RISEING DEFICITS AND THE USES OF TRANSIT SUBSIDIES IN THE UNITED STATES

By Don H. Pickrell*

TRANSIT DEFICITS AND SUBSIDIES

Transit ridership in the United States, after declining for more than 25 years, began in 1973 to climb slowly. Many cities, which had steadily curtailed transit service throughout the postwar era, restored it rapidly. Transit vehicles now operate over nearly 125,000 track and street miles, more than a third of which have been added since the low ebb of 1973 (American Public Transit Association (APTA), 1973; U.S. Department of Transportation (USDOT), 1984). Between 1973 and 1983 total operating expenditures by US transit systems nearly quadrupled (reaching $8.3 billion). A rapidly declining fraction was covered by farebox receipts. Barely two decades after first failing to meet its operating budget from fare revenues alone, the industry had an annual deficit of $5.4 billion. At the same time, losses spread from a handful of rail transit systems in the largest urban areas to operators of all modes of transit service in virtually every US city.

These escalating deficits were funded by direct payments — first from local agencies and later from states and the federal government — the cumulative total of which exceeded $25 billion by 1983 (APTA, 1981a; USDOT, 1982, 1983, and 1984). One interpretation of these events is that subsidies have escalated somewhat haphazardly in response to financial pressures caused by rising costs and declining demand; another is that their growth reflects positive government intentions to reverse historical trends by extending transit service and stabilising fares. Probably both have been true at different times. The reasons underlying the rapid increase of subsidies no doubt also differ among the various levels of US government. The growth in transit deficits raises several serious questions. For instance, how extensively have subsidies been absorbed by cost increases? Can the myriad goals of urban transport policy be promoted by expanding conventional mass transit service at costs acceptable to riders and taxpayers, who ultimately bear them in some combination? This paper identifies specific sources of the recent changes in operating expenses and farebox revenues of US urban mass transit. They can be interpreted as the causes of rising deficits, or simply as the purposes served by subsidy payments.

* John F. Kennedy School of Government, Harvard University, and Transportation Systems Center, U. S. Department of Transportation, Cambridge, Mass. The research reported here was supported by the U.S. Urban Mass Transportation Administration under its University Research Program. Douglass Lee, John Meyer, Jose Gomez-Ibanez and Larry Oxsey all offered helpful comments, but none of them can be implicated in any errors that remain.
A COMPARATIVE STATIC ANALYSIS OF GROWTH OF DEFICITS

Growth in the total operating deficit, defined as aggregate operating expenditure\(^1\) less fare revenue, of the US urban transit industry can be allocated to several basic sources by a simple comparative static model. If we express total expenditure as the product of average expenditure per vehicle-mile of service (denoted \(C\)) times the number of vehicle-miles of service operated nationwide \((S)\), and total fare revenue as the product of nationwide annual transit ridership \((R)\) times the average fare paid per rider \((F)\), the total deficit during any year is

\[ D = SC - RF \]

(Because operating expenditure exceeded fare revenue throughout the period of interest here, the deficit is defined directly as a positive quantity, rather than as a negative value for operating income.) The change between any two years in the annual deficit on transit operations is

\[ \Delta D = \Delta(SC - RF) = \Delta(SC) - \Delta(RF) \]  \hspace{1cm} (1)

During the period considered here, the rapid increase in government subsidies for urban transit made service planning decisions by transit operators largely independent of changes in the unit costs at which services could be provided. When the levels of service and unit operating expenses change independently but by non-differential amounts, the resulting change in total expenditure is given by

\[ \Delta(SC) = S \Delta C + C \Delta S + \Delta C \Delta S \]  \hspace{1cm} (2)

Similarly, the change in total revenue can be written as

\[ \Delta(RF) = R \Delta F + F \Delta R + \Delta F \Delta R = R \Delta F + (F + \Delta F) \Delta R \]

A complication is that ridership presumably depends on both the average fare level and the level of service provided, as well as upon a number of other factors influencing the demand for urban transit. Denoting these various factors as \(X_1, X_2, \ldots, X_n\), this dependence can be written

\[ R = f(F, S; X_1, X_2, \ldots, X_n) \]

From this it follows that

\[ \Delta R \equiv (\partial R/\partial F) \Delta F + (\partial R/\partial S) \Delta S + \sum_{i=1}^{n} (\partial R/\partial X_i) \Delta X_i \]

Thus the change in total revenue can be expressed as

\[ \Delta(RF) \equiv R \Delta F + (F + \Delta F)[(\partial R/\partial F) \Delta F + (\partial R/\partial S) \Delta S + \sum_{i=1}^{n} (\partial R/\partial X_i) \Delta X_i] \]

\(^1\) Only direct operating expenditure is included in this analysis, partly because the specialised nature of some transit capital facilities, such as underground or elevated rights-of-way, and the historical involvement of government in financing transit capital expenditure make it difficult to determine their opportunity costs and depreciation rates. Nevertheless, capital costs for transit are both real and substantial, and it is important to recall that they are a significant element of the full costs of transit service, despite their typical omission from estimates of operating losses such as those reported here.
or, where \( e_j \) denotes the elasticity of ridership with respect to variable \( j \),

\[
\Delta(DF) \approx R \Delta F + (F + \Delta F)[(R \Delta F/F)e_F + (R \Delta S/S)e_S + \sum_{i=1}^{n} (R \Delta X_i/X_i)e_{Xi}]
\] (3)

Substituting (2) and (3) into (1), rearranging, and collecting terms ultimately yields

\[
\Delta D \approx S \Delta C + \left[ C \Delta S - (F + \Delta F)(R \Delta S/S)e_S \right] + \Delta S \Delta C
+ \left[ R(-\Delta F) - (F + \Delta F)(R \Delta F/F)e_F \right] - [(F + \Delta F) \sum_{i=1}^{n} (R \Delta X_i/X_i)e_{Xi}]
\] (4)

Since \( D \) is positive and increasing over the period studied, each term of (4) has a convenient interpretation as expressing the contribution of specific developments to increasing deficits or, symmetrically, the disposition of rising subsidies to various uses (some of which were no doubt unintended). The first term represents the increased expenditure associated with rising operating costs per unit of service. The next bracketed collection of terms represents the net increase in expenses from an extension of urban transit services; it consists of a direct cost \( (C \Delta S) \) of service increases partly offset by fare revenue contributed by riders attracted to the new services.\(^2\) As expected, the magnitude of the new fare revenue depends on the proportional change in service offered.

The third term in (4) captures the effect of simultaneous increases in unit operating expenses and service levels, while the last two bracketed collections of terms are analogous to the net service cost term. The earlier of these two represents the net effect of fare reductions: that is, the revenue “lost” on each of the original number of trips, less the added revenue contributed by riders induced by the lower fares to make new trips. The last bracketed collection of terms summarises the contributions of changes in the various determinants of transit demand to the industry’s changing financial situation. Though their individual effects would be difficult to isolate, their combined contribution can be estimated as the residual remaining after the values of the first four terms are deducted from the change in the total deficit; these can all be computed directly from readily available data. Again, each of these effects can be interpreted in two ways, depending partly on one’s viewpoint; each can be seen as a source of increasing deficits, or as a use — intended or otherwise — of the rapidly growing subsidies appropriated by government.

Estimates of the sources of growth of deficit

From equation (4), growth in the nationwide transit deficits and subsidies can be allocated among increases in unit operating expenditure, increases in the level of service provided, declining demand for transit travel, and reduced fares. Table 1 presents estimates of the distribution of growth in deficit among these developments for the US as a whole over the period 1970-82. It shows that rapid growth

\(^2\) The need to nett out the fare revenue contributed by riders attracted by extended services from the cost of operating them was pointed out by an anonymous referee of this Journal, who also made several other valuable comments and suggestions.
in operating expenditure per vehicle-mile of transit service absorbed more than 60% of the increase in subsidy payments. About 9% apparently financed expanding levels of transit service, while simultaneous service extensions and rising unit costs interacted to absorb another 8% of the expanding subsidy funds.³

Table 1 also reports that underwriting fare reductions accounted for slightly more than 14% of the increase in governmental assistance, leaving 7.5% to compensate for declining demand for transit service, the primary cause of the industry’s long-term historical contraction (see Meyer et al., 1965).⁴ In other words, 7.5% of growth in governmental operating assistance for transit apparently financed fare reductions and service increases that by themselves would have been sufficient to sustain ridership at its 1970 level.⁵ Subsequent sections examine in more detail the developments underlying rising unit costs and declining demand for transit, and explore the changing characteristics of service patterns and fare structures over this period.

THE CAUSES OF RISING OPERATING EXPENDITURE

In order to analyse why expenditure rose so rapidly, this section develops a model of transit costs, reports estimates of its parameters for a sample of US bus transit operators, and uses them to infer the relative importance of different causes of cost escalation. Transit operators combine capital — in the form of rights-of-way, vehicles, and fixed facilities — with labour services, propulsion energy, maintenance supplies, and miscellaneous other inputs to produce passenger-carrying capacity. Apart from changing one mode for another, the prevailing technology for producing transit service offers limited opportunity for substitution among these inputs (see Williams and Dalal, 1981; Viton, 1981; and Berechman, 1982), so costs are likely to vary primarily as a function of the level of output and the prices paid for capital, labour, and operating inputs.

³ The net contribution of service extensions was estimated at a service elasticity of demand of +0.4. The resulting estimate would be lower for higher values of this elasticity (since they would suggest that more new riders are attracted by a given service increase), but various surveys (Pucher and Rotherberg, 1979; Ecosometrics, 1980) suggest that this is a reasonable estimate. Vehicle-miles as a measure of service has the advantage of ready historical availability, but it complicates intermodal or historical comparisons because the passenger-carrying capacity of transit vehicles in the same mode have increased slowly over time. A more informative measure might be miles of passenger-carrying capacity, defined to include seating capacity plus an allowance for comfortable standing loads. Yet, from the passenger’s viewpoint, vehicle-miles may be a better measure of service, since the frequency of arrivals of vehicles directly affects waiting times, while the capacity of individual vehicles becomes important only when passenger loads approach it. Consistent estimates of this concept of capacity are impossible to construct, either historically or for different transit systems at any one time.

⁴ The net contribution of fare reductions was estimated at a fare elasticity of demand of −0.3. Larger absolute values of this elasticity would reduce the estimated contribution of fare reductions, because they would suggest that the fare reductions observed attracted more additional passengers. Again, however, various surveys suggest that this is a plausible estimate.

⁵ This residual also includes any error with which the quantity ΔD is estimated by expression (4), so it may be a slight over- or under-estimate of the share of subsidy growth necessary to compensate for declining demand for transit.
### TABLE 1


<table>
<thead>
<tr>
<th>Factor</th>
<th>1970</th>
<th>1982</th>
<th>1982 $ million</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating expense per vehicle mile (1982 dollars)</td>
<td>$2.19</td>
<td>$3.54</td>
<td>2,542</td>
<td>61.1</td>
</tr>
<tr>
<td>Total vehicle-miles of service (millions)</td>
<td>1,883</td>
<td>2,131</td>
<td>384</td>
<td>9.2</td>
</tr>
<tr>
<td>Interaction of rising expenses and service</td>
<td>–</td>
<td>–</td>
<td>335</td>
<td>8.1</td>
</tr>
<tr>
<td>Average fare per passenger (1982 dollars)</td>
<td>$0.64</td>
<td>$0.51</td>
<td>587</td>
<td>14.1</td>
</tr>
<tr>
<td>Demand for transit service (million trips/year)</td>
<td>5,932</td>
<td>6,038(^a)</td>
<td>310</td>
<td>7.5</td>
</tr>
<tr>
<td>Total deficit (millions of 1982 dollars)</td>
<td>328</td>
<td>4,486</td>
<td>4,158</td>
<td>100.0</td>
</tr>
</tbody>
</table>

\(^a\) Actual 1982 ridership; estimated ridership at the 1970 average fare, assuming a fare elasticity of −0.3, would be 5,640 million trips.

Sources: APTA (1981), USDOT (1984), and computations using expression (4) in text.

The actual input mix used by transit operators may be constrained further by specific labour agreements that limit the length of the intervening break in driver shifts that are split (scheduled in two parts to encompass both morning and evening peak periods), the number and overall duration of split shifts, and the use of part-time vehicle operators (Chomitz and Lave, 1981). Because of the characteristic peaking in demand during two short but widely separated periods of the day, these provisions can sharply increase the number of driver payhours necessary to produce a schedule of service. Actual expenditure for producing the desired level of service may be still further increased if the availability of direct subsidies also induces technical inefficiency in the use of inputs.

This basic cost structure can be complicated by certain features of transit...
service itself and the environment in which it is provided, which violate the usual assumptions that firms have identical outputs and conditions for production. Outputs differ because different transit systems provide varying fractions of service during peak periods, and because some provide frequent service over dense route networks while others operate vehicles infrequently on expansive route systems. Where buses operate in conjunction with rail transit, they may provide mostly feeder and intrasuburban service, but in cities without rail service buses also serve the linehaul function. Again, different operators' inputs are not necessarily identical: capacities and operating speeds can differ substantially, depending on the ages, original design standards and levels of use and maintenance of rights-of-way and vehicles. Thus actual operating expenditure is likely to depend on the aggregate level of passenger-carrying capacity provided; prices paid for labour, fuel and other operating inputs; certain characteristics of the service and of the urban environment in which it is operated; and specific provisions in labour contracts governing driver assignments.

Econometric analysis of operating expenditure in bus transit

Table 2 reports estimates of the parameters of several linear unit expenditure functions for urban bus transit, including different combinations of these explanatory variables. These estimates are based on a sample of 68 US urban bus transit operators with fleets of 100 or more vehicles who reported the necessary financial and operating data for 1980 (APTA, 1980a and 1980b; USDOT, 1982). The assumption that unit expenditure is a linear function of factor prices and output characteristics implies a restrictive version of the underlying production function and does not allow its properties to be tested, but still allows the effects of important variables to be analysed, while partly correcting for the heteroscedasticity encountered when total expenditure is used as the dependent variable. The specifications tested are only moderately successful in accounting for variation in unit expenditure, and the effects of some factors hypothesised to have important influences cannot be reliably detected; nevertheless, the results do provide some useful information on the structure of transit operating costs, as well as on the causes of their recent escalation.

The coefficient estimates suggest that a one-dollar increase in the hourly driver wage (often a benchmark for other wage rates faced by transit operators) raises expenditure per vehicle-mile by 15 to 18 cents, which amounts to an elasticity of 0.62 to 0.71 evaluated at the variable means. This suggests that possibilities of factor substitution for driver labour are in fact very limited, since the share of labour costs in bus operating expenses averages 74% (USDOT, 1982). Thus increases in wage rates tend to be translated directly into increases in operating cost. The estimated coefficients on the fuel price variable show the expected (positive) sign and have reasonable magnitudes, but they are not much larger than their standard errors. Disregarding the associated uncertainty about their true values, the point coefficient estimates imply unit cost elasticities with respect to fuel prices of 0.11 to 0.14, approximately the average share of fuel cost in expenditure (USDOT, 1982). Again, there appears to be little opportunity to substitute other inputs for fuel, so that increases in its price are largely reflected in increases in costs.
TABLE 2


<table>
<thead>
<tr>
<th>Variable</th>
<th>Regression Coefficient (Standard Error) in Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.1507 (0.7769), 0.1428 (0.8525), 0.1296 (0.7484), 0.1370 (0.7324), 0.1208 (0.8412), 0.1267 (0.6465)</td>
</tr>
<tr>
<td>Driver wage rate ($/hour)</td>
<td>0.1820 (0.0682), 0.1628 (0.0636), 0.1499 (0.0661), 0.1614 (0.0670), 0.1770 (0.0618), 0.1689 (0.0665)</td>
</tr>
<tr>
<td>Fuel price ($/gallon)</td>
<td>0.3537 (0.2890), 0.2918 (0.3139), 0.3054 (0.3222), 0.3499 (0.3232), 0.3032 (0.3344), 0.3663 (0.3463)</td>
</tr>
<tr>
<td>Average speed of service (mph)</td>
<td>−0.1377 (0.0629), −0.1382 (0.0638), −0.1379 (0.0637), −0.1386 (0.0658), −0.1395 (0.0649), −0.1398 (0.0654)</td>
</tr>
<tr>
<td>Average age of buses (years)</td>
<td>−0.0202 (0.0232), −0.0252 (0.0286), −0.0290 (0.0264), −0.0229 (0.0227), −0.0216 (0.0268), −0.0286 (0.0249)</td>
</tr>
<tr>
<td>Ratio of peak to midday buses in service</td>
<td>0.2616 (0.1044), 0.2495 (0.1025), 0.2768 (0.1032), 0.2311 (0.1034), 0.2635 (0.1038), 0.2397 (0.1023)</td>
</tr>
<tr>
<td>Annual vehicle-miles per route-mile (thousands)</td>
<td>0.0236 (0.0041), 0.0217 (0.0042), 0.0219 (0.0052), 0.0210 (0.0052), 0.0214 (0.0053)</td>
</tr>
<tr>
<td>Rail transit dummy (1=yes; n=8)</td>
<td>0.2762 (0.1613), 0.2896 (0.1658), 0.2501 (0.1650), 0.2172 (0.1663)</td>
</tr>
<tr>
<td>Proportion of operating budget from dedicated taxes</td>
<td>0.5171 (0.2493), 0.5199 (0.2517), 0.5434 (0.2265)</td>
</tr>
</tbody>
</table>

| Spread premium time:                          |                                                         |
| 10:00 hours or less (n=16)                   | 0.4757 (0.2097), 0.4517 (0.2185)                       |
| 10:01—11:00 hours (n=33)                     | 0.1030 (0.2131), 0.1024 (0.2070)                       |
| 11:01—12:00 hours (n=18)                     | 0.1004 (0.2078), 0.0882 (0.2035)                       |

| Maximum spread time:                         |                                                         |
| 10:30—12:00 hours (n=14)                     | 0.5481 (0.2576)                                        |
| 12:01—13:00 hours (n=31)                     | 0.3398 (0.2249)                                        |
| 13:01—14:00 hours (n=16)                     | 0.2474 (0.2134)                                        |
| Adjusted R²                                   | 0.5042, 0.5789, 0.6097, 0.6623, 0.6735, 0.6925 |
| Standard error of estimate<sup>a</sup>        | 0.4497, 0.4209, 0.4138, 0.3799, 0.3731, 0.3675 |

<sup>a</sup> Around a mean of $2.12.
The average age of vehicles does not show a strong association with unit expenditures, apparently because variation in past utilization and maintenance, combined with the higher operating costs of some newer buses, offsets the expected escalation of operating and maintenance expenses with age of fleet. In contrast, the average speed of bus operation—which directly affects the number of vehicles and drivers necessary to produce the scheduled frequency of transit service—does show a significant effect on operating costs. According to the coefficient obtained with this sample, improving operating speeds by one mile per hour (or 9%) could reduce costs per vehicle-mile by about 7%; this is again consistent with the fact that direct hourly labour expenses account for about three-quarters of the total costs of operating bus transit.

Of the various service characteristics hypothesized to influence operating expenses, the degree of peaking in daily service appears to have the most pronounced effect. Its estimated coefficient suggests that, for example, a transit system with twice as many buses in peak as in midday service (approximately the sample average) will have unit costs 12-14% above those of an otherwise identical system with equal number of buses in peak and midday service. Apparently operating expenses are also slightly higher where frequent service over a dense route network is provided, even after accounting for the slower operating speed which this entails, while unit costs may also be somewhat higher for buses operated in conjunction with rail transit. In addition, there is some tentative evidence that government operating assistance for transit is partially absorbed by escalation in unit operating expenses: the coefficient estimates for the dedicated subsidy variable suggest that nearly 20% of an increase in operating assistance financed from earmarked tax sources may be absorbed by higher unit costs.6

Finally, interest attaches to the estimated effects of varying lengths of work shifts after which premium rates apply, and the maximum permissible duration of shifts. The variables describing these provisions are coded to allow varying effects of progressive changes in their restrictiveness; thus premium pay thresholds between 10 and 12 hours do not appear to increase costs significantly over situations where no restriction is specified, but to require premium pay after 10 hours or less typically raises unit costs about 23%. The pattern of estimated effects of maximum shift lengths is similar, though the results are more ambiguous: restrictions in the range of 13–14 hours may not raise unit costs significantly, but 12–13 hour maximum shifts may raise them by 15%, and costs appear to rise sharply if this rule is tightened to prohibit shifts exceeding thresholds in the range of 10.5–12 hours.

---

6 Total annual operating expenditure averages $45.92 million for the sample; annual vehicle-miles average 16.98 million. If the proportion of expenses funded by dedicated taxes rises by 0.01 (or to 0.46 million), the coefficient estimates indicate that operating costs will rise $0.0052–0.0054 per vehicle-mile, or $0.088–0.092 million in total. Thus almost 20% of the increased assistance would be absorbed by higher unit costs, leaving the remainder available to underwrite service expansions or fare reductions. Only the effect of earmarked or dedicated subsidies was tested here, because specification and estimation complications would have been posed by the presumably simultaneous relationship between total subsidy levels and total expenditure.
Inferences about the causes of rising unit operating expenditure
These estimates are based on a cross-section analysis for a single year, but they provide some useful evidence about the causes of rapid escalation in expenditure over the period studied. Most important, the average hourly driver wage paid by large bus transit authorities rose from about $4 in 1970 to nearly $8.50 by 1980 (US Bureau of Labor Statistics, 1970; Amalgamated Transit Union, 1980). If the estimated effect of wage variation on unit operating costs was stable over this period, wage increases alone may thus have accounted for nearly half the increase of approximately 170% in unit costs. Fuel costs rose rapidly at the same time, from 11 cents per gallon in 1970 to 95 cents by 1980 (APTA, 1970 and 1981b). This analysis could not identify a significant effect of differences in the price of fuel after it had once reached this higher average level; but increases of this magnitude were almost certainly responsible for some (the point coefficient estimates in Table 3 suggest about 30%) of the rapid growth in operating expenses.

The slight reduction over this period in the degree of peaking in scheduled bus service should, at least according to the estimated effect of peaking on unit costs, have reduced expenses slightly, unless it occurred in response to the increased restrictiveness of labour contracts governing driver assignments. Several transit systems imposed new provisions requiring premium pay for split shifts or restricting overall shift lengths, while others tightened existing restrictions (Chomitz and Lave, 1981). Because they require transit operators to purchase additional off-peak labour hours to maintain a given schedule of peak service, these more restrictive work rules would thus have reduced the marginal cost of expanding off-peak service. By reducing the number of vehicles in peak service that were withdrawn during the midday (thereby producing the observed decline in the peaking measure used here), operators may have minimised the effect of more restrictive work rules on average unit costs, but total operating expenditure for providing typical schedules of rush-hour service would still have risen.

Finally, the estimated effects on transit costs of earmarked operating assistance suggest that expanding government subsidies also contributed to escalating costs. Nearly 60% of the aggregate operating budget for public transit in the US is now funded by government assistance, with revenues from earmarked tax sources and formula grants each making up well over one third of the total (USDOT, 1982). Thus, even if as little as the 20% of dedicated assistance estimated here has been absorbed by higher unit operating expenditure, the growing availability of government subsidies (which were quite small at the outset of the period studied) may itself have been responsible for 5–10% of the increase in costs for providing transit service.

FACTORS REDUCING THE DEMAND FOR TRANSIT SERVICE

Urban decentralisation and changing travel patterns
The demand for transit travel declined partly because continuing dispersion of residences, employment and population-serving activities within US cities reduced travel volumes on routes where transit was competitive with private automobile travel. Table 3 shows recent shifts in residential and employment locations and
### TABLE 3

**Percentage Changes in the Geographic Distribution of Travel to Work Within US Urban Areas, 1970–1980**

<table>
<thead>
<tr>
<th>Residence Zone</th>
<th>Employment Zone</th>
<th>Central City</th>
<th>Suburban Ring</th>
<th>Entire SMSA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Central city</td>
<td></td>
<td>-16.3</td>
<td>+22.5</td>
<td>-10.7</td>
</tr>
<tr>
<td>All work trips</td>
<td>Transit work trips</td>
<td>-31.9</td>
<td>-33.0</td>
<td>-32.1</td>
</tr>
<tr>
<td>Suburban ring</td>
<td></td>
<td>+7.9</td>
<td>+33.6</td>
<td>+24.1</td>
</tr>
<tr>
<td>All work trips</td>
<td>Transit work trips</td>
<td>+4.6</td>
<td>-21.5</td>
<td>-4.0</td>
</tr>
<tr>
<td>Entire SMSA</td>
<td></td>
<td>-8.8</td>
<td>+31.5</td>
<td>+7.1</td>
</tr>
<tr>
<td>All work trips</td>
<td>Transit work trips</td>
<td>-24.7</td>
<td>-26.5</td>
<td>-24.9</td>
</tr>
</tbody>
</table>

*All US Standard Metropolitan Statistical Areas.


the resulting changes in commuting patterns. It indicates that the number of work trips originating in the densely developed central areas of the nation’s cities fell by more than 10% between 1970 and 1980, while the number originating in suburban areas increased by 24%, as large numbers of urban households continued to relocate from central cities to their surrounding suburbs. One factor contributing to this decentralisation was continuing growth in real per capita incomes: urban households, though sharply reduced in size, continued to use some of their rising incomes to purchase the more spacious and private living arrangements to be found in suburban areas.7

Though cause and effect in employment and residential relocation are difficult to distinguish, another source of decentralising urban development was the dispersion of employment into suburban areas, also shown in Table 3. The number of workers commuting to central city jobs in US metropolitan areas fell by about 9% during the 1970s, while the number commuting to suburban jobs grew by more than 31%. This occurred partly because technological innovations in manufacturing continued to stimulate industrial demand for large parcels of land, thus fostering the relocation of some manufacturing jobs outside central cities. More important, changes in economic activity concentrated increases in employment...

7 In constant dollars, the average annual income of households in US SMSAs rose by 15.8% from 1960 to 1970, and by another 9.3% to 1980. Because the size of typical urban households was falling throughout this period, real per capita income of urban residents rose by 28.2% from 1960 to 1970 and by another 27.4% from 1970 to 1980 (US Bureau of the Census, 1981 and 1984).
among lighter manufacturing, trade, and service industries. Because the production technologies of these industries are typically more land-intensive, and they depend less on access to centralised transport than traditional manufacturing activities, their increasing importance in the nation's economy contributed to rapid decentralisation of employment within its urban areas.

Table 3 also illustrates that the changing distribution of population and employment produced significant changes in patterns of travel to work. The number of work trips made entirely within the central cities of urbanised areas, the traditional stronghold of transit service and patronage, fell by more than 16% during the decade. Growth continued in the other main market in which transit often competes favourably with auto travel, radial commuting to central city jobs from suburban residences, but at a comparatively modest pace. Very rapid growth in commuting occurred in outbound radial (a 22.5% increase over the decade) and intrasuburban (33.6%) corridors, where mass transit often cannot compete favourably with automobile travel because of the characteristically diffuse patterns of trip origins and destinations. Yet Table 3 suggests that commuting by public transit declined by far more than would have resulted even from these pronounced changes in urban travel patterns: it shows sharp reductions in the use of transit within central cities (where it fell by nearly 32%), in outbound radial corridors (33%), and within suburban areas (over 21%). The only increase in commuting by transit (of about 5%) was from suburban residences to work locations in the central city; and this was largely confined to the few older, congested urban areas where employment remained highly centralised and the capacity of streets and highways serving their central areas was not substantially improved.

The sharply declining role of the transit mode implied by these data resulted partly from the growth already discussed in the incomes of urban residents, together with continued developments that facilitated automobile travel. Rising incomes affected transit demand by increasing urban travellers' valuations of the advantages offered by auto travel, particularly its minimal access and waiting times, high speed, and scheduling flexibility. Its higher travel speeds were facilitated by the additions to road and highway capacity in which most urban areas continued to invest during this period. These developments were reflected in the very rapid growth in auto ownership: per capita auto registrations in US cities rose from 0.38 in 1970 to 0.55 by 1982, or by about 45% (US Federal Highway Administration, 1980 and 1984). At the same time, substantial underpricing of much of peak hour urban auto travel, and widespread subsidisation of parking costs, may have yet further reduced the demand for transit, though it is not clear whether these inducements to auto commuting actually became stronger during this period.8

8 Straszheim (1980) estimates that peak-hour auto travellers pay only about two-thirds of the costs they typically impose, while Pickrell and Shoup (1982) report that more than three-quarters of those who commute to work by auto in urban areas use parking provided by their employers either free or substantially below the market price. Nevertheless, some estimates of the elasticity of transit demand with respect to the price of auto travel suggest that even this substantial underpricing may not have been an important cause of declining demand for transit, though there is disagreement about its exact magnitude; see Pucher and Rothenberg (1979).
Ineffective deployment of transit services

Another source of the fall in demand for transit during the past decade was apparently the failure of transit operators to adapt service policies to changing patterns of urban travel. Instead of identifying specific routes where service that was sufficiently rapid and frequent to attract acceptable ridership could be maintained at reasonable costs, most US transit operators expanded service into widespread new markets. Between 1972 and 1982, total bus route miles served by the nation's largest transit operators increased by nearly 35%, but the accompanying rise in vehicle-miles operated amounted only to about 13% (computed from APTA, 1973, and USDOT, 1984). Thus annual bus-miles operated per route mile, an index of the typical frequency of transit service, fell by nearly 20% during this period.

This decline partly reflected the widespread extension of relatively infrequent bus service into expanding suburban areas, though some transit operators simultaneously reduced service frequencies on central city routes. Another contributing factor was probably the new bus transit systems in many smaller cities and towns which provided infrequent service over expansive route networks. The demand for transit depends in part on the frequency of service (since headways directly affect passenger waiting times); thus extending infrequent transit service into areas of dispersed travel patterns and convenient auto travel probably caused an even greater reduction in demand than would necessarily have resulted from shifting travel patterns and other economic developments in US urban areas.

CHANGES IN TRANSIT FARE STRUCTURES

Constant-dollar average fares fell during the 1970s, partly because transit operators used expanding government assistance to stabilise basic cash fares (which had risen throughout the preceding decade) during a period of sustained price inflation. Table 4 reports recent changes in basic adult fares, as well as average revenue actually received, for transit trips in 26 large urban areas which together account for approximately two-thirds of US transit ridership. It indicates that, after adjusting current dollar base fares for inflation, fares in cities with both rail and bus service rose only slightly between 1974 and 1981, while fares typically charged by large bus transit systems actually fell somewhat. Table 4 also indicates that growth in real average revenue per passenger — a better measure of the typical fare actually paid — fell by 4% over this period in cities offering rail service, and by more than 17% in large bus systems.9

Actual fare revenue per passenger fell, partly because many transit operators introduced selective fare reductions for specific groups of riders. By 1981, all these 26 cities offered at least off-peak fare discounts to senior citizens (a federal mandate), while 11 offered lower fares to children, and 22 carried students at reduced fares (APTA, 1981c). These three groups consistently accounted for

---

9 These differences would be even more pronounced if changes in transit fares were compared with growth in other consumer prices in large urban areas which rose approximately 25% faster than the implicit price deflator for Gross National Product (US Council of Economic Advisers, 1984).
### TABLE 4

**Changes in Transit Fare Structures and Revenue per Passenger for Transit Systems in 26 Large US Urban Areas**

<table>
<thead>
<tr>
<th>Urban Areas</th>
<th>Year</th>
<th>% of Nationwide Transit Riders</th>
<th>Average Adult Fare(^a)</th>
<th>Average Revenue per Passenger(^a)</th>
<th>Number of Cities with Zone Charges</th>
<th>Free Transfers</th>
<th>Peak Fares</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 cities with bus and rail rapid transit(^b)</td>
<td>1974</td>
<td>49</td>
<td>0.53</td>
<td>0.52</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1981</td>
<td>48</td>
<td>0.56</td>
<td>0.50</td>
<td>3</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>20 cities with bus transit only(^c)</td>
<td>1974</td>
<td>17</td>
<td>0.56</td>
<td>0.49</td>
<td>19</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1981</td>
<td>21</td>
<td>0.51</td>
<td>0.40</td>
<td>14</td>
<td>12</td>
<td>1</td>
</tr>
</tbody>
</table>

\(^a\) Adjusted to 1981 dollars, using change in the implicit price deflator for Gross National Product.

\(^b\) New York, Chicago, Philadelphia, San Francisco-Oakland, Boston and Cleveland.

\(^c\) Los Angeles, Detroit, St. Louis, Baltimore, Houston, Minneapolis-St. Paul, Dallas, Seattle, Milwaukee, Atlanta, Cincinnati, San Diego, Buffalo, Miami, Kansas City, Denver, Indianapolis, San Jose, Portland (Oregon) and Washington, D.C.

Sources: Gomez-Ibanez (1975); APTA (1976, 1981c); USDOT (1983).

20–30% of total ridership in these cities even before they were offered reduced fares (APTA, 1976), so the proliferation of discounts was clearly one reason why actual fare revenue per passenger failed to keep pace with increases in basic fare levels. At the same time, many transit systems offered weekly or monthly passes that entitled their holders to ride free or at a substantial discount from the basic adult fare, yet these passes were commonly priced at or below the equivalent of one round trip per weekday of the period for which they were valid (APTA, 1981c).\(^10\)

Another important source of the growing gap between transit operating expenses and fare revenue was the widespread elimination of higher fares for particularly costly forms of transit service: longer journeys, travel during morning

---

\(^10\) The few statistics available on how frequently purchasers use their passes suggest that it far exceeds one round trip per weekday, thus substantially reducing the effective fare per trip. For example, the Massachusetts Bay Transportation Authority estimates that monthly passes are used for 62 one-way rides on average (see Boston Metropolitan Area Central Transportation Planning Staff, 1982).
and evening rush hours, and trips entailing transfers. Table 4 illustrates how extensive these changes in fare structures were: in 1974 transit operators in virtually all these 26 cities imposed zone fares as a means of charging higher prices for longer trips, but several abandoned them by 1981. Requiring passengers who transfer from one route or mode to another during a single trip to pay an extra charge, or even another full fare in addition to the base fare, was once another common mechanism for charging higher fares to passengers with complex travel demands, yet Table 4 shows that free transfers became much more usual among the nation's largest transit systems during the 1970s. Finally, the Table also reports that most of the few examples of peak fare surcharges — probably the most important mechanism for charging higher prices for trips that are particularly costly to accommodate, since vehicle and labour requirements are determined by the level of peak hour ridership — were eliminated from large US transit systems between 1974 and 1981.

HOW UNUSUAL HAS THE US EXPERIENCE BEEN?

One obvious question raised by these developments in operating expenses, service policies, and fare structures is whether they are the product of circumstances peculiar to the United States, or simply reflect inevitable broader developments in the cost structure and demand for transit in urbanised nations. This question has important implications for the likely future trend of transit deficits, the design and funding of government subsidy policies to aid individual nations' transit industries, and the formulation of transport policies in urbanised nations of the world. The limited available evidence indicates that, though most of the developments reported in previous sections have been particularly damaging to transit finances in the United States, they are evident to some degree in a broad sample of transit systems operating in other nations. Table 5 reports changes over a recent ten year period in summary financial and operating statistics for 75 US bus transit operators, as well as for a sample of approximately 50 European and Asian urban bus transit systems. It is difficult to ensure comparability of these data between the two samples (especially since financial data must be adjusted to reflect changing price levels and exchange rates); but some inferences can be drawn.

As the Table indicates, real operating expenditure per unit of bus service rose by nearly 40% among US transit firms over this period, primarily (as reported in section 3) in response to rapid increases in wage rates and fuel prices coupled with a decline in labour productivity. Cost escalation among European and Asian bus operators was only about half as rapid during this period, despite closely comparable growth in real wage rates among their employees. Apparently the difference is largely due to variation in labour productivity records between the two samples: Table 5 shows that a rough index of labour utilisation, vehicle-kilometres of service produced per labour-hour, fell by about 17% in the US sample, but rose by nearly 15% in the sample of European and Asian operators. Though it is difficult to judge in the absence of more complete evidence, some of this difference is probably due to the success of US transit workers in negotiating the
TABLE 5

Changes in Factors affecting Deficits for US and Other Bus Transit Operators

<table>
<thead>
<tr>
<th>Variable</th>
<th>US Bus Transit Operators&lt;sup&gt;a&lt;/sup&gt;</th>
<th>European and Asian Bus Transit Operators&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1971</td>
<td>1981</td>
</tr>
<tr>
<td>Operating expense per vehicle-km</td>
<td>$1.01</td>
<td>$1.41</td>
</tr>
<tr>
<td>Compensation per labour-hour</td>
<td>$8.69</td>
<td>$11.47</td>
</tr>
<tr>
<td>Vehicle-km per labour-hour</td>
<td>4.05</td>
<td>3.35</td>
</tr>
<tr>
<td>Annual vehicle-km per million persons served</td>
<td>25.9</td>
<td>26.5</td>
</tr>
<tr>
<td>Passengers carried per vehicle-km of service</td>
<td>1.8</td>
<td>2.0</td>
</tr>
<tr>
<td>Average fare revenue per passenger carried</td>
<td>$0.47</td>
<td>$0.28</td>
</tr>
<tr>
<td>Fare revenue as % of operating expenses</td>
<td>82%</td>
<td>39%</td>
</tr>
<tr>
<td>Operating deficit per passenger carried</td>
<td>$0.09</td>
<td>$0.43</td>
</tr>
</tbody>
</table>


protective but productivity-reducing work rules already discussed. Part of the difference may also be explained by the replacement of on-board conductors by automated fareboxes among transit systems operating in some European and Asian cities. (Strictly speaking, this represents a substitution of capital for labour rather than an improvement in productivity, yet it would still appear as an improvement in the index of labour utilisation.)

It is interesting to see from Table 5 that one measure of the per capita level of bus transit service, annual kilometres of bus service per million persons served, remained strikingly similar between the two samples over the decade. Of course,
the patterns of transit service within individual urban areas — route layouts, service frequencies, relative peak and off-peak service levels and so forth — may vary widely between cities having the same value of this index. Nevertheless, the close comparability between the samples both in average per capita service level and in its change over the decade suggests surprising similarity in overall level of service, as well as in the tendency to expand service roughly in proportion to increases in urban population.

In contrast, Table 5 indicates that the frequency of passenger boardings is sharply higher among the sample of European and Asian bus systems than among their US counterparts, though some slight improvement in this index was evident in both groups of cities over the decade. This pronounced difference no doubt partly reflects differences between the two samples in factors affecting auto ownership and use, such as personal incomes and the capacity of streets and highways, as well as differences in the spatial distribution of employment and population, patterns of transit service, and fare structures. Whatever its exact source, the maintenance in US cities of service levels comparable to those in areas where transit is much more intensively used was apparently not a major reason for the more rapid increase in transit deficits in the US, since roughly comparable changes took place in these factors in Asian and European cities.

Table 5 also indicates that US transit operators used some of the rapidly expanding government subsidies to reduce real fare levels much more sharply than was typical in the sample of European and Asian nations: while real fare levels fell by about 14% among the latter, US operators reduced average fares by more than 40% over the decade. Nevertheless, Table 5 reports that the net effect of these differences in cost escalation, service and utilisation trends, and fare changes on overall financial performance was not entirely dissimilar between the US and international samples. Farebox coverage of operating expenses fell from about 82% to 39% for US bus systems, and from 80% to 59% for the Asian and European sample; as a result, real subsidy levels per passenger nearly quintupled in the US, and nearly tripled in the international sample of cities. Thus it appears that at least one of the basic factors producing the sharp rise in US transit operators' expenses (rising wages), and the mix of purposes for which they used the remaining subsidy funds, were similar to developments in other cities around the world.

**IMPLICATIONS FOR TRANSIT SUBSIDY POLICY**

The most surprising finding from this analysis is that the main source of the US transit industry's historical contraction and financial deterioration, the continuing decline in demand for urban transit, absorbed so little of the rapid escalation in government subsidies for transit. Instead, rapidly rising expenses per unit of output, expanded service levels, and reduction and simplification of fares together absorbed nearly all the growing pool of operating assistance that governments offered in their attempts to reverse the extended postwar decline in transit ridership. Some factors contributing to rapid cost escalation — part of the growth in wage rates, perhaps some of the decline in labour productivity, and certainly the sharp increase in energy costs — were outside the control of the transit industry.
RISING DEFICITS AND THE USES OF TRANSIT SUBSIDIES

D. H. Pickrell

itself. Yet the very rapid growth in deficits over the past decade apparently owes as much to the policies followed by both government transport agencies and transit operators as it does to the continuing changes in urban development that fostered the industry’s earlier, more gradual decline. A further implication is that government subsidy programmes can become substantially more effective if transit operators can gain some control over operating costs, adapt their service offerings to changing patterns of demand for urban travel, and restructure fares to recognize variation in the costs of providing different services.

A revision of the mechanisms for funding and distributing government subsidies could also contribute to achieving these results, since the earlier analysis adds slightly to the accumulating evidence that growth in the availability of operating assistance may itself be a cause of the escalation in costs and deficits (Barnum and Gleason, 1979; Bly et al., 1981; and Pucher, 1982). The earmarking of specific tax sources for automatic distribution as transit assistance often permanently exempts subsidies from the fiscal scrutiny normally applied to periodic budget appropriations. Similarly, state and federal transit assistance programmes often distribute operating grants according to formulae that fail to take into account the financial or operating performance of their recipients; these need to be revised to re-establish incentives for transit operators to control operating costs and tailor their services to changing patterns of demand for travel.

REFERENCES

American Public Transit Association (1981c): “Transit Fare Summary: Fare Structures in Effect on February 1, 1981”.


Boston Metropolitan Area Central Transportation Planning Staff (1982): Environmental and Socioeconomic Impact Report of the Fare Increase on the Massachusetts Bay Transportation Authority.


297


298