The Rationalization of Urban Transit: Toward a Benefit Incidence Analysis

*by Greg Mason**

1. INTRODUCTION

THIS PAPER INTRODUCES benefit incidence analysis as applied to urban transit rationalization. The past decade has witnessed an increasing number of urban transit properties requiring extensive operating assistance from senior levels of government. Coupled with an era of fiscal restraint, many transit operators are seeking ways to realign their services to increase revenues, reduce costs or provide increased service within a constrained deficit level.

The first section of the paper reviews the nature of transit deficits and evaluates some of the commonly voiced defences of increased operating subsidies. The next section evaluates the concept of transit rationalization. The essence of the argument is that rationalization in transit must proceed by a recognition that transit service is in fact a heterogeneous commodity and that realignment requires the elimination of some routes and the augmentation of others. The third section of the paper reviews benefit incidence methodology and the key concept of a Lindahl tax price elasticity to allocate income equivalent benefits to various income groups. The fourth section of the paper applies the benefit incidence methodology to a simple numerical example to illustrate how it may be applied to transit.

2. URBAN TRANSIT IN CANADA: RECENT EXPERIENCE

As shown below in figure 1, urban transit in Canada was generally profitable until 1970–71. Of course, some properties have reported operating deficits as early as 1960, but the fiscal crisis in transit is a generally acknowledged fact of the seventies.

Coincident with accelerating deficit levels have been studies to both explain and justify the required operating subsidies. There are four widely accepted explanations for the losses presently incurred by urban transit.

1. Simple measures of labour productivity indicate that when compared to manufacturing and other industrial averages, urban transit has lagged far behind. It is alleged by some that the changing form of labour relations, and especially labour contracts which preclude the use of part-time employees which allows the property to meet peak hour loads, is a major factor in accelerating costs. The case is strengthened by the observation that wage costs form about 80% of the total costs in urban transit.

2. Land-use and subdivision plans which prevent the efficient delivery of transit to a growing porportion of the urban population. The result is that many properties have increased the total vehicle miles delivered by expanding their low-density suburban feeder routes. The modern suburb has also created an entire class of "captive" car users, for whom transit is an expensive alternative in terms of time and convenience. Not only do these suburban routes fail to attract revenue passengers, they typically are very much more expensive than more centrally routed transit services.

3. In the seventies, public policy dictated that fares be constrained to assist in the redistribution of income. Any fare increases brought significant political pressure upon the operator as selected groups were identified as important losers in any fare adjustment. On average fares have risen by 50–70%, while costs have accelerated by at least 200%.

4. Finally, many researchers have called attention to the significant subsidy government provides private transportation users. Not only is it alleged that car users fail to pay the full costs of the infrastructure provided by government (roads, bridges, traffic law enforcement) but they also generate important external effects (pollution and congestion). This excessive subsidy of the private auto has made it impossible for transit to significantly increase its ridership in the last thirty years.

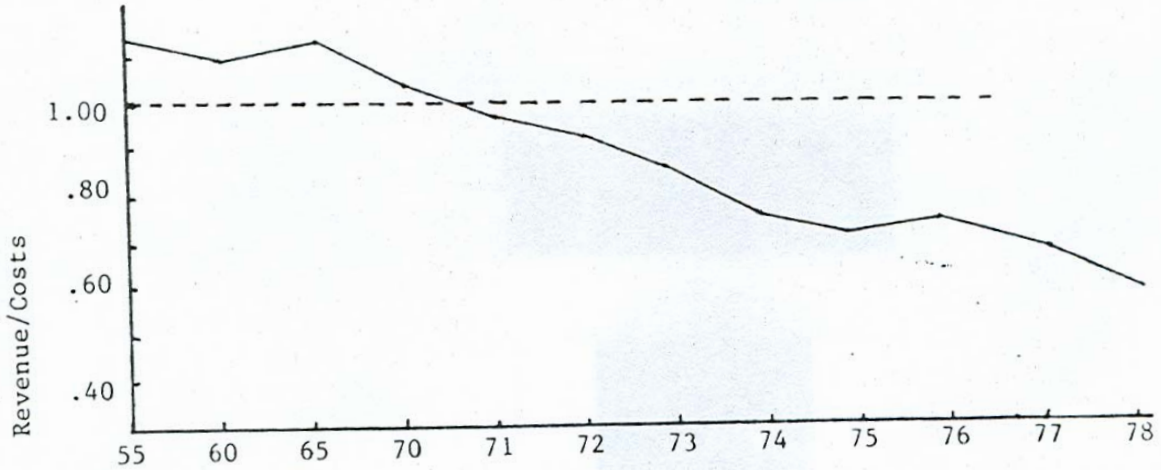
Of all the reasons for the increased deficits, certainly the failure to maintain fares with inflation must rate as the most important. The fact that declining fares (real) have failed to attract patronage means that the other explanations advanced also have had significant roles in creating deficits.

Along with the deficits have come justifications of the operating subsidies. As long as transit made profits few felt any compunction to argue that is should be a tool of social policy. Since 1970 three main arguments have been advanced.

1. Transit is alleged to be a decreasing cost industry. If efficient (marginal cost) pricing is to be adopted, then inevitably the industry will record losses. The divergence between margin-

**University of Manitoba.*

REVENUE/COST RATIOS FOR URBAN TRANSIT IN CANADA



Source: CUTA Fact Book, 1978

FIGURE 1

al and average cost will be most significant for those properties which have not fully exploited the inherent economies of scale in the service.

2. Since the second world war, it is alleged that planners have systematically favoured the car and according transportation and land—use policy has misallocated resources. By under-pricing transit, these previously wrong decisions may be reversed and the city be made more "compact" and efficient.

3. Transit is alleged to be an effective means of redistributing income. Therefore, subsidies to transit are an important social policy tool.

Each of these propositions has a certain amount of credibility, however on the whole they are overstated.

First, the argument that efficient pricing will require less than average cost fares ignores the theorem of the second best which argues that global efficiency requires that if one aspect of a set of services is "under-priced" (cars), then it does not make sense to under-price other aspects of the service. More importantly, the existence of economies of scale in transit is not well empirically defined. Most likely evidence of decreasing costs is very sensitive to the output unit chosen. Furthermore, what might appear as a declining costs for the entire range of services provided by an operator could well disappear when costing is done on a route by route basis. If economies of scale are proved to be a general feature of transit services, then efficient pricing will dictate losses and the state has a choice of subsidising the service or creating a regulated monopoly with the inevitable service constraints.[1]

The second allegation that transit must be underpriced to correct for land-use and transportation miscues is very weak. There is little empirical evidence that the long run process of land-use planning is sensitive to a low fare policy. Certainly, the cross price elasticity of demand for car use (with respect to transit fares) is very low.

Finally, the arguments that transit is a useful vehicle for income redistribution fails to receive very strong support from studies of transit patronage. Most researchers, find that the income distribution of transit users corresponds quite closely to the income distribution in general. Some even allege that the provision of transit benefits is slightly progressive (Frankena 1978). The fact that proposals for fare increases are met with stiff opposition on the grounds of supposed income distribution problems does suggest that transit planner must be very circumspect about fare policies to alleviate deficits.

The existing literature suggests that theoretical justifications for increased operating subsidies are relatively weak. Coupled with fiscal restraint, transit planners have little choice but to rationalize transit service.

3. TRANSIT RATIONALIZATION

The term rationalization has come to mean cutback and restraint, yet in reality it means a reallocation of resources away from unproductive activities to areas which yield higher private and social returns. The literature in transit rationalization usually analyses the cost structure of transit, indentifies areas which are important and growing and then suggests ways of taming these areas of increase. For example several analysts have pointed to the exsiting management-labour environment which has produced serious conflict; strikes not only eliminate revenues for the duration of

the strike, but can seriously impair ridership for years to come. They suggest that a general revision 6of labour contracts and the management atmosphere is required to produce a disciplined cost structure. Others point to articulated vehicles, improved fuel efficiencies, network optimizing procedures and promotion as devices which must be used to reduce operating deficits.

As valuable as all the above procedures are, none examines what is the fundamental point. To what extent ought transit be collectively provided? Furthermore, should rationalization not be a reallocation of transit services away from higher income groups toward lower income groups, away from choice transit users toward captive transit and auto users? The essence of the transit rationalization process is the selected elimination of some routes and the augmentation of other aspects of the service. Rather than seeking system wide economies through revisions to present labour contracts, a process which has a very high probablity of engendering bitter disputes, and further erosion of ridership, government should view transit from the perspective of a benevolent monopolist, who does not seek to extract all the consumer surplus and maximize profits, but merely wishes to breakeven from farebox revenues or even to maintain deficit growth in real terms.

An interesting experiment is reported by Garbade and Soss (1976) who approach the problem of deficit minimizing by creating an heursitic programming problem based upon a simultaneous equation model of urban transit in New York. The objective is to simulate deficits over a planning horizon of seven years under a variety of projections on the exogneous variables of the system such as ridership and cost inflation. They are seeking a set of fares which will constrain deficits to a pre-set inflation in costs (6%). Once the revenue function has been estimated based upon a demand function estimated from time series data, the mathematical program

Minimize: Average Fare
Subject to: Average deficits = S

is solved heuristically for the set of fares which provide the optimum. They discover that a fare increase of 50% will be required over the period 1974-1980 merely to hold costs constant in real terms.

The approach to rationalization they employ is to analyse what fare increases will be required to hold deficits constant, and this is one part of the question. Alternately one can view transit, not as a homogeneous service offered to each urban voter (consumer) equally, but a heterogeneous product differentiated at least by route and time of day. The process of transit rationalization requires that it be viewed as many services, each of which must be subjected to an individual benefit-cost evaluation.

Proposed here is a benefit incidence methodology which forms one part of the transit service evaluation. Given that transit operators in Canada are unlikely to be able to raise fares to cover increases most transit operations must rationalize operations with a minimum level of perceived social welfare as an important constraint. By disaggregation of the transit service and application of a benefit imputation to each individual service, the operator can impute to various income groups income equivalents for the service.

4. BENEFIT INCIDENCE METHODOLOGY

Benefit incidence analysis imputes income equivalents to each income group, resulting from the collective provision of goods. Important questions involved in the approach are

a. Valuing the income redistributive impacts of public goods;
b. Evaluating "mixed" goods in terms of the proportion of "publicness" contained in the good or service.

1. Transit as a Public Good

The definition of public good has undergone frequent refinement in the last few decades. Two key attributes have now been isolated—the concept of externality and the concept of appropriation of property rights in exchange.

The first condition for a good or service is that it must have significant externalities in production. At the polar extreme, the good or service must be such that the consumption by any voter does not impair the consumption enjoyed by any other. For example, aside from infrequent anomalies, radio and television are public goods in this sense, for the power of the signal is not diminished by the numbers who consume the service.

Second, public goods are difficult for the private market to organize in exchange. The costs of transition are too high or the technical attributes such that private individuals are unable to provide the product or service and charge sufficient to cover costs. These enforced losses incurred by the private provision of a service at times is taken to be evidence that a service is inherently a public good.

Other consequences are also due to publicness. Often a service is provided uniformly, or with some minimum amount, simply because failure to consume the service may impair health. The most common examples are sanitation and fluoridation of water supplies.

In addition, Musgrave (1959) has identified a class of services termed "merit" goods which are provided because it is believed that certain intergenerational benefits accrue to society. Education and public housing are among the common examples here.

Transit is often identified as having significant aspects of publicness and therefore should be provided by the state. But some important qualifications should be noted. First, some aspects of mass transit have and are being organized by private individuals. Van pooling is an important counter example to the usual argument that mass transit must be collectively provided. Second, it is clear that transit does have important externalities in the alleviation of congestion faced by car users, but the externalities obtained by the transit user have limits. Clearly, service attributes such as headways, speed and comfort are affected by the number of users. Thirdly, to provide the extent of service presently installed by most transit operators implies that no private organization could charge to cover all benefits, but most likely for most properties there is a set of services which could be profitably organized by the private market. There is nothing inherent in transit service which precludes any recapture of the property rights exchanged through the price system.

Most analysts now agree with Samuelson (1969) who maintains that there are a few polar public goods such as defence and possibly public health which are inherently public and must be collectively provided. The overwhelming majority of goods and services in our economy are mixed goods, even those which are exclusively provided by the private sector. An important aspect of benefit incidence analysis of mixed goods is the measurement of the degree of publicness possessed by a good which is collectively provided.

2. Measurement of Publicness

Most empirical attempts to measurement publicness of collectively provided goods exploit the externalities possessed by the vast majority of goods and services. The underlying assumption is that as the number of consumers increase, goods which become progressively restricted in supply to the average voter exhibit the attributes of a private good, while commodities which remain available despite the number of consumers have a high degree of public content.

The actual measurement of publicness reported by researchers such as Borcherding and Deacon (1972) and Deacon (1978, 79) exploits cross sectional data and a basic crowding formula given by

$$Q = N^{-a} X$$

where Q is the consumption of a service by the median voter, N the total number of people consuming the service and X the expenditure by a municipality on the service. The parameter "a" is the elasticity of consumption with respect to a change in the number of people who consume the good and varies between 0 and 1. A value of 0 indicates that a service is purely public while a value of 1 indicates a purely private good. For the most part, the estimates of "a" lie between .8 and 1, indicating the collectively provided goods, at least those provided by municipalities in the United States, have a high degree of privateness. Unfortunately, because transit expenditures are also funded by senior levels of government, the estimates of transit service publicness has not yet been accomplished.

The measurement of the elasticity of externality, another way of looking at the parameter "a," enables one to predict how increasing the number of consumers reduces the benefits received by typical voters. Coupled with what is termed a Lindahl pricing rule and a procedure of assigning benefits to the public and private components of transit, enables one to estimate the benefits received from transit service received by various income groups.

3. Benefit Incidence of Mixed Goods

If a good is completely private, valuing the benefits received by the consumer is relatively straightforward. Assuming that the service is provided under conditions of constant costs (not realistic for most goods), then the benefits obtained by the average voter are merely the total cost of providing the good divided by the number of voters. For a given income class, one use the cost of providing that group with the service or the number of consumers times the uniform price of providing that good.

Aside from the problems associated with free riders which makes the assignment of quantities consumed by a given income class somewhat problematic, many services are provided in uniform quantity. Sanitation is generally provided to all urban households in uniform quantity. If we assume that otherwise the demand for sanitation services is a linear function of housing expenditures and that these generally have an income elasticity of around 1, then it appears that this service has a progressive impact upon income distribution with lower income households obtaining more than they would otherwise consume and higher income households less. Higher income households are made worse-off by collectivization.

For public goods the valuing of income equivalents is more problematic. Historically the approach has been to argue that the income equivalents for a given good are allocated in inverse proportion to the marginal utility of

income. This conclusion is arrived at by assuming that each voter has a two part utility function (separable) consisting of a group of private goods and one public good. If the public good is provided at minimum cost (constant costs) the total benefits to the community equal the total costs of providing the good. The marginal rate of substitution between public and private (or more precisely the income spent on private goods) will be in direct proportion to the ratios of the marginal utilities of each type of good. Separability implies that the enjoyment of the public goods is unaffected by the amount of private goods consumed and vice-versa. Thus, since everyone consumes the same amount of a public good by definition, it can be shown that benefits in terms of income equivalents are inversely proportional to the marginal utility of income.

This approach to benefit incidence analysis is somewhat obscure since despite the possibility of measuring marginal utilities of income (Sato, 1972), most transit practitioners would regard it as a rather emphemeral reed upon which to argue for rationalization of certain routes. Fortunately, it is possible to use the more common notions of fare and income elasticities.

The first step in the explanation revolves around the Lindahl tax price equilibrium which argues that the tax price levied for pure public goods should be responsive to the price and income elasticities of demand. Given that pure public goods are such that the consumption by any individual leaves unaffected the consumption of others, the Lindahl pricing formula indicates that the tax price levied should be given by the formula

$$\gamma = \frac{-\beta}{\eta}$$

where β is the income elasticity of demand, η the price elasticity of demand and γ the elasticity of the tax price with respect to income.

Dean (1980) has shown how this formulation can be used to produce a benefit imputation for publicly provided services. The steps in the analysis are twofold;

1. An estimate of the degree of publicness of a good. By the Lindahl pricing rule, the parginal cost of any collectively provided service can be decomposed into a private and public prices for the mixed good. The public price is interpreted as the sum of Lindahl tax prices for the public component. Once the price and income elasticities of a service are known, the Lindahl tax price structure is known and using revealed preference one can allocate benefits across income classes.

2. The second empirical estimate required is the demand relationship which yields the price and income elasticities of demand required for the Lindahl tax price elasticity.

5. BENEFIT INCIDENCE APPLIED TO TRANSIT

Assuming that transit is viewed as a heterogeneous collection of services, differentiated by routes and time of day, and that rationalization is to proceed by realigning routes in accordance with some measure of income redistribution, the most important step in the procedure is to develop the various empirical estimates. In particular the transit operator must gauge the mixture of the transit services offered and the price and income elasticities which characterize aspects of the service.

For the most part, there seems little prospect of analyzing each route for degree of publicness, especially since this estimate has not been developed for transit in general. Since most collectively provided services appear to lie in the range of .75–1.00, its seems reasonable to suppose that the majority of the benefits are private.

The measurement of price (fare) and income elasticities are less problematic, as most operators have developed these estimates for their operations as a whole. Few transit operations fail to collect detailed route operating statistics, and the response of various routes to changes in fares and average wage rates of riders should not be a difficult figure to obtain. In some properties, where labour disputes have disrupted service for some period of time, the estimates of price and income elasticity by route can be corroborated by examining how routes recover ridership after a prolonged interruption in service.

Benefit incidence is a politically malleable analysis. What it permits is the translation of different services into income equivalents. The transit operator can "fine-tune" the service by realigning routes toward those services which confer benefits on the lower income groups, or toward other income groups. There is no presumption as to social objective in the methods, except that contained in the neo-classical assumptions underlying the concept of a Lindahl tax price and the notion of constant costs of transit operation.

A Numerical Example

To further illustrate how the methods may be applied a simplified numerical example is presented to simulate the calculations **hence the measure of publicness and the required elasticities have been obtained.** Consider an operator who provides three routes with total costs of service of $100,000, $150,000 and $200,000 respectively. Assume also that we wish to consider the benefits received by two individuals of $5000 and

$10,000 income and that each route has the same number of riders (10). Finally assume that we have established that the degree of publicness is .75. Table 1 summarizes the basic data.

Note that the column totals of 7,500, 11,250 and 15,000 are the per user private costs for each route. The differences between these totals and the total per user cost of the service on each route is the public benefit which is allocated according to the Lindahl tax price formula.

From the route by route studies, the tax price elasticity may be established as in Table 2.

For routes 1, 2 and 3 the sums of 2,500, 3,750 and 5,000 must be allocated for each user according to income. A Lindahl tax price elasticity of 1.0 implies that if income doubles, the benefits conferred are doubled, but if the tax price elasticity is 2.0 the benefits increase threefold for a doubling in income. This is because the tax price increases by 200% for a 100% increase in income. Similarly, if the tax price elasticity is 4.0, as income doubles, benefits rise by a factor of 5. The following table summarizes the distribution of public benefits by income group.

Now one merely aggregates the private benefits calculated in Table 1 and the public benefits calculated in Table 3 to arrive at a distribution of total benefits by type of service. Of course, the disaggreation can be pursued by considering time of day differentiation, but this implies that the transit operator is willing to engage in an extensive market survey of the users. It remains a mystery why many jurisdictions do not, as a matter of routine survey the users on a regular basis. Certainly, private operators of such a complex service would not fail to analyze the operation comprehensively. Many of the elasticities required to do a proper analysis of the benefits conferred by a service require that the demand for the service be carefully monitored. Thus a pre-requisite for this, and in fact any approach to transit rationalization, is a comprehensive market survey.

6. CONCLUSION

Recent advances in assigning income equivalents to mixed goods which are collectively provided offer hope to transit operators who wish to calculate how social welfare is augmented by the service they offer. It is assumed that most operators will not be permitted to behave like monopolists, and face continued fare constraint. Therefore, rationalization inevitably implies the realignment of services; some aspects must be cutback while other aspects can be augmented.

Benefit incidence analysis permits the transit operator to calculate income equivalent benefits for each service, and on the basis of these, to realign services according to some political objective. Although benefit incidence analysis cannot by itself indicate the appropriate rationalization of services, it does offer a useful way of regarding the services provided and actually requires comparatively little empirical sophistication.

FOOTNOTES

1. The theory of the second best has failed to receive widespread endorsement as a useful applied policy prescription. In general, the argument that transit will incur losses if marginal cost pricing is used is valid as long a decreasing costs do in fact exist. For larger operators, the rate of decline in costs as output increases may be relatively slight suggesting that marginal costs may not be very divergent from average costs. Also, it should be recognized that the actual measurement of economies of scale is very sensitive to the output unit.
2. In principle, transit services could be organized by a monopolist. Many defenders of transit fare increases note that the low fare elasticity implies that fare increases will bring increased revenues without jeopardizing transit use. It is dangerous to extent elasticity estimates at a point to the entire demand curve, for at some fare level, one can expect the elasticity of demand to increase. Furthermore, once a former transit user has found other means of transport (bought a car), the short-run fare elasticity is zero. There probably exist some set of routes and fares with-in all properties which could be organized by a monopolist, although on average, service would decline and fare levels increase.
3. If the public service is provided at constant cost and if the marginal cost is normalized to equal 1, then the total cost of providing service (Marginal cost × quantity) will equal the benefits (quantity times the sum of individual marginal rates of substitution between income on private goods and the public good) or,

$$\left(\sum_i \mathrm{MRS}^i\right)(G) = \mathrm{MC}_G \times G$$

the ratios of the marginal utilities of public goods and income equal the marginal rates of substitution between public and private goods. Since everyone must consume equal quantities of a pure public good, the marginal utility of the public good (MU_G) is constant which implies

TABLE 1

User	Income ($)	Private Good Benefits (.75 × cost) by Route($)		
		1	2	3
A	5,000	3,750	5,625	7,500
B	10,000	3,750	5,625	7,500
		7,500	11,250	15,000

TABLE 2

Route	Income (β)	Elasticity Price (η)	Tax Price (γ)
1	1.0	- .25	4.0
2	1.0	- .50	2.0
3	1.0	-1.00	1.0

$$\sum_i \frac{K}{MU_Y^i} = MC_G$$

and leads to the allocation rule that benefits are received in inverse proportion to the marginal utility of income (recall that benefits equal the cost of providing the service).

BIBLIOGRAPHY

1. Borcherding, T.E. and R.T. Deacon, "The demand for the services of non-Federal government," *American Economic Review,* Vol. 62, pp. 891-901.
2. Deacon, R.T. (1978), "A demand model for the local public sector," *Review of Economics and Statistics,* vol 60, pp. 184-192.
3. _____ (1979), "The expenditure effects of alternate public supply institutions," forthcoming, *Public Choice.*
4. Dean, J. (1980), "Benefit incidence methodology for mixed goods," forthcoming, *Public Finance Quarterly.*
5. Frankena, M. (1973), "Income distributional effects of urban transit subsidies," *Journal of Transport Economics and Policy,* vol. 7 pp. 215-230.
6. Garbade, K.D. and N.M. Soss (1976) "Fare policies for mass transit deficit control: analysis by optimization," *Transportation Research,* vol 10, pp. 237-247.
7. Musgrave, R. (1959) *The Theory of Public Finance,* New York, McGraw-Hill
8. Samuleson, P. (1955) "The pure theory of public expenditure" *Review of Economics and Statistics* vol
9. Sato, K. (1972) "Additive utility functions with double-log consumer demand functions" *Journal of Political Economy,* vol. 80, pp 102-124.

TABLE 4

Income($)	Benefits by Route($)		
	1	2	3
5,000	500	937.5	1,650
10,000	2,000	2,812.5	3,350
	2,500	3,750	5,000