [3]), so rail transit may be a normal good.⁹

Cost

Regulations constrain rail transit agencies from abandoning or adding track and stations to optimize their operations; hence, it is inappropriate to assume that they are in long-run equilibrium. We therefore specify a short-run total cost function where we include the network variables discussed previously to control for the capital inputs of track and stations that are fixed in the short-run. Formally, short run total costs, C_{it} , for transit system i during year t are expressed as:

$$C_{it} = C(Q_{it}, w_{it}, Z_{it}, Y_{it}; v_{it}), \quad (2)$$

where Q_{it} is output in passenger miles, w_{it} contains factor prices, Z_{it} contains exogenous network variables, Y_{it} contains other exogenous influences on cost, and v_{it} is an error term.¹⁰

To the extent that rail transit fares are exogenously determined, it is reasonable to treat output as exogenous in the cost function.¹¹ We include factor prices for labor and fuel. The price of labor is given by the hourly wage for transit workers, including fringe benefits. Because different rail transit systems use different combinations of fuel types (gas, electricity, kerosene, ethanol, and so on), we computed a standardized price per kilowatt-hour of energy using the appropriate physical constants (i.e., KWH equivalents) for each one.¹² We capture the effect of capital on short-run costs by including the network variables that we used in the demand model. Holding output constant, we expect networks with greater density, connectivity, or more closely situated stations to have higher costs. The a priori effect of network linearity on costs is not clear; greater linearity requires a system to use less rolling stock to serve its stations, while greater non-linearity may reduce costs by enabling a system to improve the efficiency of its scheduling and the allocation of its equipment. Finally, the other exogenous variables we include in the cost function are

instead of average travel time to work but this variable was poorly measured and had virtually no effect on demand. The poor measurement of distance also precluded us from using average speed.

⁹ We explored additional and alternative variables to capture service quality and city characteristics but they were statistically insignificant and did not affect the other parameters. The variables we explored were: a dummy variable for systems with predominantly heavy rail service because heavy rail systems can operate at higher speeds than light rail systems; average snowfall and a light rail dummy variable interacted with snowfall or cold temperatures because light rail stations tend to be above ground; a dummy variable for new systems—those that had operated for five years or less—because some travelers might be attracted to rail because of its novelty; and the age of a city—measured by a dummy variable equal to one if the city had a population of at least 100,000 in 1900 and zero otherwise—to capture development patterns that may be favorable to rail use.

¹⁰ Urban rail transit cost functions have been estimated previously by Pozdena and Merewitz [11], Viton [12], and Savage [13].

¹¹ Although absolute measures of peak and off-peak passenger miles were not available, we tried to specify peak and off-peak output by interacting passenger miles with peak-to-base ratios, but this did not enable us to obtain reliable estimates of distinct outputs. We also specified peak and off-peak vehicle miles (as noted, seat miles were not available), but these measures of capacity are apparently too crude to improve the model. We also tried to control for different “qualities” of output by including the heavy rail dummy noted in footnote 9 to capture potentially faster travel times provided by systems that offered predominantly heavy rail service, but the dummy was insignificant.

¹² We interacted the heavy rail dummy with fuel price because heavy and light rail have different fuel intensities, but the interaction term was statistically insignificant. We also specified a city’s debt to revenue ratio to control for a city’s overall financial health. A city in poor financial health may defer its system’s maintenance, which could increase costs. However, we found that this variable was statistically insignificant.

Table 1
Variables, sample means, and sources*

Variable	Mean	Source
Fare per passenger mile	0.197	National Transit Database
Short-run total costs (millions)	393	National Transit Database
Passenger miles (millions)	528	National Transit Database
Total track (miles)	118.2	National Transit Database
Total stations	60.4	Transit System Network Maps
Service area (square miles)	1132	National Transit Database
Links between stations	57.74	Transit System Network Maps
Transit system age (years)	30.79	Transit System Agencies
Average energy cost per KWH	0.002	National Transit Database
Average hourly wage plus benefits	32.14	National Transit Database
Metropolitan area population (millions)	3.713	US Census Bureau
Median annual household income	50,613	Texas A&M Real Estate Center
Average regional gasoline price per gallon	1.44	US Department of Energy, Annual Energy Review
Average travel time to work (minutes)	25.24	Department of Transportation, Journey to Work Survey
Average number of days annually below 32° F	63.2	National Meteorological Survey
Rail peak-to-base ratio	2.21	National Transit Database
Bus peak-to-base ratio**	1.85	National Transit Database
Municipal debt-to-revenue ratio	1.16	US Census of Governments
Number of observations	194	

* Fares, costs, wages, and incomes are in 2000 dollars.

** In the few cities where multiple bus systems served the rail transit agency's service area, we selected the largest bus system.

transit system age and the average number of days below 32 degrees Fahrenheit. Newer systems may depreciate less rapidly than older systems but have higher operating costs.¹³ Colder climates are likely to accelerate the depreciation of a transit system's capital stock but have an uncertain effect on operating costs.

Sample

We include the twenty-five U.S. urban rail transit systems that were in operation between 1993 and 2000 generating 194 observations (a few systems began operations midway through the decade).¹⁴ These systems are distinct from various commuter rail systems in the nation that mainly link distant suburbs and nearby urban areas to a city's rail and bus transit system. The sample accounts for all urban travel by light and heavy rail transit during the period.

Data sources for the variables used in the demand and cost models and their sample means are presented in Table 1. We measure fares on a passenger-mile basis. As an alternative, we measured fares per trip but our basic findings were unaffected. Short-run total costs are defined as the sum of operating costs and depreciation and amortization of rail transit capital. To construct

¹³ In our sample discussed below, system age and the costs of depreciation are positively related. The age of a transit system is based on the year that a system began offering rail service to the public.

¹⁴ We began our panel in 1993 because some variables in the National Transit Database were defined differently before that year. We did not extend our sample beyond 2000 because we thought it would be difficult to obtain long-lived information about the welfare properties of transit systems in the tumultuous period that followed. Finally, we did not include the King County (Seattle) transit system and the Detroit Metropolitan transit system because they just began rail operations in 2000.

Table 2

Minimum and maximum values over the sample period for selected variables and systems*

Agency	Min. fare	Max. fare	Min. passenger miles	Max. passenger miles	Min. total cost	Max. total cost
New York City Transit	0.23	0.26	5570	8320	1520	4370
Mass. Bay Transit Authority (Boston)	0.17	0.35	536	668	312	1010
MTA Maryland (Baltimore)	0.13	0.17	84.8	130	185	253
Niagara Frontier (Buffalo)	0.25	0.29	15.1	19.5	34.4	59.6

* Fares per passenger mile are in 2000 dollars, passenger miles are in millions, and total costs are in millions of 2000 dollars.

the network variables, we obtained annual transit network maps for each system and traced their expansion using detailed city maps. It is useful to note that although our sample includes twenty-five systems, trips on New York City's system account for roughly two-thirds of the nation's rail transit passenger miles. Table 2 shows that during our sample period, average fares, passenger miles, and costs have increased by varying degrees for a representative system from each size classification.

3. Estimation results

We explored various functional forms for the demand model including linear, log-linear, quadratic, and flexible forms with interacted network variables. The best statistical fit, which also captured the heterogeneity of the systems, was obtained with a simple linear functional form that allowed the coefficients for fares and the network variables to vary by system size. As noted, the New York City system is quite extensive and carries far more passengers than any other system. This system along with the Washington, San Francisco, Chicago, Philadelphia, Boston, Atlanta, and Northern New Jersey systems comprise the so-called "Big 8" U.S. transit systems; we therefore identified some distinct effects for New York's system and defined "large" systems to include the other seven systems. We then divided the remaining seventeen systems into seven "medium" systems and ten "small" systems.¹⁵

Estimation results are presented in Table 3. We controlled for fixed system, state, and year effects, but only the year effects were statistically significant so they are included in the final model. The high R^2 is not attributable to any particular socioeconomic characteristic(s) that may be believed to correlate strongly with passenger demand. For example, when we omitted population from the specification, the R^2 was still quite high; a small decrease in R^2 also occurred when we omitted income. Note that a strong relationship between rail transit demand and population

¹⁵ Large systems were defined as having more than 250 million passenger miles annually. These systems included Washington, DC Metro, BART (SF Bay Area), Chicago Transit, MBTA (Boston), MARTA (Atlanta), SEPTA (Philadelphia) and PATH (Northern New Jersey). Medium systems were defined as having between 70 million and 250 million passenger miles annually. These systems included GCRT (Cleveland), MTA (Baltimore), PATCO (Southern New Jersey), San Francisco Municipal Railway, San Diego Trolley Inc., Los Angeles Metro and Metro-Dade Transit (Miami). Small systems were defined as having between 2 million and 70 million passenger miles annually. These systems included TriMet (Portland, OR), Santa Clara Co. Transit (San Jose), Staten Island Rapid Transit, Sacramento Regional Transit, Niagara Frontier Metro (Buffalo), Bi-State Development Agency (St. Louis), DART (Dallas), PA Allegheny Co. (Pittsburgh), RTD (Denver) and NJTransit (Newark).

Table 3
Demand coefficients, 1993–2000* (dependent variable is billions of passenger miles)

Variable	NYC Transit	Large systems ^a	Medium systems ^b	Small systems ^c
Fare (\$)	−37.1 (11.6)	−1.48 (0.521)	−3.44 (1.17)	−0.495 (0.223)
Track density (D)**	0.544 (0.133)	0.544 (0.133)	0.061 (0.126)	−0.184 (0.207)
Average distance between stations (η)**	−0.112 (0.057)	−0.112 (0.057)	−0.137 (0.047)	−0.115 (0.054)
Network connectivity (Γ)**	4.94 (1.93)	4.94 (1.93)	0.013 (0.094)	0.206 (0.163)
Network linearity (π) (greater values indicate less linearity)**	0.090 (0.027)	0.090 (0.027)	0.017 (0.017)	0.016 (0.016)
Constant	10.6 (3.40)	1.43 (1.95)	3.56 (1.62)	2.92 (1.52)
Common parameters				
Rail peak-to-base ratio			(0.053) (0.030)	
Bus peak-to-base-ratio			0.180 (0.068)	
Average number of days below 32 degrees			0.002 (0.001)	
Metropolitan area population (millions)			0.023 (0.014)	
Median household income in metro area (\$thousands)			0.003 (0.004)	
Average regional gasoline price (\$)			−0.371 (0.308)	
Average travel time to work (minutes)			−0.218 (0.105)	
Average travel time to work squared			0.004 (0.002)	
Year dummies			Included	
Number of observations	194		R^2	0.991

* Newey–West robust AR(1) standard errors shown in parentheses.

** We could not reject the hypothesis that the network coefficients were equal for NYC and large systems.

^a Large systems include Washington, DC Metro, BART (SF Bay Area), Chicago Transit, MBTA (Boston), MARTA (Atlanta), SEPTA (Philadelphia) and PATH (Northern New Jersey).

^b Medium systems include GCRT (Cleveland), MTA (Baltimore), PATCO (Southern New Jersey), San Francisco Municipal Railway, San Diego Trolley Inc., Los Angeles Metro and Metro-Dade Transit (Miami).

^c Small systems include TriMet (Portland, OR), Santa Clara Co. Transit (San Jose), Staten Island Rapid Transit, Sacramento Regional Transit, Niagara Frontier Metro (Buffalo), Bi-State Development Agency (St. Louis), DART (Dallas), PA Allegheny Co. (Pittsburgh), RTD (Denver) and NJTransit (Newark).

does not always exist (e.g., Los Angeles has a large population, but small transit ridership, while the opposite is true for Cleveland).

Turning to the coefficients, average fares have a negative and statistically significant effect on demand. The demand elasticities, computed by sample enumeration, are -1.3 for New York City, -0.97 for the large systems, -5.4 for medium systems, and -3.2 for small systems. Previous studies of large urban transit systems based on data in the 1970s and early 1980s tended to find that the demand for rail was inelastic (Winston [14] and Small [15] provide surveys). However,

given that these systems now offer more extensive suburban service and that populations in a metropolitan area are more geographically dispersed, rail faces more intense competition from auto, which would cause the demand elasticity to increase. For instance, Voith [16] finds that the elasticity of demand for the SEPTA commuter rail system, which includes but is not limited to serving Philadelphia, is -1.6 . In addition, real transit fares have risen substantially in the past two decades, which would also cause the demand elasticity to increase.¹⁶ It is therefore plausible that we now find that the demand elasticity for New York City and the large systems clusters around unity, with the New York City system having the most elastic demand among these systems because many travelers have the option of using auto, taxi, bus, or even walking to get to their destinations.¹⁷

We are not aware of previous estimates of demand elasticities for medium and small rail systems. In any case, these systems have a tiny share of travelers in the cities they serve (less than 1 percent), which undoubtedly contributes to their high elasticity of demand (i.e., modest responses in ridership to fare changes are proportionally large for smaller rail transit systems). In addition, these systems face considerable competition from auto because they operate in cities where residential and commercial development is predominantly influenced by and caters to automobile transportation. Despite their dependence on the auto, cities with medium and small rail systems tend to experience less congestion than cities with large rail systems.

As noted, we have assumed that the average fares for each system classification are exogenous in the demand equation. We tested this assumption with a Hausman specification test using as instruments for each fare variable input prices (wages and the composite energy price) and the municipal debt to revenue ratio (as a measure of a city's overall financial health). We could not reject the assumption of exogeneity at a high level of confidence.¹⁸

The network variables for New York City and the large systems are statistically significant and have their expected signs. We found that it was statistically justifiable to allow these systems to have the same coefficients for the network variables. Greater track density, connectivity, and (non-)linearity increase demand for rail, while a greater average distance between stations reduces demand. Greater average distances between stations also reduce demand for rail in medium and small systems. But the other network variables tend to have statistically insignificant effects on demand, possibly because these systems are not well developed and compete only for a small, specialized segment of travelers.

We did not reject the hypothesis that each of the remaining influences has the same effect on demand across rail transit systems. As expected, we find that an increase in the peak-to-base frequency ratio for rail or bus increases rail demand and that cities with larger populations and that experience more days below freezing have more rail users.¹⁹ Rail demand is also positively

¹⁶ According to the American Public Transit Association, real average transit (bus and rail) fares have increased 37.3 percent from 1980 to 2000. Unfortunately, data for just rail fares are not available during this period. In any case, the increase in transit fares undoubtedly reflects an increase in rail fares.

¹⁷ According to the Federal Highway Administration *Means of Travel to Work* survey in 2000, New York City leads the nation in the share of commuters who get to work by bus (6.8 percent), by walking or biking (5.9 percent), or by taking a taxi (0.8 percent).

¹⁸ The Hausman test statistic was 8.33 while the chi-squared critical value at the 99 percent level is 13.3. In addition, we found that the demand elasticities for each system based on instrumental variable estimation were very similar to the elasticities reported here.

¹⁹ It is possible that the weather variable is actually capturing a "rust belt/sun belt" effect of traditional cities in the northeast versus the more auto-friendly cities of the west and south. However, its effect persisted when we included regional dummies in the specification. Those dummies tended to be insignificant.

related to a city's median household income, indicating that rail transit is a normal good, but the effect is not precisely estimated. We find that the decline in real gas prices during our sample period caused rail demand to increase. This finding is consistent with the notion that gas prices tend to affect transit use by influencing the overall level of economic activity. In addition, because most suburban-based rail trips are combined with an auto trip, lower gas prices could increase rail demand by expanding its catchment area (Baum-Snow and Kahn, 2005). Finally, we find that the demand for rail decreases as average commute time increases, but this effect diminishes as commutes get longer and travelers are more willing to shift to rail to avoid roadway congestion even if they have to drive to outlying stations.²⁰

We initially explored estimating a cost function for rail transit using either a translog or quadratic flexible functional form along with factor demand equations for fuel and labor. However, we found that most of the interaction terms were poorly estimated and that the predictions of transit costs were implausible because New York City's dominant system made it impossible to specify a meaningful average (mean) transit system for the country. Hence, we settled on a simple, linear specification that fit the data well and produced plausible predictions of transit costs.

The coefficient estimates are presented in Table 4. Final estimations include state and year fixed effects (system fixed effects were statistically insignificant, possibly because the twenty-five systems in our sample are located in only thirteen distinct states and most transit systems within a state are subject to the same state regulations). In contrast to the demand model, we do not find that any of the determinants of short-run total cost vary by system size. Consistent with the presence of large deficits, the estimated marginal cost per passenger mile of 65 cents is considerably above the average fare per passenger mile of 20 cents reported in Table 1. The short-run scale elasticity, based on enumerating the ratio of the estimated marginal cost coefficient to average costs for each system in the sample, was 1.45 for New York City transit and 0.96 for the remaining systems. These estimates indicate that none of the systems is operating at short-run minimum cost.²¹ (Of course, the more relevant benchmark for production efficiency is long-run minimum cost.) Part of the problem is likely to be rail transit's low average load factor—less than 20 percent as of the mid-1990s (Winston and Shirley [3]). Since then, the Federal Transit Authority has stopped requiring transit systems to report load factor data.

Although some of the coefficients for the factor prices and network variables are imprecisely estimated, all of these influences have their expected signs. Increases in energy prices and wages raise costs. Greater track density and connectivity raise the cost of a system, while costs fall as distances between stations increase and as a network becomes more non-linear. Finally, we find that system age and weather are inversely related to costs but that these variables have interactive effects that may increase costs. Newer systems may have higher operating costs than older systems because they use more costly technologically sophisticated equipment and their agencies have less experience in troubleshooting these systems. On the other hand, older systems

²⁰ If the mode share of rail transit were influencing average commute time, we would expect that commutes that take longer would be associated with more demand for rail. However, we find the opposite relationship between commute time and demand, suggesting that average commute time is likely to be exogenous because of rail's small mode share. As a further check, we estimated a model that specified average *auto* commute time, which would be more likely than average commute time to be exogenous. We found that this variable had little effect on the coefficients, although t-statistics were somewhat lower than those obtained using average commute time.

²¹ We can reject the hypothesis that the non-New York transit systems are minimizing short-run costs (i.e., their scale elasticity estimate is equal to 1) at a high level of confidence. Our findings are broadly consistent with Viton [12], who finds diseconomies for the New York City system and slight scale economies for some smaller systems.

Table 4
Cost coefficients, 1993–2000* (dependent variable is total cost in millions of dollars)

Variable	Coefficient
Passenger miles (millions)	0.648 (0.063)
Energy price per KWH (\$)	7250 (7160)
Hourly wage plus benefits (\$)	2.183 (2.055)
Track density (D)	124 (74.0)
Network connectivity (Γ)	103 (48.5)
Average distance between stations (η)	−20.4 (18.3)
Network linearity (π) (greater values indicate less linearity)	−40.8 (18.4)
Average number of days below 32° F	−2.250 (1.690)
Transit system age	−9.047 (3.196)
Average number of days below 32° F*transit system age	0.168 (0.048)
Constant	514 (300)
Year dummies	Included
State dummies	Included
Number of observations	194
R^2	0.961

* Newey–West robust AR(1) standard errors shown in parentheses.

depreciate at faster rates than newer systems. Weather therefore interacts with system age because colder climates accelerate the depreciation of a system's capital stock.²²

4. Net benefits of urban rail transit systems

The net benefits of an urban rail transit system are the difference between users' consumer surplus and agency deficits that must be covered by publicly funded subsidies. (We will account for externalities later.) Given the inverse of our estimated demand function, $p(q)$, and the estimated short-run cost function, C , evaluated at the equilibrium output, q^* , we can express a system's net benefits, NB , as:

$$NB(q^*) = \left(\int_0^{q^*} p(q) dq - p \cdot q^* \right) + p \cdot q^* - C(q^*). \quad (3)$$

²² The weather variable may also be capturing regional differences in costs. But its effect persisted when we included regional dummies to the specification. We could not obtain reliable estimates of both regional and state dummies; we obtained a better fit and greater statistical significance using the state dummies instead of the regional dummies.

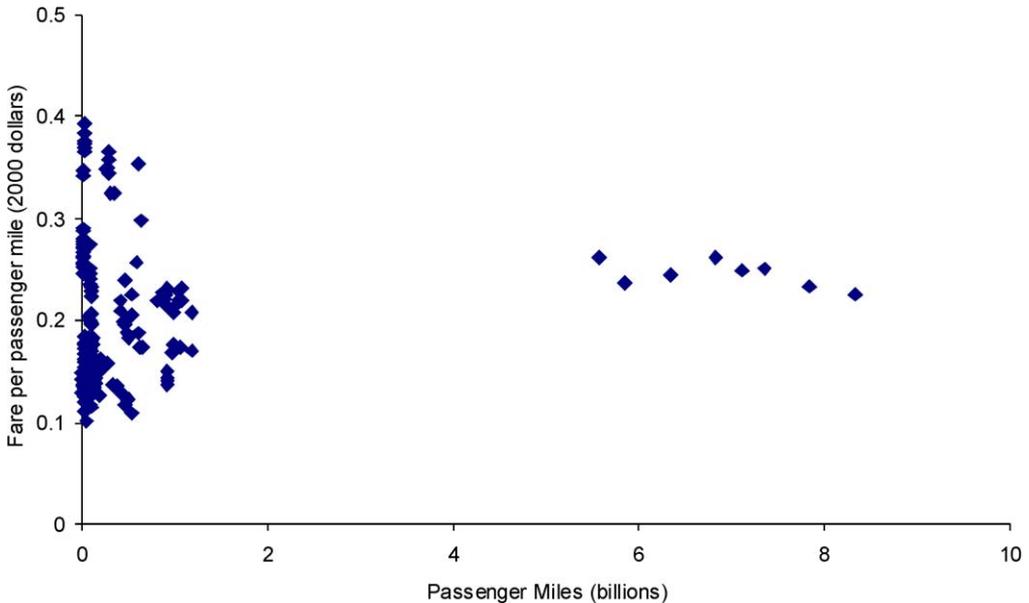


Fig. 1. Fares and traffic on US rail transit networks, 1993–2000. (The eight outlying points correspond to the New York City Transit Agency, which serves roughly two thirds of total national rail transit traffic.)

Note agency deficits are the difference between agency revenues and estimated short-run total costs, which include depreciation and amortization of transit capital. We therefore assume that transit systems would be able to divest themselves of these costs if their service were abandoned.

A valid concern can be raised about the reliability of using Eq. (3) to calculate net benefits because we are integrating our estimated demand curve from values of $q = 0$ over the equilibrium output levels. But as shown in Fig. 1, the vast majority of our data points lie in a small neighborhood around $q = 0$, implying that our coefficients are most reliable at small levels of traffic. In addition, it is highly unlikely that demand could be infinite around $q = 0$ because urban travelers do not place a large value on rail compared with their next best alternative (Winston and Shirley [3]). According to the 2001 *National Household Travel Survey*, the median annual incomes of rail users exceed \$50,000, which indicates it would be misleading to suggest that a perceptible fraction may be captive to rail because they cannot afford a car.²³ As traffic levels significantly expand, calculation errors have a much smaller effect on the estimated triangular region of surplus because the area that is added falls with the square of q .

In Table 5, we present estimates of the net benefits for each system in the year 2000.²⁴ Although New York City's transit system and some of the large systems do provide substantial

²³ Median annual incomes of rail transit users may be considerably higher than \$50,000 because the highest income bracket reported in the *National Household Travel Survey* is "greater than \$50,000."

²⁴ The estimates measure the absolute rather than the marginal benefits to travelers provided by rail. Thus, we are overestimating rail's benefits because if rail transit were eliminated, some users would shift to other modes of transportation and their loss in consumer surplus would only be the difference between the consumer surplus from rail and the consumer surplus from the next best modal alternative. On the other hand, we are also underestimating rail's benefits to some extent because some travelers who never use rail may nonetheless value it as an (unexercised) option. Given rail's very small market share (indicating its lack of competitiveness with auto) and its low likelihood of being a viable alternative for the

Table 5

Social net benefits of transit, 2000* (figures in parentheses include an exhaustive public spending cost of 10.2%)

City (agency)	Consumer surplus	Transit agency deficit	Net benefits	Congestion savings to road users	Social net benefits
New York (NYC Transit)	850	2500 (2750)	-1650 (-1900)	1195.7	-454.3 (-704.3)
Washington, DC** (METRO)	281	657 (724)	-376 (-443)	181.3	-194.7 (-261.7)
SF Bay Area (BART)	542.5	624 (688)	-81.5 (-145)	181.3	99.8 (36.3)
Chicago (CTA)	391	644 (710)	-253 (-319)	272.8	19.8 (-46.2)
Boston (MBTA)	256	701 (772)	-445 (-517)	64.4	-380.6 (-452.6)
Atlanta (MARTA)	120	424 (467)	-304 (-347)	45.5	-258.5 (-301.5)
Philadelphia (SEPTA)	54	365 (402)	-311 (-348)	77.0	-234 (-271)
Northern New Jersey (PATH)	62	141 (155)	-79 (-93.4)	6.3	-72.7 (-87.1)
Los Angeles Metro	17	477 (526)	-460 (-509)	383.8	-76.2 (-125.2)
San Diego Trolley	6.8	47.5 (52.3)	-40.7 (-45.6)	16.4	-24.3 (-29.2)
Portland, OR (TriMet)	4	213 (235)	-209 (-231)	9.1	-199.9 (-221.9)
Baltimore (MTA Maryland)	6	198 (219)	-192 (-212)	14.9	-177.1 (-197.1)
Miami-Dade Transit	Negligible***	141 (156)	-144 (-158)	16.9	-127.1 (-141.1)
San Francisco (Municipal Railway)	3	276 (305)	-273 (-301)	51.5	-221.5 (-249.5)
St. Louis (Bi-State Dev. Agency)	Negligible***	160 (176)	-159 (-176)	5.0	-154 (-171)
Southern New Jersey (PATCO)	2.57	8.77 (9.66)	-6.2 (-7.14)	Negligible	-6.2 (-7.14)
Cleveland (GCRT)	2	115 (127)	-113 (-125)	7.5	-105.5 (-117.5)
Dallas (DART)	13	443 (488)	-430 (-475)	18.2	-411.8 (-456.8)
Sacramento RT	Negligible***	96.7 (107)	-100 (-110)	4.0	-96 (-106)
San Jose (Santa Clara Co. Tr.)	Negligible***	202 (223)	-201 (-222)	11.5	-189.5 (-210.5)
Pittsburgh (PA Allegheny Co.)	Negligible***	127 (140)	-126 (-139)	3.6	-122.4 (-135.4)
Denver (RTD)	Negligible***	259 (277)	-260 (-285)	5.6	-254.4 (-279.4)
Staten Island (SIRT)	3.4	22.5 (24.8)	-19.1 (-21.3)	Negligible	-19.1 (-21.3)
Buffalo (Niagara Frontier)	Negligible***	51.2 (56.5)	-51.5 (-56.7)	Negligible	-51.5 (-56.7)
Newark (NJTransit)	Negligible***	55.1 (60.8)	-54.2 (-59.8)	1.2	-53 (-58.6)
Total			-6338 (-6984)	2573	-3842 (-4496)

* All figures are in millions of 2000 dollars.

** The actual congestion savings to drivers for Washington, DC, are unavailable, so we use the estimated savings for a comparable metropolitan area with a similar transit system (SF Bay Area).

*** Negligible consumer surplus (i.e., estimated consumer surplus is less than the error term).

consumer surplus, their benefits pale in comparison with the massive deficits that these agencies face. In the final year of our sample, no transit agency provided positive net benefits. As shown in Fig. 2, the welfare performance of the largest systems has deteriorated over time. As of 1994, these systems combined to produce a small welfare loss.²⁵ But in 1995 rail transit budget deficits ballooned because vehicle miles on existing systems increased, while traffic and revenues both fell. Transit agencies found new sources of funding for the deficits by wresting bridge and tunnel tolls away from local governments, and either raising fares or municipal taxes. This auxiliary source of revenue exhibited a sharp, fourfold increase from 1993 to 1995.²⁶

Transit deficits are funded largely through taxation, which generates a deadweight loss that has been estimated by Allgood and Snow [17] to range from 7.8 to 12.6 cents per marginal dollar of tax revenue. We use the mean of the estimates, 10.2 cents, to compute these additional costs. As shown in Table 5, rail systems are even less socially desirable when we account for the cost of raising public funds. By 2000, the losses attributable to all rail transit systems amounted to roughly \$7 billion, broken down as \$1.9 billion for New York, \$2.0 billion for other large systems, \$1.4 billion for medium systems, and \$1.7 billion for small systems.

It could be argued that some rail transit costs amount to transfers to labor in the form of rents—that is, above market wages. Accordingly, we should include benefits to labor that would be subtracted from the welfare losses. However, elimination of rail service would lead to a concomitant expansion in bus service and jobs for rail transit workers at their former pay because unions typically represent all transit workers and state-level regulations mandate changes in pay-scales to be the same for rail and bus occupations. Hence, rail workers are likely to maintain a sizable fraction of their rents.

Externalities

Rail transit systems compete with automobiles and light trucks for passengers. By taking some travelers off the road, rail reduces congestion and enhances its social desirability. Winston and Langer [18] estimated econometric models of the determinants of congestion costs for motorists, truckers, and shippers based on their value of travel time delays and the difference between free flow speeds and actual speeds within and between major urban areas. The authors estimated the impact of a city's rail transit system on road users' congestion costs by including rail's total directional route miles in the specification and calculating the difference between congestion costs with and without a given system. We present the estimates for the systems here in the last two columns of Table 5. Congestion cost savings amount to an additional \$2.5 billion in social benefits and reduce total losses to \$4.5 billion, but social net benefits are still negative for all systems except BART.

In most urban areas it is not surprising that rail transit does not attract enough auto travelers to be socially desirable, but readers may question that this finding also applies to New York City. We therefore provide a check on our estimate of the congested-related benefits from the NYC rail transit system by calculating the congestion costs that would result if the system were closed.

many travelers who do not live and work sufficiently close to rail stations, our measure is most certainly overestimating rail benefits more than it is underestimating them. Thus, rail's actual benefits are likely to be smaller than estimated here.

²⁵ Using a different methodology than used here, Winston and Shirley [3] found that rail transit's net benefits were roughly zero in 1990. This is consistent with the findings in Fig. 2 for the early 1990s.

²⁶ *American Public Transportation Association*, 2000, Statistics Table 66. Note the deficits did not increase because of a change in accounting for capital expenditures.

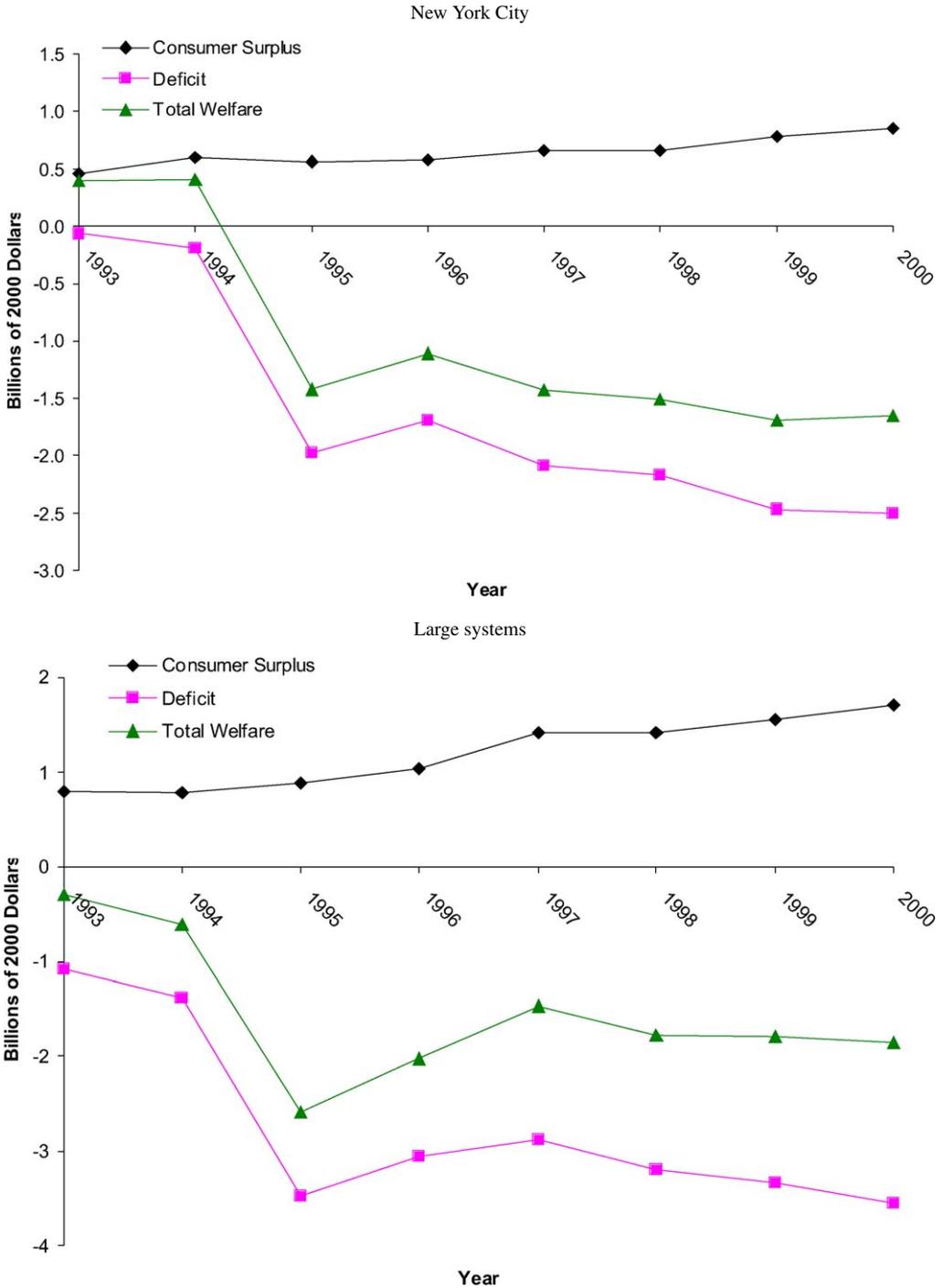


Fig. 2. New York City and large transit system welfare. (Large transit systems include Metro (Washington, DC), BART (SF Bay Area), CTA (Chicago), MBTA (Boston), MARTA (Atlanta), SEPTA (Philadelphia) and PATH (Northern New Jersey).)

We use a congestion cost function, apparently first proposed by William Vickery and used in the standard Urban Transportation Planning Program package provided by the U.S. Department of Transportation to state and local agencies, which specifies delays as proportional to the fourth power of the traffic volume-capacity ratio. Based on this function, we find that if all NYC subway users were forced to travel by auto, congestion costs would increase \$1.3 billion (in 2000 dollars). This figure compares quite favorably as an upper bound to our estimate of \$1.2 billion in congestion cost savings provided by the NYC subways. We note that the cost estimate is an upper bound because the New York City system is unique among rail transit systems in its large share, 57 percent, of non-commuters who comprise its ridership compared with the roughly 20 percent share of non-commuters on other rail systems in the country.²⁷ A greater share of non-commuters who are displaced reduces the costs of highway congestion because these users have the flexibility to avoid peak-period travel (Winston and Langer [18]).

In theory, rail transit could provide additional external benefits besides reducing roadway congestion, but empirical evidence of these benefits is weak. First, it has been claimed that by attracting auto users, rail reduces emissions. But given its low load factor, which includes a large share of users who keep older cars to get to suburban rail stations, its high consumption of electricity, whose generation produces pollution, and its consumption of smaller amounts of heavy petrochemicals, such as kerosene and bunker fuel, a greater share of rail ridership has, at best, an ambiguous effect on the environment.²⁸ In addition, the construction and expansion of new and existing rail systems is very energy intensive. For instance, Tri-County Metropolitan Transit Agency claims that under the best case scenario, the proposed north light-rail line in Portland, Oregon would save the equivalent of 7875 gallons of gasoline per day. But the agency also calculates the energy cost of building the line to be 32 million gallons of gasoline. Thus, even using the most optimistic estimates and assuming no depreciation of the capital stock, it would take a minimum of 15 years to even begin to achieve energy savings—and concomitant reductions in emissions—from this rail line.

It has also been argued that rail transit improves the safety of urban travel by reducing traffic on the road. But motorists absorb (internalize) most of the cost of accidents through various types of insurance. And the net improvement in safety to those who switch to rail is small because travelers are still exposed to losses in property and bodily harm from transit accidents and serious crimes on trains and in stations.²⁹

Finally, it has been suggested that rail has contributed to commercial development. But case studies have yet to show that after their construction transit systems have had a significant effect on employment or land use close to stations and that such benefits greatly exceed the benefits from commercial development that would have occurred elsewhere in the absence of rail construction (Bollinger and Ihlanfeldt [20], Charles [21]).³⁰

²⁷ These figures were obtained from the New York City Metropolitan Transit Agency and the 2000 National Transit Database.

²⁸ Mannering and Winston [19] estimate a duration model of vehicle ownership and find that, all else constant, households who reside in large metropolitan areas, which tend to have rail transit systems, hold on to their cars longer than households who reside in smaller metropolitan areas.

²⁹ According to calculations from the *National Transportation Statistics Report*, Bureau of Transportation Statistics, 2000, urban fatalities per passenger mile are slightly *higher* for light rail than for automobile, and comparable for heavy rail and automobile. The small incidence of light rail accidents suggests caution in drawing conclusions from these data. In any case, it is far from clear that rail transit is safer than automobile travel.

³⁰ As noted in footnote 4, rail's effect on housing prices is an alternative (indirect) measure of its benefits to users.

Notwithstanding the economic efficiency considerations that are unfavorable to rail transit, supporters of these systems claim they are attractive on distributional grounds because they contribute to the mobility of low-income residents. But the median income of rail transit users exceeds the median income in the general population. In addition, rail transit systems have difficulty keeping up with and responding to changes in job growth; thus, they are unable to provide the poorest residents access to employment opportunities in outlying suburbs (Winston and Shirley [3]).

Optimization

Rail transit in its current form does not generate nearly enough benefits to consumers to offset its massive deficits. However, it is possible that improvements in transit networks, which have not been optimized, could close the gap between benefits and costs such that some systems are socially desirable. Using our demand and cost models, we simulate optimal networks to explore this possibility.

The net benefits produced by transit systems in short-run equilibrium are given by Eq. (3). In the long run, it is reasonable to assume that transit agencies can adjust track length, stations, and so on, which affects demand and costs (assuming the political impediments to structural network change can be overcome). We therefore optimize the net-benefits equation with respect to network variables to see if net benefits can improve.³¹ For example, we obtain the welfare maximizing level of track by differentiating net benefits with respect to total track length, assuming the ratio of stations to track and overall connectivity are held constant. These assumptions allow us to shrink or expand a transit network while keeping its “shape,” which is influenced by the physical and geographical characteristics of the city it serves. We find for all systems that the welfare maximizing level of track is zero (i.e., we obtain corner solutions for all systems). In other words, no optimal size for any transit system exists that enables it to generate enough consumer surplus and revenue to offset its short-run total cost. (Our finding would be even stronger if we performed the optimization using long-run instead of short-run total cost.)

Another approach to reducing the transit budget deficit would be to raise fares. Setting prices at marginal cost would produce a three-to fourfold increase in fares for the vast majority of systems in our sample—an increase that would result in such a massive attrition of ridership that realistic improvements in social welfare would be generated only by abandonment.³² Small increases in fares would have a modest impact on deficits. If neither optimal pricing nor optimal investment can enable rail systems to generate net benefits, then their social desirability is clearly in question.

5. Discussion and policy implications

What factors contribute to rail transit’s social undesirability? Rail’s budget balance is inherently strained by the high costs of building and maintaining a network to serve urban and

³¹ This analysis considers strategies to improve the net benefits of rail transit systems. We cannot evaluate social net benefits, which account for congestion cost savings to road users, because we do not know how these savings would vary with respect to changes in network variables. It is likely that network optimization would call for reductions in track and stations, which would undoubtedly reduce congestion cost savings to drivers. Hence, our estimate of current congestion savings to drivers serves as an upper bound.

³² Winston and Shirley [3] found that rail transit use would still fall sharply if fares were raised to marginal cost and if motorists were charged efficient (marginal cost) congestion tolls.

suburban travelers and by the inefficiencies associated with low load factors and excessive labor expenses. Rail is unable to generate revenues to cover these costs because it must offer low (subsidized) fares to compete with the convenience and flexibility of autos.

Since the 1970s, deficits expanded as rail costs rose while demand fell. Aging systems have incurred high costs to repair and maintain their systems. For example, the Washington Metro has spent nearly \$1 billion in recent years to improve system reliability and ease crowding with little to show for project expenditures.³³ The projected cost of new systems to the public and the federal government has often been underestimated (or understated) by transit promoters (Pickrell [22], Flyvbjerg, Holm, and Buhl [23]). The public finally rebelled against cost overruns as Los Angeles county voters in 1998 temporarily halted extensions of their rail system by denying the use of a county transit sales tax for additional subway projects. But this ballot measure did not prevent the use of other funds (federal and state) or apply to light rail. Recently, the New York Metropolitan Transportation Authority had a showdown with transit labor over escalating costs. After leading a strike, the union was able to negotiate a contract for transit workers that maintained most of their benefits.³⁴

The demand for rail has continued to shrink because transit networks are unable to keep up with changing land use and travel patterns that have decentralized residences and employment. Indeed, less than 10 percent of the nation's employment in metropolitan areas is located in old central business districts. Baum-Snow and Kahn [24] point out that in cities with rail systems that have not changed their networks, rail's share has declined as former patrons and jobs have moved beyond rail's catchment areas. Even in cities that have built and expanded their rail systems, persistently declining population densities in catchment areas have prevented rail from attracting even a modest (e.g., 2 percent) share of travelers. Bertaud [25] argues that Atlanta's Regional Transportation Plan to expand the existing transit network is not warranted because the current spatial structure of Atlanta is incompatible with a sizable transit market share.

Generally, rail cannot be relied on to expand its system in a timely fashion to attract a potentially large pool of riders when an opportunity exists. For instance, the Capital Center, located in a suburb outside of Washington, DC, was regularly used since the early 1970s to showcase popular entertainment and professional basketball and hockey. Metro service to the arena was likely to attract considerable ridership. After decades of planning and delay, the Metro did open a rail station in 2005 at the site of the Capital Center—which unfortunately had been demolished three years earlier.

In addition, major extensions have not attracted the expected ridership. For example, BART extended its system to the San Francisco airport in 2003. Prior to construction, BART projected roughly 18,000 daily boardings to the airport by the year 2010 and that the service, in conjunction with three other new stations on the airport extension, would be profitable. As of 2006, there are 7000 daily boardings, indicating that the 2010 projection is unlikely to be realized, the route is losing money, while BART is embroiled in a fight over funds.³⁵

Why do existing systems continue to expand and new systems get built despite rail's negative contribution to social welfare? Rail transit enjoys strong support from urban planners, who wish to discourage auto use, from suppliers of transit capital and labor, who receive economic

³³ Lindsey Layton, Jo Becker, Efforts to repair aging system compound metro's problems, *Washington Post* (June 5, 2005) A1.

³⁴ Steven Greenhouse, New York transit deal shows union's success on many fronts, *New York Times* (December 29, 2005).

³⁵ Rachel Gordon, BART ridership to airport fails to take off, *San Francisco Chronicle* (July 8, 2006).

rents, from civic boosters, who perceive that a rail system adds prestige to their city, and from city officials, who support investments in a transit system that serves the downtown core because it may help the downtown remain vibrant or keep it from decaying. Until recently, the public has rarely rebelled against the actual costs of new systems or system extensions.³⁶ In fact, opinion polls suggest that a majority of residents in a city tend to support rail transit regardless of whether they actually use it on a regular basis.³⁷ We speculate that the public may support rail transit because it overestimates rail's ability to mitigate automobile externalities and because it is "rationally ignorant"—that is, the costs of transit subsidies (relative to other subsidies in the U.S. economy) are too small to merit the attention of most residents in a metropolitan area. Facing little resistance from the public, transit advocates aggressively explore alternative avenues to fund a new system or extend an existing one. And once a system is built, it is difficult to partially or completely abandon it, so it continues to receive public support despite its rising welfare costs.³⁸

Rail transit also benefits from substantial congressional support. Title III of the 1982 Surface Transportation Assistance Act granted a "golden penny" to transit (i.e., transit received one cent of the five cent increase in gasoline taxes). Since then, federal transportation legislation passed every six years has set aside for transit 20 percent of all revenues from gasoline tax increases used to rehabilitate and build the nation's highways (Dunn [27]). Congress has therefore solidified rail's participation in a multimodal urban transportation system and ensured that its funding will be entwined with highway funding.³⁹ Recently, rail transit has gained another source of funds through capital earmarks, which members use to benefit constituents in their districts. In fiscal year 2004, congressional earmarks for new construction and expansion of fixed guideway capital alone exceeded \$1.5 billion (calculated from the *Transportation, Treasury, and Independent Agencies Appropriations Bill Earmarks*, 2004).

Could any system be transformed to have a positive effect on social welfare? We are unable to find ways to significantly raise the net benefits of the nation's transit systems given their current operations. However, recently privatized rail transit systems in foreign cities, notably Tokyo and Hong Kong, have been able to eliminate deficits by reducing labor and capital costs and by introducing more comfortable cars and remote payment mechanisms, among other innovations, that have reduced operating costs and expanded ridership.

³⁶ As it turns out, the 1998 Los Angeles ballot measure was only a temporary setback for rail transit in LA. In the time since the measure passed a new light rail line has been put into operation, another one is being constructed, and the newly elected mayor, Antonio Villaraigosa, has pledged to extend the Red Line subway along Wilshire Boulevard to the Pacific Ocean.

³⁷ For example, Steven Ginsberg, Commuters like metro more than they use it, *Washington Post* (March 5, 2005) A1, reports that in a recent *Washington Post* poll, only 9 percent of Washingtonians said they regularly use the Metro to get to and from work, while two-thirds said that public transportation is not an option for them. However, 58 percent of respondents said they would support more funding for Metro, even if that means higher taxes.

³⁸ Baseball stadiums also enjoy strong public support despite their dubious welfare properties. Indeed, even the most successful baseball stadiums, such as Baltimore's Camden Yards, do not generate enough revenues to justify their substantial costs. In fact, every independent economic analysis of new stadium construction has failed to find measurable positive effects on output or employment, and some analyses have even found negative effects (Zimbalist [26]). Nonetheless, cities continue to build new stadiums with public funds.

³⁹ States have also tied rail funding to highway funding. For example, environmental regulators in Massachusetts ruled that Boston's "Big Dig" project to depress the central artery could not proceed unless the state expanded Boston's rail transit system. On the other hand, given its tiny share of passengers, it is doubtful that rail has had much effect on funding for highways. Thus, it is difficult to argue that in the absence of rail transit funding for highways would be measurably greater.

We therefore investigated which, if any, U.S. rail transit systems would become socially desirable assuming privatization reduced short-run total costs 20 percent—a plausible estimate based on U.S. and foreign experience with bus transit privatization (Winston and Shirley [3]). With the exception of BART, which already generates small net benefits, we found that such a cost reduction would result in only the New York City and Chicago systems producing positive net benefits.

We are not aware of any public officials who have endorsed complete privatization of rail transit. On the other hand, a few have encouraged bus transit agencies to contract with private companies in an effort to reduce costs. Private contracting would be a politically more feasible alternative to privatization, but it appears that at best it would enable only a few rail systems to be socially justified.

Because no policy option exists that would enhance the social desirability of most urban rail transit systems, policymakers only can be advised to limit the social costs of rail systems by curtailing their expansion. Unfortunately, transit systems have been able to evolve because their supporters have sold them as an antidote to the social costs associated with automobile travel, in spite of strong evidence to the contrary.⁴⁰ As long as rail transit continues to be erroneously viewed in this way by the public, it will continue to be an increasing drain on social welfare.

References

- [1] T. Litman, Evaluating rail transit criticism, Victoria Transport Policy Institute, October 2004.
- [2] J. Richmond, A whole-system approach to evaluating urban transit investments, *Transport Reviews* 21 (2001) 141–179.
- [3] C. Winston, C. Shirley, *Alternate route: Toward efficient urban transportation*, Brookings Institution, Washington, DC, 1998.
- [4] I. Savage, Management objectives and the causes of mass transit deficits, *Transportation Research* 38A (2004) 181–199.
- [5] I. Savage, A. Schupp, Evaluating transit subsidies in Chicago, *Journal of Public Transportation* 1 (1997) 93–117.
- [6] J. Kain, Cost-effective alternatives to Atlanta's rapid transit system, *Journal of Transport Economics and Policy* 31 (1997) 25–49.
- [7] P. Viton, On competition and product differentiation in urban transportation: The San Francisco bay area, *Bell Journal of Economics* 12 (1981) 362–379.
- [8] R. Diaz, Impacts of rail transit on property values, in: *American Public Transit Association Rapid Transit Conference Proceedings*, May 1999.
- [9] N. Baum-Snow, M. Kahn, The effects of new public projects to expand urban rail transit, *Journal of Public Economics* 77 (2000) 241–263.
- [10] P. Hagget, R. Chorley, *Network Analysis in Geography*, St. Martin's Press, New York, 1969.
- [11] R. Pozdena, L. Merewitz, Estimating cost functions for rail rapid transit properties, *Transportation Research* 12 (1978) 73–78.
- [12] P. Viton, On the economics of rapid-transit operations, *Transportation Research* 14A (1980) 247–253.
- [13] I. Savage, Scale economies in United States rail transit systems, *Transportation Research* 31A (1997) 459–473.
- [14] C. Winston, Conceptual developments in the economics of transportation: An interpretive survey, *Journal of Economic Literature* 23 (1985) 57–94.
- [15] K. Small, *Urban Transportation Economics*, Harwood Academic Publishers, Switzerland, 1992.

⁴⁰ Richmond [28] summarizes the results of interviews he conducted with influential policymakers and residents in Los Angeles, who were convinced that building a rail subway system in the Southland was strongly justified. It may be argued that the benefits from rail transit are likely to be greater as road congestion continues to increase. However, existing transit systems (e.g., Washington Metro and BART) have exhibited a limited capability of handling additional patronage without experiencing over-crowding and delays in boarding and exiting trains. In addition, the benefits that transit may generate from reducing growing congestion costs are likely to be offset by its rising deficits.

- [16] R. Voith, The long-run elasticity of demand for commuter rail transportation, *Journal of Urban Economics* 30 (1991) 360–372.
- [17] S. Allgood, A. Snow, The marginal cost of raising tax revenue and redistributing income, *Journal of Political Economy* 106 (1998) 1246–1273.
- [18] C. Winston, A. Langer, The effect of government spending on road users' congestion costs, *Journal of Urban Economics* (2006), in press.
- [19] F. Mannering, C. Winston, Brand loyalty and the decline of American automobile firms, *Brookings Papers on Economic Activity: Microeconomics* (1991) 67–114.
- [20] C. Bollinger, K. Ihlanfeldt, The impact of rapid rail transit on economic development: The case of Atlanta's MARTA, *Journal of Urban Economics* 42 (1997) 179–204.
- [21] J. Charles, The mythical world of transit oriented development, *Cascade Policy Institute Policy Perspective* no. 1019, October 2001.
- [22] D. Pickrell, *Urban rail transit projects: Forecast versus actual ridership and costs*, US Department of Transportation Washington, DC, 1990.
- [23] B. Flyvbjerg, M. Holm, S. Buhl, Underestimating costs in public works projects: Error or lie? *Journal of the American Planning Association* 68 (2002) 279–295.
- [24] N. Baum-Snow, M. Kahn, The effects of urban rail transit expansions: Evidence from sixteen cities, 1970 to 2000, *Brookings–Wharton Papers on Urban Affairs* (2005) 1–60.
- [25] A. Bertaud, Clearing the air in Atlanta: Transit and smart growth or conventional economics? *Journal of Urban Economics* 54 (2003) 379–400.
- [26] A. Zimbalist, *May the best team win: Baseball economics and public policy*, Brookings Institution, Washington, DC, 2003.
- [27] J. Dunn, *Driving forces: The automobile, its enemies, and the politics of mobility*, Brookings Institution, Washington, DC, 1998.
- [28] J. Richmond, The mythical conception of rail transit in Los Angeles, *Journal of Architectural and Planning Research* 15 (1998) 294–320.