



ELSEVIER

Transportation Research Part A 38 (2004) 181–199

TRANSPORTATION
RESEARCH
PART A

www.elsevier.com/locate/tra

Management objectives and the causes of mass transit deficits

Ian Savage *

*Department of Economics and the Transportation Center, Northwestern University, 2001 Sheridan Road,
Evanston, IL 60208-2600, USA*

Received 8 April 2002; received in revised form 1 November 2002; accepted 30 October 2003

Abstract

Based on data for the period from 1948 to 1997, exogenous decreases in demand and increases in costs are estimated to have reduced the annual profitability of the Chicago Transit Authority (CTA) by \$1 billion. Half of this decline was recouped by reductions in service, increased fares and increased productivity. Even more would have been recouped had the CTA not given away earlier productivity gains during the 1970s. This was during a period when subsidies were increasing rapidly, and seemingly without constraint. When faced with financial challenges, management has preferred to increase fares rather than reduce service levels to the detriment of social welfare.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Transit; Subsidies; Chicago; Management objectives

1. Introduction

Subsidies to urban transit systems increased rapidly during the 1970s. The causes have been hotly debated. Some hold that purely exogenous factors, such as increased automobile ownership, depopulation and deindustrialization of cities, and inflation-driven cost increases, were to blame. Others argued that the decline was exacerbated by endogenous decisions by managers who exercised poor cost control, made unwise service expansions, and hesitated to increase fares in an inflationary era. Pickrell (1985), using aggregate data, found that the latter predominated. However, he was unable to establish whether the decisions made by managers were “a *source* of increasing deficits, or a *use*—intended or otherwise—of the rapidly growing subsidies” (emphasis in original). This paper attempts to untangle the cause or effect conundrum by applying Pickrell’s

* Tel.: +1-847-491-8241; fax: +1-847-491-7001.

E-mail address: ipsavage@northwestern.edu (I. Savage).

methodology to a lengthy time series data set for one, large, transit agency. The data is for 50 years, and covers the period before, during, and after the run up in subsidies.

2. Previous literature

Pickrell (1985) concluded that the rise in aggregate operating deficits in the United States between 1970 and 1982 had been due to increased unit costs (representing 61% of the increased deficits), increased service provision (17%), reduced real fares (14%), and adverse exogenous demand conditions (8%). The rise in transit costs has spawned a large literature. The literature attempts to see whether transit has moved away from the most efficient methods of production. Changes in total costs are disaggregated into the effects of changes in output, changes in the characteristics of output, and changes in productivity of various factors of production (see Oum et al. (1992) for a theoretical survey, and Obeng and Sakano (2002), for a recent application to transit).

This paper, like Pickrell's, deals with costs in a less sophisticated way. However, it explicitly considers decisions by managers on the price and quantity of output. Theoretically, Nash (1978) and Glaister and Collings (1978) show that there are multiple combinations of fare and level of service that can satisfy a given subsidy constraint. In selecting their preferred combination, managers can be motivated to maximize social welfare, the numbers of passengers carried, or the amount of service provided. The preponderance of the empirical analyses to date concludes that transit agencies have typically opted to maximize level of service as opposed to social welfare (see Glaister (1987), for evidence from Britain and Australia). As a result, social welfare is lower than it might otherwise be.

3. Choice and relevance of the case study

In most cities, it is difficult to obtain a lengthy time series of comparable data. Typically, there have been mergers of neighboring companies, expansion of service into newly-developed suburbs, regionalization of finances, and privatization and contracting of service. Consequently there are changes in corporate structure and data inconsistency. To conduct a long time series analysis in cities such as New York, London or Sydney would essentially require comparing apples and oranges.

The Chicago Transit Authority (CTA) provides a rare exception. Its basic structure has changed little in more than 50 years. It has been publicly owned since 1947, was not permitted to expand geographically, has been untouched by privatization, and retains a statutory monopoly. While regional forms of transportation planning and finance did emerge during the 1970s, the CTA remains a distinct corporate and operating entity. It is also a very large operating agency, ranking fourth in ridership in North America. Currently, it provides service with a peak-vehicle requirement of 1630 buses and 990 railcars, operating on seven routes, to approximately four million people living in the City of Chicago and the older inner suburbs. Over the course of the past half century, there have been wild swings in both management policy and market conditions which provide a rich source of data for analysis.

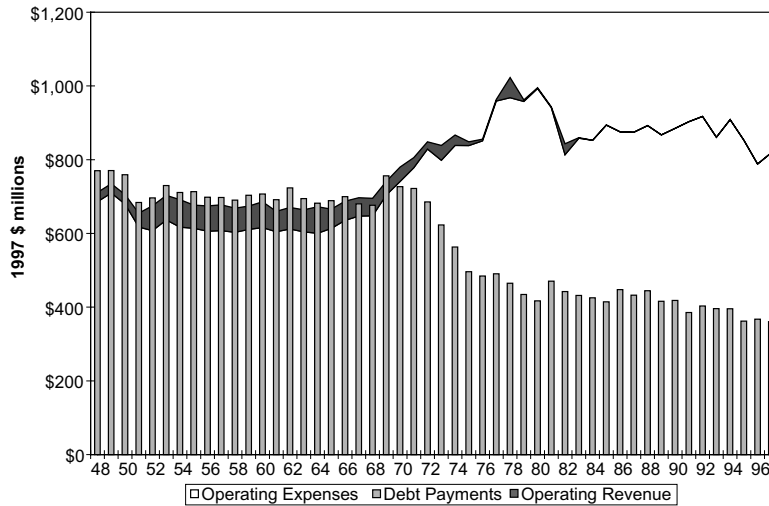


Fig. 1. Operating revenues and expenses in 1997 dollars.

But how generally applicable are the findings? Although, some of the specifics are unique to Chicago, the overall pattern is surprisingly similar to other older “traditional” cities in the northeastern United States and Canada, and comparable cities on other continents. For example, the graph of the constant-dollar operating revenues and costs for the CTA, shown in Fig. 1, is almost identical to that for the entire US transit industry (Winston and Shirley, 1998, Figs. 1 and 2). Starting in the mid 1960s and lasting into the early 1980s, subsidies increased considerably as revenue fell and costs increased. Evidence from Europe, Canada, Australia and New Zealand is broadly similar (Bly and Oldfield, 1985, see especially Fig. 1). Costs and subsidies increased considerably over the same time period. In most countries revenue also fell in real terms (although this was not the case in Britain and a few other countries, where fares were increased, but even there revenue did not keep pace with the increased costs).

4. Analytical model

The model is based on the identity that operating profit/deficit is equal to total farebox revenue minus total operating cost:

$$\pi = P(q_R(P, M_R, X_i) + q_B(P, M_B, X_i)) - (C_{RX} + C_{RN})M_R - (C_{BX} + C_{BN})M_B \tag{1}$$

where

- π profit (positive) or deficit (negative)
- $q_R(\cdot)$ a demand function explaining rail passenger trips
- $q_B(\cdot)$ a demand function explaining bus passenger trips (the generic term “bus” is used to indicate service on the surface streets which at various times has been provided by a combination of streetcars, trolley buses and motor buses)
- P common price (fare) on both modes (the CTA has always had a single flat fare)

M_R	railcar miles
M_B	bus miles
X_i	a vector of exogenous demand factors
C_{RX}	exogenously determined cost per railcar mile
C_{RN}	endogenously determined cost per railcar mile
C_{BX}	exogenously determined cost per bus mile
C_{BN}	endogenously determined cost per bus mile

Exogenously determined costs are defined as the unit costs that the CTA inherited on its formation, adjusted to reflect nationwide changes in the unit costs of inputs. For example, real wages of all workers have increased over the past 50 years, and the CTA would have to respond to this in order to attract employees. Of course, actual unit costs may be more than or less than the exogenously determined costs. The difference between actual and exogenously determined costs will be called endogenous costs. For example, if the introduction of productivity-enhancing technology or work practices reduces actual unit costs below exogenously determined unit costs, then the value of the variables representing endogenous-determined costs (C_{RN} and C_{BN}) will be negative. Alternatively, if labor is used less efficiently, or paid wages that are in excess of those in comparable occupations, actual unit costs will exceed the exogenously determined ones and C_{RN} and C_{BN} will be positive.

Our interest is not in Eq. (1) *per se*, but rather in how it changed from year to year. Using the expression Δ to represent the change in the value of a variable from one year to the next (for example, subtracting 1949 from 1950), the “first difference” of the profit function is:

$$\begin{aligned} \Delta\pi &= q_R(\cdot)\Delta P + P\Delta q_R(\cdot) + \Delta P\Delta q_R(\cdot) + q_B(\cdot)\Delta P + P\Delta q_B(\cdot) + \Delta P\Delta q_B(\cdot) \\ &\quad - M_R(\Delta C_{RX} + \Delta C_{RN}) - (C_{RX} + C_{RN})\Delta M_R - (\Delta C_{RX} + \Delta C_{RN})\Delta M_R \\ &\quad - M_B(\Delta C_{BX} + \Delta C_{BN}) - (C_{BX} + C_{BN})\Delta M_B - (\Delta C_{BX} + \Delta C_{BN})\Delta M_B \end{aligned} \quad (2)$$

Collecting terms on the right-hand side and disaggregating the change in demand into its constituent parts produces:

$$\begin{aligned} \Delta\pi &= q_R(\cdot)\Delta P + (P + \Delta P) \left[\partial q_R(\cdot)/\partial P \Delta P + \partial q_R(\cdot)/\partial M_R \Delta M_R + \sum_i \partial q_R(\cdot)/\partial X_i \Delta X_i \right] \\ &\quad + q_B(\cdot)\Delta P + (P + \Delta P) \left[q_B(\cdot)/\partial P \Delta P + \partial q_B(\cdot)/\partial M_B \Delta M_B + \sum_i \partial q_B(\cdot)/\partial X_i \Delta X_i \right] \\ &\quad - M_R(\Delta C_{RX} + \Delta C_{RN}) - (C_{RX} + C_{RN})\Delta M_R - (\Delta C_{RX} + \Delta C_{RN})\Delta M_R \\ &\quad - M_B(\Delta C_{BX} + \Delta C_{BN}) - (C_{BX} + C_{BN})\Delta M_B - (\Delta C_{BX} + \Delta C_{BN})\Delta M_B \end{aligned} \quad (3)$$

This equation can be made more intuitively appealing by substituting to remove the differential terms. For example, price elasticity for the rail system (ε_{PR}) is defined as:

$$\varepsilon_{PR} = (\partial q_R(\cdot)/\partial P) \cdot (P/Q_R)$$

where Q_R is the number of rail passenger trips. By manipulation:

$$\partial q_R(\cdot)/\partial P = \varepsilon_{PR} Q_R / P$$

which can be substituted into Eq. (3). There will be similar definitions for the other rail demand variables, and for the bus mode. Consequently Eq. (3) can be rewritten as:

$$\begin{aligned}
\Delta\pi = & Q_R\Delta P + (P + \Delta P) \left[\varepsilon_{PR}Q_R\Delta P/P + \varepsilon_{MR}Q_R\Delta M_R/M_R + \sum_i \varepsilon_{RX_i}Q_R\Delta X_i/X_i \right] + Q_B\Delta P \\
& + (P + \Delta P) \left[\varepsilon_{PB}Q_B\Delta P/P + \varepsilon_{MB}Q_B\Delta M_B/M_B + \sum_i \varepsilon_{BX_i}Q_B\Delta X_i/X_i \right] - M_R(\Delta C_{RX} + \Delta C_{RN}) \\
& - (C_{RX} + C_{RN})\Delta M_R - (\Delta C_{RX} + \Delta C_{RN})\Delta M_R - M_B(\Delta C_{BX} + \Delta C_{BN}) \\
& - (C_{BX} + C_{BN})\Delta M_B - (\Delta C_{BX} + \Delta C_{BN})\Delta M_B \tag{4}
\end{aligned}$$

Terms can be collected in such a way that the change in profitability can be decomposed into six effects. The first two are exogenous, and the other four are under the control of the CTA:

- A. changes in the exogenous unit cost of existing service ($-M_R\Delta C_{RX} - M_B\Delta C_{BX}$),
- B. revenue change from various exogenous demand factors ($\sum_i (P + \Delta P) \cdot \Delta X_i / X_i [\varepsilon_{RX_i}Q_R + \varepsilon_{BX_i}Q_B]$),
- C. changes in the endogenous unit cost of existing service ($-M_R\Delta C_{RN} - M_B\Delta C_{BN}$),
- D. changes in rail service, composed of the net effect of the increased mileage at the original unit costs, the effect of increased unit cost, the revenue consequences of generated ridership at existing prices, and the effect on generated revenue of price changes ($\Delta M_R [P \cdot \varepsilon_{MR}Q_R/M_R - (C_{RX} + C_{RN})] + \Delta M_R [\Delta P \cdot \varepsilon_{MR}Q_R/M_R - (\Delta C_{RX} + \Delta C_{RN})]$),
- E. changes in bus service, calculated in a similar fashion ($\Delta M_B [P \cdot \varepsilon_{MB}Q_B/M_B - (C_{BX} + C_{BN})] + \Delta M_B [\Delta P \cdot \varepsilon_{MB}Q_B/M_B - (\Delta C_{BX} + \Delta C_{BN})]$),
- F. changes in fares, comprising the revenue effects from changing price to existing riders and from generated/discouraged riders ($\Delta P(Q_R + Q_B) + (P + \Delta P) \cdot \Delta P/P [\varepsilon_{PR}Q_R + \varepsilon_{PB}Q_B]$).

While Eq. (4) is developed as an accounting identity, in practice it is unlikely to fully explain the change in profitability. There will be a residual, unexplained, amount when the demand variables fail to capture all of the changes in demand. There are many sources for error. Elasticities are econometrically estimated with inherent standard errors. The original data may contain measurement errors, and there will always be random demand shocks. Some of these shocks may be exogenous (a snow storm, or the soccer World Cup) while others (such as a strike) may be endogenous. The success of a decomposition will be judged by the size of residual.

5. Data

The analysis uses annual data for the 50-year period, from 1948, the first full-year of operation, to 1997. Published data are available on operating profit/loss, price (calculated as farebox revenue divided by total passenger trips), railcar and bus miles in revenue service, bus and rail unlinked passenger trips, and actual operating cost per railcar and bus mile. All dollar amounts are adjusted to 1997 prices using the consumer price index. Appendix A contains additional notes on sources and specific data issues and problems.

Exogenous costs per bus mile and per railcar mile were calculated by taking modal unit costs in 1948, and forecasting them forward using an index of exogenous factor input prices. The index was composed of two parts. The first part is an index of real employee expense per labor hour for the entire economy (see Gordon, 1995). This index was applied to 90% of the 1948 unit costs.

Direct labor expenses represent just less than 80% of operating costs, and the companies that supply other goods and services that the CTA purchases will be subject to similar wage pressures. The second part of the index, applied to the remaining 10% of the 1948 costs, is the real value of the fuel and power component of the national producer price index. Endogenous unit costs are the difference between the calculated exogenous unit costs and the observed actual costs.

The split between endogenous and exogenous cost changes will be sensitive to the choice of indices. Ideally, one would wish to decompose costs into more than two parts, and to apply indices that might be better tailored to the Chicago region as opposed to national trends. Despite this limitation, the calculated endogenous costs do appear to accord with observed events such as the introduction of new technology and changes in labor contracts.

The revenue side of the model includes variables (the X_i variables) that represent the many changes in society that have been unfavorable to transit. There has been a suburbanization of residential location. The population of the City of Chicago fell from 3.6m to 2.7m. Jobs, even for city residents, have also moved to the suburbs. In 1950, 87% of the jobs in Cook County, which covers Chicago and the inner ring of suburbs, were located in the City. Now that proportion is only 49%. Reverse commuting to low-density areas poorly served by public transit is common. City dwellers have also become more motorized, with auto ownership (as measured by the State of Illinois, Department of Motor Vehicles) rising from 170 per thousand population in 1948 to 400 per thousand. All of these changes are highly collinear to each other. In this analysis, the ratio of jobs in Chicago to total jobs in Cook County will be used as the primary variable, and should be interpreted as representing the whole host of societal changes and not just the suburbanization of jobs.

Two other exogenous demand variables will be used. Transit ridership is largely composed of the journey to work, and increased unemployment during recessions temporarily reduces travel. A “recession” dummy variable equals one for those years in which real Gross National Product per head of population in the United States fell. Data were obtained from the annual *Economic Report of the President*. By this definition, recessions occurred in 1949, 1954, 1957–1958, 1970, 1974–1975, 1979–1982 and 1990–1991. Finally, Chicago and other US cities witnessed a wartime transit boom (Kain, 1999, p. 363). Transit use in Chicago was 20% higher in 1946 than it was in 1941. Much of this additional traffic had dissipated by 1950 as gasoline rationing came to an end, and automobile factories reverted to civilian production. A spline dummy variable with the value of two in 1948, one in 1949, and zero thereafter is used to capture this “post-war” effect.

Demand elasticities appear in the underlying demand functions that are part of Eq. (1), and also in Eq. (4). Analytically we are interested in obtaining a good fit for the demand equations and minimizing the residual, or error, term when actual data are used in Eq. (4). Econometric efficiency can be improved by jointly estimating the demand functions and Eq. (4) using Zellner’s seemingly unrelated equations model. The estimation would have cross-equation constraints to equate estimated demand elasticities, and the variables in Eq. (4) that do not contain elasticities would have their coefficients constrained to equal +1 or –1 as appropriate.

Such an estimation is highly desirable for uni-modal transit systems. However, problems emerge for multi-modal systems because of very high collinearity between many (if not all) parts of the vector of variables explaining the demand for the rail and bus system in Eq. (4). Variables in the bus and rail parts of the equation share the same $\Delta P/P$ and $\Delta X_i/X_i$ components, and only differ because they are multiplied by the ridership on each mode, which are themselves highly

Table 1

Regression on annual modal passenger trips (first difference equation, continuous variables in logarithms, *t*-statistics in parentheses)

	Rail system	Bus system
Price	−0.256 (3.77)	−0.457 (6.17)
Service level	0.334 (3.00)	0.556 (3.17)
Recession dummy	−0.006 (0.70)	−0.012 (1.27)
Ratio of Chicago/Cook County employment	0.566 (1.69)	1.097 (2.72)
Post-war spline dummy	0.068 (2.91)	0.036 (1.37)
Adjusted R^2	0.51	0.65
Observations	49	49

collinear. Data from Chicago resulted in correlations of greater than 0.9. Moreover, the structure of the dummy variables led to perfect collinearity. Estimation of Eq. (4) would require dropping or combining variables which would remove the attractive cross-equation feature of joint estimation.

Consequently, for multi-modal agencies, elasticities have to be estimated solely from the modal demand functions (provided, of course, that disaggregated data are available). Ridership is regressed on price, service level and the three exogenous demand variables. A first difference equation was specified for all variables (including the dummies). The continuous variables are specified in logarithms, which will produce constant elasticities. An assumption is made that the marginal cost (and hence supply) function with respect to ridership is constant and approximately zero. Hence we are uniquely identifying the demand function.

A good fit, and attractive results were obtained because, at various times, real fares and service levels have both increased and decreased by large amounts, and the associated changes in demand were generally consistent with the usual assumptions made by economists. The results, shown in Table 1, indicate a price elasticity of −0.256 for the rail system and −0.457 for the bus system. This is in line with the extensive literature on the subject (Goodwin, 1992). The rail system's ridership is much more inelastic as it serves longer trips (averaging 6 miles) on radial routes to the CBD, whereas the bus system has much shorter trips (averaging 2 miles) in the neighborhoods where walking and the automobile are more effective competitors. Estimated service level elasticities are 0.334 for the rail system and 0.556 for the bus system, again consistent with the mainstream literature (Webster and Bly, 1980), and showing that rail ridership is more inelastic than bus ridership. Exogenous societal change, represented by the suburbanization of jobs variable, is particularly strong, especially for the bus system.

6. Overall results

In 1948, the CTA collected \$770m (at 1997 prices) from the farebox and other commercial forms of revenue such as advertising and charters. Operating costs were \$686m. The \$84m operating surplus was used to repay bondholders, provide for capital expenditures, and pay the City a “municipal compensation” fee for the use of the streets. By 1997, farebox and other commercial revenues had declined to \$360m, while operating costs had increased to \$819m. The

Table 2
Decomposition of the change in annual profitability

Category	Change
Exogenous change in unit cost of existing service	–\$575m
Revenue change from exogenous demand factors	–\$399m
Total exogenous changes	–\$974m
Endogenous change in unit cost of existing service	+\$91m
Changes in rail service	–\$8m
Changes in bus service	+\$182m
Changes in fares	+\$192m
Total endogenous changes	+\$457m
Unexplained changes in demand	–\$25m
Overall change in profitability	–\$542m

\$458m operating loss is funded by a local sales tax levy, grants from the State of Illinois, and small City and federal contributions (the capital needs of the agency have long since been funded by federal and local grants rather than by internally-financed bond issues). The decomposition of the \$542m decline in annual profitability into its six constituent parts is shown in Table 2.

The largest component is a \$575m decline caused by exogenous cost factors. Driven primarily by increases in economy-wide real wages, the CTA's cost base would have been expected to double over the 50 years. The company also suffered a \$399m loss caused by identified exogenous demand factors. Combined, the CTA suffered an annual billion dollar challenge to its profitability from factors outside of management control. Partly counteracting these negative effects were decisions to reduce bus service and raise real fares leading to positive financial effects of \$182m and \$192m respectively. The CTA was also able to offset \$91m of the unit cost increases by endogenous actions, such as changes in productive and technology.

The residual revenue reduction of \$25m is unidentified as to cause, and will probably have an endogenous as well as an exogenous component. The residual is relatively small. The correlation between actual and predicted changes in profitability is 0.94, implying that Eq. (4) can be thought of as having a r^2 of 0.88.

7. Disaggregated analysis by time period

While the overall results are interesting, the richness of the data is revealed when the 50-year period is split into smaller segments. As already discussed, and as illustrated in Fig. 1, the run up in subsidies mainly occurred between 1965 and 1980. Prior to 1965 and after 1980, both revenues and costs appear to be relatively stable. However, total revenue and cost figures mask a much more complex story. An indication of this complexity is given in Fig. 2 which shows some key operating statistics as indices with their value in 1948 equal to 100.

There have been wild fluctuations on these indices. Demand, as measured by total passenger trips, fell considerably in the decade after the Second World War, and has suffered a continual decline ever since. It is now only a quarter of what it was in 1948. The provision of bus service has

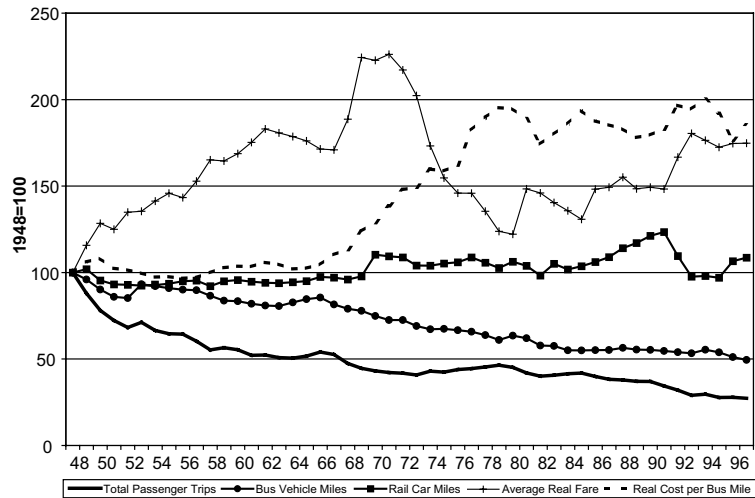


Fig. 2. Indices of CTA key indicators with 1948 = 100.

also gradually been decreased, but less than proportionately. Rail service has been expanded with new line construction at the end of the 1960s, in the early 1980s and mid 1990s. Real fares were raised considerably during the late 1940s and 1950s. There were dramatic increases in the late 1960s, but these were more than matched by reductions in the 1970s when monetary fares were kept almost constant despite high inflation. Since 1980 real fares have increased, but remain below those of the late 1960s. Unit costs (shown here for the bus system) remained constant throughout the 1950s and early 1960s, and then almost doubled between 1965 and 1980. In 1980 the increases were arrested and have since remained relatively stable.

This suggests that a much finer subdivision of the 50 years is necessary. Six periods have been chosen (1948–1957, 1958–1964, 1965–1970, 1971–1980, 1981–1990, 1991–1997) to coincide with dramatic changes in CTA policy. The disaggregation of changes in profitability within each of these periods is shown in Table 3. The remainder of the paper is concerned with analyzing this table and understanding and formalizing the management strategies that may have produced these results.

7.1. Responding to exogenous challenges (1948–1957)

Fig. 1 might suggest that 1948–1957 was a very stable period with total revenues and costs changing little. However, this disguises considerable challenges. Nationally, real wages increased by 30%, forcing the CTA to increase wages in response. Unit costs increased by almost 25%, and would have led to a decline in profitability of \$171m, were it not for an extraordinary increase in productivity. Much of the \$184m endogenous cost reduction came from the introduction of new technology into a cash-starved and physically outmoded system. Two-person operated streetcars were replaced by one-person operated trolley buses and motor buses, and new rail rolling stock eliminated the need to have a staff member on each car to operate the doors. Train crews were reduced from up to eight to a basic two-person crew. Many lightly used transit stations and little-used branch lines were closed.

Table 3

Decomposition of the change in annual profitability by time period (millions of 1997 dollars)

	1948–1957	1958–1964	1965–1970	1971–1980	1981–1990	1991–1997	Overall
Exogenous changes in unit cost of existing service	-171	-96	-140	-138	3	-34	-575
Post-war effect	-70	na	na	na	na	na	-70
Chicago/Cook employment	-68	-65	-56	-63	-43	-33	-328
Recessions	-8	7	-8	3	0	4	-1
Endogenous changes in unit cost of existing service	184	65	-16	-218	76	0	91
Changes in rail service	7	1	-25	8	-43	43	-8
Changes in bus service	13	11	18	54	50	37	182
Changes in fares	172	62	83	-203	43	35	192
Unexplained demand effects	-54	7	48	-4	21	-43	-25
Total profit change	6	-8	-96	-562	109	9	-542

(1) The Chicago/Cook County employment variable should be interpreted as including the effects of City depopulation and changes in automobile ownership.

(2) Data are rounded, so columns and rows may not add up exactly.

Exogenous challenges were as acute on the demand side. The wartime bulge in transit use dissipated, televisions became widespread, workplaces moved from six-day to five-day weeks, and the automobile became a more effective competitor after the City started to construct large parking garages downtown in 1952. If one includes several large negative residual demand effects in the early 1950s, exogenous demand changes reduced profitability by almost \$200m. Indeed, almost half of the negative exogenous demand reduction over the history of the CTA occurred in these first 10 years. To counter this decline, real fares increased by 50%. Revenue was boosted by \$172m, which almost compensated for the revenue lost to exogenous changes. Despite a loss of two-fifths of the ridership, the quantity of service provided fell relatively little, by only 5% on the rail system and 10% on the surface system, producing a minimal \$20m increase in profits.

The behavior of the CTA in response to these exogenous shocks would seem to accord with Nash's (1978) model of output maximization subject to an exogenous budget constraint. In formal terms, the situation could be characterized by a model with a demand function of:

$$Q_{R+B} = q(P, M_R, M_B, X_i)$$

and a total cost function:

$$TC = c(M_R, M_B, W, T)$$

where W is a vector of prices of inputs such as labor, fuel and equipment, and T is technology. The maximization problem is postulated to be:

$$\text{Max } \xi_1 = (M_R, M_B) + \lambda_1 [P^* q(P, M_R, M_B, X_i) - c(M_R, M_B, W, T) - \pi] \quad (5)$$

where the π represents the budget constraint. Analytically this represents a surplus if π is positive, or a subsidy if π is negative. Typically economists assume that price, service levels and technology (P , M_R , M_B and T) are endogenously determined, and the budget constraint, input prices and demand effects (π , W and X_i) are exogenously determined.

The budget constraint was certainly exogenous and fixed. The CTA was formed as a municipally owned yet commercial company. It had to enter into agreements with commercial banks to issue bonds in order to purchase bankrupt streetcar and elevated rail companies in October 1947 (these companies had been operated by court-appointed receivers since the early 1930s), and a small, profitable, motor bus company in 1952. The trust agreements with the banks required the CTA to pay approximately \$50m a year, in 1997 dollars, to service the bonds, and to set aside up to 8% of revenue (or about \$55m) in a depreciation trust fund to provide for capital needs including repayment of financial instruments issued for the purchase of rolling stock. Between 1948 and 1957 the CTA met this requirement almost exactly.

Wages appear to be exogenously determined as changes in real earnings at the CTA mirror national trends. The first order conditions of Eq. (5) have certain relevant predictions (see Nash, 1978, for details). First, the CTA will act in a unit cost minimizing fashion. The logic is that more output can be produced when unit costs are low. Second, cost reducing or revenue enhancing technology will be deployed to the extent commercially possible. Third, maximization of output within the budget constraint will result in increased fares. Indeed one of the first order conditions suggests that fares will be increased until marginal revenue with respect to price equals zero, a revenue maximizing strategy. I would argue that CTA's response to exogenous challenges bears out these predictions.

Contrariwise, it should be said that in 1947 fares were probably too low as a result of wartime inflation and strict price regulation by the Illinois Commerce Commission. The latter was a response to the public outcry over price gouging by streetcar companies in the 1890s. Exempted from price regulation, the CTA was able to raise prices for the first time in years. However, the CTA's preference for raising fares over cutting service will also become apparent in later time periods.

The overall impression is that the CTA was managed in a very efficient and businesslike manner. In part this was due to the personalities in control of the organization. During the crucial time between 1949 and 1954 the chairman was a retired railroad CEO. The general manager for the first 17 years was recruited from the Cleveland system, which had always been very commercially oriented, and even 40 years after his retirement remains highly revered in transit circles in Chicago.

7.2. The end of the beginning (1958–1964)

There continued to be substantial exogenous demand losses, primarily on the bus system. Albeit, the losses were less dramatic than the late 1940s. Construction of interstate highways on the edge of the city led to a relocation of blue-collar jobs from City neighborhoods to the inner ring of suburbs. Annual revenue fell by \$65m. Bus service levels were reduced in response but, proportionately were only half of the decline in bus ridership, and only saved \$11m. Consistent with the output maximizing model, the loss in revenue was recouped by continued increases in real fares that netted \$62m. Fares were raised 20%, including the institution of a charge for transfers, which previously had been free.

Unlike the earlier period, exogenous cost increases (–\$96m) were only partially offset by productivity improvements (+\$65m). Real labor cost per employee started to increase faster than in the economy in general, at 2.7% per annum compared with the national 2.2%. To some extent,

this was due to the introduction of the controversial Cost of Living Allowances (COLA), which adjusted wages for inflation independently of annual collective bargaining. The scope for productivity enhancement was largely exhausted as streetcars were eliminated.

Overall, annual profitability declined by just \$8m. However, there were signs that management would like to break out of the exogenously determined profit constraint. The State of Illinois was, unsuccessfully, approached three times over the period to obtain small operating subsidies. Moreover, in 1957 the CTA produced an expansive 20-year plan that was in stark contrast to the productivity-based modernization plan of the previous decade. Unlike the earlier plan, there was no suggestion that it should be funded by the CTA's own depreciation trust fund. A request to divert gas taxes from the State highway trust fund for capital investment was unsuccessful. At that time, the balance of power in the State of Illinois was held by "downstate," or rural, legislators who opposed funding to benefit residents of Chicago.

7.3. Desperate attempts at survival (1965–1970)

The primary challenge in this period was on the cost side. The strong economy led to an increase in exogenously determined costs of \$140m. A quarter of all exogenous cost increases occurred in just five years. Unlike earlier periods, productivity enhancements could not recoup this amount, despite the replacement of trolley buses by motor buses. Indeed from 1968, endogenously determined cost changes changed from aiding profitability to hurting it. Successful strike action coupled with the COLA adjustments saw wages grow at greater than the national rate. Indicatively, the 1971 *Annual Report* stated that it had a full complement of bus drivers for the first time since 1950. Analytically, the CTA was moving away from its efficient production possibility curve. The cost-minimization assumption of Eq. (5) no longer held. Some observers have attributed the lack of financial cost controls to the retirement of commercial era managers, and the ascendancy of those with engineering rather than managerial backgrounds.

On the demand side, unexplained positive demand residuals (+\$42m) compensated for the continued exogenous losses (-\$56m). Modest improvements in financial conditions due to reductions in bus service (+\$18m) were obliterated by the first rail transit expansions funded by federal capital grants. While these new services were in corridors that are among the most heavily used in the system, they would not be justified on a commercial basis. Total ridership on the rail system actual fell after the expansions were opened.

Starting in 1965, small subsidy funds were obtained from the State under the premise that they were compensation for reduced fares offered to school children. While the subsidies did triple in real terms, they remained small (\$25m by 1970). The subsidies were insufficient to counter the cost increase, and the CTA filled this void in characteristic fashion by increasing fares. Fares increased by 25% in real terms. In 1970 the base one-ride ticket was almost \$1.90 in 1997 prices, a level without parallel before or since. Despite a resulting \$83m revenue boost, the company was unable in 1970 to meet the provisions in its agreements with the banks to provide for principal and interest on its bonds. Indeed it was barely able to cover operating costs. It was technically bankrupt.

7.4. Subsidies galore (1971–1980)

The 1970s were wild years, with a complete change of policy from the past. Overall, annual profitability fell by \$562m during this decade compared with a net gain of \$20m in the other 40 years. About a third of this decline was due to exogenous causes. Depopulation of the city following the social turmoil of the late 1960s reduced demand by \$63m, and the two energy price shocks increased costs by \$138m. A third resulted from holding nominal fares constant in an inflationary era. Fares were devalued by 45%, and revenue fell by \$203m as a result. The final third was due to endogenous cost increases of \$218m.

The change in policy can be associated with the appointment of a federal transit planner and lawyer as chairman in 1971. He campaigned for a “new vision” of public transit to obtain city, state and federal funding of both capital and operating expenditures (see Allen, 1996; Young, 1998). He was successful due to the confluence of two events. The first was political redistricting following the 1970 census which shifted the political power from rural areas to metropolitan Chicago. The second was the convergence of city and suburban interests, as the funds that bailed out the CTA also provided support for failing bus companies in the suburbs, and the continuance of commuter train service on some failing railroads. Public funding for all of these operations was channeled through a Regional Transportation Authority (RTA) that was narrowly approved by referendum in 1974.

Public funds were used to pay off the remaining bonds, provide for depreciation, and increasingly cover operating costs. Operating subsidies increased 50-fold in monetary terms and 23-fold in real terms. The astonishing real increase of almost 40% per annum appeared to be unconstrained by political forces. It is almost as if the budget constraint became endogenous! In addition, nominal price was held constant, and betterment of the workforce seems to have become a de facto objective (and hence factor prices became endogenous). The new maximization problem could be written as:

$$\text{Max } \xi_2 = \alpha(M_R, M_B) + \beta W + \lambda_2 [P^* q(P, M_R, M_B, X_i) - c(M_R, M_B, W, T) - \pi(?)] \quad (6)$$

where α and β are weights attached to the two objectives by management. The endogenous variables are service levels (M_R, M_B), input prices (W) and technology (T) which, not surprisingly, became less relevant. An unusual feature of the 1970s was a relatively aggressive line in containing output, especially on the bus system (reducing losses by \$60m). However, this is not totally inconsistent with an output maximizing objective, given that price was no longer a choice variable, and funds were needed for labor betterment.

Labors gains in the early 1970s were in the form of hourly wages. However, with national wage controls urged as part of the fight against inflation, the union turned its attention to improvements in fringe benefits, especially sickness benefits, and liberalized scheduling. The markup of fringe benefits over wages, which had been about 20% in 1960s increased to 46% by 1980. The growth in per employee compensation relative to national trends is shown as the solid line measured on the left-hand axis in Fig. 3. This figure also shows operating subsidies in constant 1997 dollars as the dashed line measured on the right-hand axis. It is quite remarkable how the relative compensation only started to rise after subsidies were established, and then rose rapidly until subsidies were effectively capped in 1980. Similar findings in a European, Canadian and Australasian context can be found in Bly and Oldfield (1985, especially Fig. 4).

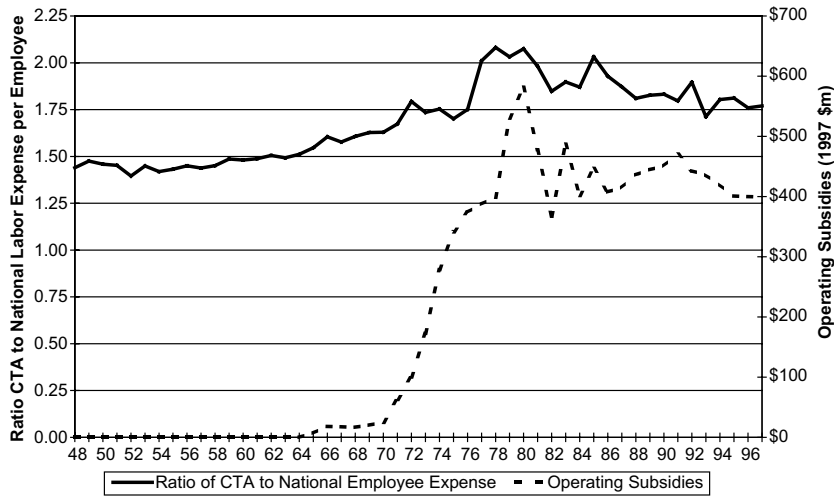


Fig. 3. Ratio of CTA to national labor cost versus operating subsidies.

Why did this increase occur? The most obvious reason is that the unions felt that they should secure their share of the new public funding, and management was unable or unwilling to prevent this. Allen (1996) asserts that the City wished to purchase labor peace in an era of persistent work stoppages by other public sector workers, notably teachers, and wildcat transit strikes in 1968 and 1974. Management acquiesced because they wanted to build political goodwill to secure additional capital funds. In addition, untimely deaths and political pressures removed the visionary managers of the early 1970s, and they were succeeded with people with operating experience and little incentive to exercise fiscal control as subsidies increased dramatically from year to year.

7.5. Free fall arrested (1981–1990)

At the turn of the new decade, three major events contributed to the end of the free fall in finances, and a reversion to the management objectives pursued before 1970. First, by 1979 it was clear that the RTA, the regional funding agency, was no longer financially viable due to unreliable taxing authority, and mounting demands for financial support from the CTA, suburban bus companies, and commuter rail operations. Second, a transit strike in December 1979, while bringing a short-term victory and higher wages, ultimately proved disastrous for the union. The public became cognizant of the considerable gains by organized labor in the previous 15 years, and how this had contributed to the run up in subsidies. Third, labor lost its political as well as its public support, as the death of the legendary Mayor Richard J. Daley marked the beginning of the end for “labor-oriented, big-city Democratic Party machine politics.”

When the funding crisis was finally resolved in 1983, conditions were placed on the CTA and the other operators to prevent a continuation of the trends of the 1970s. For the RTA as a whole, the operating revenue to cost ratio was legally required to be 50%. The CTA had to contribute to achieving this requirement. Moreover, total operating subsidies were tied to a proportion of the sales tax levy collected within City boundaries. In effect the budget constraint had again become exogenous and dependent on the same factors that influence demand, namely population and

employment within the City of Chicago, and the state of the economy. Not surprising, it seems that there was a return to a version of the maximization model of the 1950s:

$$\text{Max } \xi_3 = (M_R, M_B) + \lambda_3 [P^* q(P, M_R, M_B, X_i) - c(M_R, M_B, W, T) - \pi(X_i)] \quad (7)$$

where price has become a choice variable again, and input prices start to fall and trending back toward national levels (albeit that they still have a long way to go). There is strong evidence, consistent with this model, that management worked to reduce costs, and favored fare increases over service reductions.

Almost three-quarters of the \$109m reduction in the annual operating deficit during the 1980s was due to cost reductions. Management was able to recoup about a third of the endogenous cost inefficiencies that occurred in the 1970s (+\$76m). The CTA also benefitted from falling fuel prices and stagnant real wage growth elsewhere in the economy, which led to negligible exogenously determined cost changes (+\$3m). The first signs of gentrification and urban renewal reduced the rate of the exogenous decline in demand (-\$43m). This was almost exactly counteracted by a 20% real fare increase. While there continued to be reductions in bus service, the \$50m in savings were diverted to the rail system which expanded to the far northwest side and O'Hare Airport in 1982 and 1984.

7.6. *Budget crises (1991–1997)*

One consequence of the new budget constraint is that an economic downturn hurts the CTA in two ways: journey-to-work demand falls, and the amount of subsidy decreases due to reduced consumer spending. The recession of the early 1990s was particularly severe as it disproportionately affected white-collar downtown workers, and had lingering effects. This was compounded by residual losses in demand caused by a disastrous attempt to revise the fare structure with the aim of increasing the yield from peak-period riders. Consequently, while gentrification of the neighborhoods surrounding century-old transit infrastructure led to the smallest exogenous loss in demand in half a century (-\$33m), a residual decline in demand of \$43m and an endogenous cost increase of \$34m led to budget problems.

While fares were, characteristically, increased by 17% leading to net revenue gain of \$35m, more was needed. Consequently, a hard-fought process led to the implementation of service reductions, saving \$37m on the bus system and \$43 on the rail system. One might observe that the agency only turned to service reductions when it was in dire straights. With this combination of actions, overall profitability changed little (decreasing by only \$9m), but this had to be set against a decline in subsidy funds. There were operating deficits, after subsidies, in 1991–1995 and in 1997, which were only met by “creative accounting” involving the capital budget, and a windfall due to the closure of one rail line for reconstruction.

7.7. *A reprieve (1998–2004?)*

The analysis of this paper ends at the CTA's 50th anniversary in 1997. The years since then have witnessed a reprieve from the budget troubles of the earlier parts of the 1990s. The economic expansion, the reversal of the population decline of the city, and gentrification led to the first increases in ridership since the 1970s. The full effects of the service reductions, combined with

productivity enhancements (automatic ticketing, one-person train operation) kept costs in check. Moreover, buoyant consumer spending increased the availability of subsidy funds. Fares have been kept constant, and service has been augmented in expanding neighborhoods.

However, some observers are waiting for the bubble to burst. The economic downturn led to reduced ridership in 2002, although consumer spending (and hence sales taxes) remains strong. There is a concern that there might be a return to the persistent unfunded deficits of the previous decade. In preparation for this eventuality, the Republican-led RTA has promoted a discussion of competitive contracting of bus service, using the system in London as the model, as a means to reduce costs. The union contract currently forbids such a move, and the recent boom has placed the issue on the back burner. The onset of a significant recession will surely herald the next chapter in this unfolding drama.

8. Conclusion

In 1948 the CTA made an operating surplus of \$84m. By 1997, it lost \$458m a year. Exogenous factors played a large role. The relocation of homes and jobs and other social trends reduced revenues by \$399m a year, a large proportion of which occurred in the decade following the second world war. In addition, economy-wide doubling of real wages and fuel prices increased costs by \$575m. Three-quarters of the exogenous cost increases occurred during the booming years of the 1950s and 1960s. Management was able to recoup about half of the \$1 billion exogenous effects by increasing fares, reducing service and improving productivity.

However, between 1965 and 1980, they gave away nearly all of the productivity gains obtained in earlier years. Had the CTA been able to maintain unit costs at their mid 1960s level, it would be \$158m to the better after allowing for exogenous effects. That is to say, subsidies could be a third lower. Consequently, the citizens of Chicago would either be paying lower sales taxes and/or receiving more and better service at lower fares. Alternatively, some of the funds would be available to deal with the backlog of capital expenditures in the century-old system.

Many commentators have associated the national gains by transit labor in the 1970s to a shift from private to public ownership. This is clearly not the case in Chicago. The CTA appears to have been a well-run public agency for most of its history. The period of labor gains occurred when the budget constraint was not exogenously determined. The lessons from the 1970s were learned by the proponents of privatization in the 1980s, who argued that subsidies and cost minimization could only coexist when there is a competitive contracting relationship between the financing/planning agency and the operating companies.

Finally, for most of its existence, management has preferred to raise fares rather than cut service. Although service has been reduced, particularly on the bus system, average load factors for both bus and rail are only half what they were in 1948. It appears to be more politically palatable to share the misery among many anonymous people by means of a fare increase than to affect a smaller, identifiable, group by a service cut. This seemingly worldwide phenomenon has made riders worse off as the price/output combination moves further away from that which would maximize social welfare given the budget constraint. The size of this welfare loss in Chicago in the early 1990s was estimated by Savage and Schupp (1997) to be \$82m a year in 1997 dollars.

Acknowledgements

My interest in this subject was sparked by a Northwestern University doctoral student, Eric Petersen, who also initially collected the data from the primary sources (see Petersen, 1997). Northwestern University's transportation library and the CTA library have been invaluable. Many friends in the CTA planning department provided considerable assistance in clarifying questionable data items from many decades ago. Anonymous referees contributed valuable insights. Needless to say the inferences and opinions are solely my own.

Appendix A. Additional notes on data and sources

Until 1976, excepting 1975, data on financial and operating performance were drawn from company *Annual Reports*. As the financial situation worsened, the first casualty was public disclosure of financial information. Fortunately, the CTA archives contain the legally required auditor's letter which provides the necessary information for 1975 and 1977–1986. Since 1987 the Federal Transit Administration (formerly the Urban Mass Transportation Administration) standardized "Section 15" reports have been used.

Ridership is measured as unlinked passenger trips (a ride on the bus and then on the rail system counts as two trips). These data are available since 1961, when a charge was introduced for transfers. Prior to that time, data were only collected on the number of passengers whose trip originated on a specific mode. Fortunately both data were collected in parallel between 1961 and 1976. A ratio of unlinked to originating trips of 1.6 for the bus system and 1.54 for the rail system can be calculated, and applied retroactively to produce unlinked trip figures for 1948–1960. This paper does not deal with subcontracted paratransit services for elderly and physically challenged individuals. Neither the cost nor revenues from these services appear in this analysis. Compared with traditional transit services, paratransit is only a small part of transit operations in Chicago representing about 3% of the CTA's budget.

Between 1982 and 1988, reported data on railcar mileage are very suspect with wild changes which are inconsistent with documented service variations. Internal CTA railcar odometer data were consulted, and appear to be more reflective of reality. These data have been substituted for the traditional measures that are derived from the schedules for 1982–1988.

Cost data has only been disaggregated by mode since 1982. Cost per bus mile (C_B) and cost per railcar mile (C_R) has to be inferred for the earlier period. Between 1982 and 1997, the average ratio of modal unit costs (C_R/C_B) was 0.89. This ratio is used to modify the identity:

$$TC = C_R M_R + C_B M_B$$

where TC is the reported total operating cost to:

$$TC = 0.89 C_B M_R + C_B M_B$$

Collecting terms, the estimated unit cost of bus service (C_B^*) in any year between 1948 and 1981 is:

$$C_B^* = TC / (0.89 M_R + M_B)$$

and the estimated cost per railcar mile (C_R^*) is $0.89 C_B^*$.

The measure of exogenous wage pressures is a modification by Gordon (1995) to the average hourly wage for private non-agricultural industries. The motivation for this modification was the apparent decline in national average wages starting from the early 1970s. Macroeconomists argue that this was due to myriad factors including the move from heavy industry to service occupations; the increased proportion of salaries based on commissions; the rise of part-time work; entry of women into the labor force; and a geographic redistribution of jobs to the southern United States where wages and the cost of living are lower. Gordon's modification is based on the labor share of the national income accounts, and includes allowance for increases in overtime payments, and employer-paid fringe benefits and social security contributions.

The ratio of the workforce in the City of Chicago to that in the whole of Cook County, which covers both Chicago and the inner ring of suburbs, is derived from the State of Illinois, Department of Employment Security *Where Workers Work*. This publication covers private sector employees. Before 1991 workers were recorded as working at the main location of the reporting firm, but since then has tried to reflect the actual location of employment. Data are available for 1955, 1957, 1959–1979, 1981–1986 and 1988 onwards. Interpolated data are used for 1956, 1958, 1980 and 1987. The *United States Census* also measures location of employment at the turn of each decade. This source covers both private and public sector employees. Where a direct comparison can be made between the two sources, the Census provides a ratio that is 6% lower than that obtained from *Where Workers Work*. The 1950 Census is used to estimate a figure for 1950, and hence by interpolation figures for 1948–1949 and 1951–1954 were obtained. Fortunately, given the data difficulties, the rapid suburbanization of jobs did not begin in earnest until the late 1950s.

Inferences from the analytical model are complemented by reference to primary and secondary sources. The chairman's letter in CTA *Annual Reports* present a "public face" on the objectives of the agency until these were discontinued in the mid 1970s. There are also some excellent secondary sources. Allen (1996) makes astute observations as a political scientist in recounting events, especially those of the 1970s, in a thesis supervised in part by transportation economists, and based on extensive interviews with those involved. Young (1998), a longtime transportation writer for the *Chicago Tribune*, provides incisive analysis of the political shenanigans that have been a feature of transit in Chicago for over a century.

References

- Allen, J.G., 1996. From Centralization to Decentralization: The Politics of Transit in Chicagoland. Ph.D. Thesis, Massachusetts Institute of Technology, USA.
- Bly, P.H., Oldfield, R.H., 1985. Relationships between public transport subsidies and fares, service, costs and productivity. Research Report 24, Transport and Road Research Laboratory, Crowthorne, UK.
- Glaister, S. (Ed.), 1987. Transport Subsidy. Policy Journals, Newbury, UK.
- Glaister, S., Collings, J.J., 1978. Maximization of passenger miles in theory and practice. *Journal of Transport Economics and Policy* 12, 304–321.
- Goodwin, P.B., 1992. A review of new demand elasticities with special reference to short and long run effects of price changes. *Journal of Transport Economics and Policy* 26, 155–169.
- Gordon, R., 1995. The American Real Wage Since 1963: Is it Unchanged or Has it More than Doubled? Mimeo, Northwestern University.

- Kain, J.F., 1999. The urban transportation problem: A reexamination and update. In: Gómez-Ibáñez, J., Tye, W.B., Winston, C. (Eds.), *Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer*. Brookings Institution Press, Washington, DC, pp. 359–401.
- Nash, C.A., 1978. Management objectives, fares and service in bus transport. *Journal of Transport Economics and Policy* 12, 70–85.
- Obeng, K., Sakano, R., 2002. Total factor productivity decomposition, input price inefficiencies, and public transit systems. *Transportation Research: Part E* 38E, 19–36.
- Oum, T.H., Tretheway, M.W., Waters II, W.G., 1992. Concepts, methods and purposes of productivity measurement in transportation. *Transportation Research: Part A* 26A, 493–505.
- Petersen, E., 1997. Transit subsidies in New York and Chicago: local, state and federal assistance in a historical context. *Journal of Public Transportation* 1, 23–41.
- Pickrell, D.H., 1985. Rising deficits and the uses of transit subsidies in the United States. *Journal of Transport Economics and Policy* 19, 281–298.
- Savage, I., Schupp, A., 1997. Evaluating transit subsidies in Chicago. *Journal of Public Transportation* 1, 93–117.
- Webster, F.V., Bly, P.H., 1980. *The Demand for Public Transport*. Transport and Road Research Laboratory, Crowthorne, UK.
- Winston, C., Shirley, C., 1998. *Alternative Route: Toward Efficient Urban Transportation*. Brookings Institution Press, Washington, DC.
- Young, D.M., 1998. *Chicago Transit: An Illustrated History*. Northern Illinois University Press, DeKalb, IL.