THE EFFICIENCY OF PUBLIC TRANSPORT OBJECTIVES AND SUBSIDY FORMULAS

By Mark W. Frankena*

Recent articles in this Journal (Glaister and Collings, 1978; Nash, 1978) have analysed the efficiency of passenger-mile maximisation as an objective for a public transport firm. Glaister and Collings considered a firm which operates two or more services subject to an overall budget constraint; Nash considered related issues for a firm which operates a single homogeneous service. But the conditions under which maximisation of passenger miles would be inefficient, and the nature of the inefficiency, have not been clarified. The present article fills this gap and extends the analysis to issues raised by the practice of subsidising public transport firms. We begin by analysing the conditions under which maximisation of ridership by a public transport firm which operates a single service would be economically efficient. We then consider whether a subsidy to such a firm would increase efficiency, and which subsidy formula would be best. Because little additional effort is required, we analyse the objective of vehicle-mile maximisation as well as the objective of ridership maximisation.

The model used here will be interpreted as an aggregate model of an urban public transport system, but it could obviously be applied to non-urban public transport systems as well. The model determines the aggregate city-wide levels of public transport ridership and service, and the fare. The model does not consider temporal or intraurban spatial variations in the fare, service, or demand for rides, and does not allow fare discrimination between riders. The demand for rides is determined by the fare and the number of vehicle miles of service, and these are chosen by a non-profit monopolistic transport firm which maximises an objective function subject to a budget constraint. The budget constraint requires that the transport firm’s costs must not exceed its revenues from fares and any subsidies.

The analysis of the normative properties of a model of a public transport firm which operates a single service is of interest for two reasons. First, this model has been used extensively in the literature on urban public transport (Frankena, 1978, 1980, 1981; Nash, 1978, Nelson, 1972); but attention has been concentrated almost exclusively on the positive rather than the normative properties of the model. Second, the model of a non-profit monopolistic transport firm which selects a combination of fare, vehicle miles of service, and ridership is similar to models which have been applied to both

* Associate Professor of Economics at the University of Western Ontario, London, Canada.
profit-maximising and non-profit monopolies which select a combination of price (which corresponds to the fare in the transport model), product quality (vehicle miles of service), and product quantity (ridership) (see Hansmann, 1981; Newhouse, 1970; Sheshinski, 1976; Spence, 1975).  

1. MAXIMISATION OF VEHICLE MILES AND OF RIDERSHIP

Assume that the inverse demand function facing the public transport firm is $F = F(R, M)$ and that $F_R < 0$ and $F_M > 0$, where $F$ is the fare per ride, $R$ is number of rides, $M$ is vehicle miles of service, and $F_R$ and $F_M$ are partial derivatives. (Alternatively, $F$ could be interpreted as the fare per passenger mile and $R$ as the number of passenger miles, as in Nash (1978). Although models developed in terms of number of rides and of number of passenger miles are different empirically if rides are not of uniform length, they are formally identical.) Also assume that the total production cost function is $C = C(R, M)$ and that $C_R \geq 0$ and $C_M > 0$. Given these assumptions, in the absence of subsidies the transport firm chooses among combinations of fare, vehicle miles, and ridership which satisfy the inverse demand function and the budget constraint $FR - C = 0$. The combination of $F, M$ and $R$ which the transport firm selects depends on the firm's objective function, which is assumed to take the form $\max U = U(F, M, R)$ where $U_F \leq 0, U_M \geq 0$, and $U_R \geq 0$. $U = R$ corresponds to ridership maximisation and $U = M$ corresponds to vehicle-mile maximisation.

We are interested to see how the combination of $F, M$, and $R$ selected by a transport firm which maximises ridership or vehicle miles would compare with the "second-best efficient" combination, that is, the combination which maximises the net social benefit from transport service under the firm's budget constraint. The net social benefit is measured by the total amount of money that transit users would be willing to pay for their rides, minus the total production cost for those rides:\footnote{In fact, the non-profit model can be used in a wide range of monopolised industries, from public transport to the performing arts. For example, a number of the normative results derived in the original version of this paper (Frankena, 1980, Appendix 8A) and the positive results derived in a companion paper (Frankena, 1980, chapter 8; Frankena, 1981) were arrived at independently by Hansmann (1981) in the context of a model of a performing arts company.}

$$\int_{0}^{R} F(Z, M) dZ - C(R, M)$$

Corresponding to each level of $M$ is a different inverse demand curve for transport.

\footnote{By comparison, previous models of this type have assumed that $C = \alpha M$, so that $C_R = 0$ (Frankena, 1978, 1981; Nash, 1978; Nelson, 1972).}

\footnote{One should use the inverse income-compensated demand curve to measure willingness to pay. This measure assumes that there are no externalities of the type that would arise if there were unpriced automobile congestion. The implication of such externalities is considered below. Finally, it ignores the costs imposed by transport riders on other transport riders and on other road users, e.g., when an extra transport rider delays a transport vehicle while boarding.}
rides, $F = F(R, \bar{M})$, in $(R, F)$ space. As $M$ is increased, the demand curve shifts outward. Figure 1 shows demand curves corresponding to four different levels of $M (M_1 < M_2 < M_3 < M_4)$, drawn on the assumptions that $F_{R_R} = 0$ and $F_{R_M} > 0$.

Also corresponding to each level of $M$ there is an average cost function $C(R, \bar{M})/R$ (not drawn in Figure 1), which would normally intersect the demand curve twice for feasible levels of $M$. Each intersection would determine a combination of $F$ and $R$ for which the transport firm would break even, given the level of $M$. In the light of our assumptions about the transport firm's objective function, we are concerned only with the intersection at the higher level of $R$ along a demand curve. We have used the letters $A, B, D, K$ to indicate the relevant points along the four demand curves in Figure 1 at which the transport firm would break even. For other levels of $M$, the break-even points would lie on the dashed line $ABDK$.

Along line $ABDK$, which is tangent to the demand curve for $M = M_4$, point $K$ would be chosen by a firm maximising vehicle miles of service, and point $D$ would be chosen by a firm maximising ridership. By contrast, a firm maximising the net social benefit would choose point $B$. At point $B$ the net social benefit is measured by the
shaded triangle \( (BGF_2) \), which is larger than the corresponding triangle for any other point along line \( ABDK \).4

Thus, in the example given in Figure 1, a transport firm which maximises either vehicle miles of service or ridership will set \( F \) and \( M \) above the second-best efficient levels, and thus will waste resources. However, the location of the second-best efficient operating policy in relation to the policy that would maximise ridership depends upon the assumptions one makes about the demand function. In Figure 1, the demand curve becomes flatter as \( M \) increases; that is, \( F_{BM} > 0 \). If the slope of the demand curve does not change as \( M \) increases and the demand curve shifts out \( (F_{BM} = 0) \), the operating policies corresponding to ridership maximisation and second-best efficiency will coincide. If the demand curve becomes steeper as \( M \) increases \( (F_{BM} < 0) \), the second-best efficient operating policy will involve a higher level of \( F \) and \( M \) than the policy which would maximise ridership, but will still involve a lower level of \( F \) and \( M \) than the policy which would maximise vehicle miles of service.

Intuitively, how the slope of the demand curve changes as \( M \) increases depends upon how inframarginal users compare with marginal users in the value they place on the increase in vehicle miles of service or the resulting savings in travel time. For example, given the level of vehicle miles of service, suppose that if the fare is high the only people who will ride buses will be low-income people who place a low value on their travel time. Suppose that if the fare is reduced the new riders will be higher-income people who place a higher value on their travel time. If the level of vehicle miles of service is now increased, the demand curve will shift up. However, the left end of the demand curve may not rise very much, because low-income (inframarginal) riders may not be willing to pay much for improved service. By contrast, the right end of the demand curve may rise a good deal, because higher-income (marginal) riders may be willing to pay more for improved service. If so, the demand curve will become flatter as the number of vehicle miles of service increases, as in Figure 1. In this case, a transport firm which maximises ridership offers more than the second-best efficient level of service (and charges more than the second-best efficient fare) because of the disproportionately large valuation placed on improved service by marginal users. This extra service is not efficient, because it costs more per rider than the average user would be willing to pay for it.

Furthermore, the information about demand behaviour which is required to determine the second-best efficient operating policy is greater than the information about demand behaviour which is required to maximise ridership. To determine the efficient policy, one must know how all inframarginal users value improved transport service or the resulting reduction in travel time, while to maximise ridership one needs information only on the relevant marginal users. This helps to explain why transport firms find it more convenient to maximise ridership than to maximise net social benefits.

---

4 To prove this, observe that at any point \((R,F)\) at which the transport firm would break-even the net social benefit \( V_f \) is determined by \( V_f = 0.5(G - F)R \). All break-even points with a consumer surplus equal to that at \( B \) therefore must lie along the line \( TT' \), which is a rectangular hyperbola when viewed with respect to an origin located at point \( G \) and axes formed by the lines \( R = 0 \) and \( F = G \). Since the loci of break-even points \( ABDK \) and \( TT' \) are tangent at \( B \), all break-even points other than \( B \) would yield lower net social benefits than \( B \).
THE EFFICIENCY OF PUBLIC TRANSPORT

Mark W. Frankena

Our basic conclusion is that maximisation of vehicle-miles is always inefficient, and that maximisation of ridership is inefficient unless one makes a rather restrictive assumption about the demand function. However, the direction of any deviation from second-best efficiency which would result from the objective of maximising ridership cannot be determined without empirical knowledge about the form of the demand function.

2. REFORMULATION USING THE SPENCE–SHEHINSKI MODEL

The preceding analysis of a non-profit monopolistic transport firm which selects a combination of fare, vehicle miles of service, and ridership is similar to the analysis carried out independently by Spence (1975) and Sheshinski (1976) for a profit-maximising monopoly which selects a combination of price, product quality, and product quantity. The only difference between the transport model analysed here and Sheshinski’s model is in the monopoly’s objective function. Vehicle miles of service in our model plays the role of product quality in Sheshinski’s model; but one could replace vehicle miles of service by a more general scalar index or vector of quality variables including in-vehicle travel time, walking and waiting time, comfort, and adherence to schedules.

In this section we reanalyse the issues discussed in Section 1, using a geometric model in \((R,M)\) space developed by Sheshinski. That model has the disadvantages that it suppresses the price variable and does not explicitly show the role of the sign of \(F_{RM}\), but it has the advantage of being able to show the effect of a subsidy on net social benefits, which is the topic of Section 3.

The locus of combinations of \(R\) and \(M\) which satisfies the transport firm’s break-even budget constraint in the absence of subsidies is given by:

\[ F(R,M)R - C(R,M) = 0 \]  
(1)

This can be regarded as an iso-profit contour defined by the firm’s profit function:

\[ \pi(R,M) = F(R,M)R - C(R,M) \]  
(2)

Consequently, in order to locate the points which would satisfy (1), it is helpful to find the combination of \(R\) and \(M\) which would maximise the transport firm’s profit. For an interior solution, the first-order conditions for profit maximisation are:

\[ \pi_R = F + F_R R - C_R = 0 \]  
(3)

\[ \pi_M = F_M M - C_M = 0 \]  
(4)

We assume that the second-order conditions hold globally. For any given value of \(M\), equation (3) determines the level of \(R\) which maximises profits. Similarly, for any given value of \(R\), equation (4) determines the level of \(M\) which maximises profits.

Since we want to analyse this model geometrically, we now want to identify the locus of points \((R,M)\) which satisfies equation (3), and similarly the locus for equation (4). To do so, we must specify the sign of \((F_M + F_{RM} R - C_{RM})\)—see Sheshinski (1976, p. 130); we assume that it is positive. This gives us the case in Sheshinski’s Figure 1. The loci for \(\pi_R = 0\) and \(\pi_M = 0\) in our Figure 2 now both slope up, and \(\pi_R = 0\) is
FIGURE 2
Efficiency of Alternative Objectives

steeper at the intersection, point N, which is the combination of R and M which maximises the transport firm’s profits.

We can now add the transport firm’s break-even contour (1) to Figure 2. Sheshinksi’s profit-maximising monopolist operates at point N, but a non-profit monopolistic transport firm would choose the point along the break-even contour which maximises its utility function $U = U(F, M, R)$. If the transport firm maximises vehicle miles of service, it will select point P, where the break-even contour and the locus $\pi_M = 0$ intersect. If the transport firm maximises ridership, it will select point Q, where the break-even contour and the locus $\pi_M = 0$ intersect. If the transport firm has an objective function characterised by what we defined in Frankena (1981) as “low
THE EFFICIENCY OF PUBLIC TRANSPORT

Mark W. Frankena

fare preference”, it will choose a point along the break-even contour right of \( \pi_R = 0 \) and below \( \pi_M = 0 \).

In order to determine the point along the break-even contour which is second-best efficient and to analyse the efficiency of maximising \( M \) or \( R \), we now wish to add iso-net social benefit contours to Figure 2.

The net social benefit of transport service is:

\[
V(R,M) = \int_0^R F(Z,M) dZ - C(R,M)
\]

The first-order conditions for maximisation of \( V \) with respect to \( R \) and \( M \) are:

\[
V_R = F - C_R = 0
\]

\[
V_M = \int_0^R F_M(Z,M) dZ - C_M = 0
\]

We assume that the second-order conditions hold globally.

We now wish to add to Figure 2 the locus of points which satisfy equation (6), and similarly the locus for equation (7). To do this, we must specify the sign of \( F_M - C_{RM} \)—see Sheshinski (1976, p. 130), and we assume that it is positive, as it would be for the model in Frankena (1981), where \( C_{RM} = 0 \). We must also specify the sign of \( F_{RM} \) which determines how the slope of the transport demand function changes as \( M \) increases. We assume \( F_{RM} > 0 \), which is the assumption made in our Figure 1. These assumptions give us the case analysed by Sheshinski in his Figure 1(a). In this case, the loci for \( V_R = 0 \) and \( V_M = 0 \) in our Figure 2 both slope up, and \( V_R = 0 \) is steeper at the intersection, point \( H \), which is the combination of \( R \) and \( M \) which maximises the net social benefit.\(^\text{5}\) Furthermore, \( V_R = 0 \) lies to the right of \( \pi_R = 0 \) and \( V_M = 0 \) lies below \( \pi_M = 0 \). We have added one iso-net social benefit contour \( V(R,M) = \bar{V} \) around point \( H \).

In order to achieve a second-best efficient allocation of resources subject to its budget constraint, the transport firm should operate at point \( E \), where it is on the highest attainable iso-net social benefit contour. Under the assumptions made in Figure 2 (notably the assumption that \( F_{RM} > 0 \)), the points that would maximise vehicle miles of service and ridership both involve an inefficiently high level of \( F \) and \( M \). The intuition behind this result was discussed in the preceding section. In the case at hand, an objective function involving a particular specification of “low fare preference” (see Frankena, 1981) could have led the transport firm to choose point \( E \). However, one could not specify what transport objective function characterised by low fare preference would achieve this result unless one knew the transport demand and cost functions.

If we changed one of the assumptions made above and assumed that \( F_{RM} < 0 \), we would obtain a new diagram similar to Figure 2 except that the locus \( \pi_M = 0 \) would lie below the locus \( V_M = 0 \) (see Sheshinski’s Figure 1(b)) rather than above it. With this change, the tangency between the break-even contour and the iso-net social benefit

\(^\text{5}\) This assumes that there are no real resource costs involved in raising revenue to finance any subsidy required to achieve the efficient allocation.
contour would lie between the points of maximum vehicle miles and maximum ridership along the break-even contour. In that case, ridership maximisation would lead to inefficiently low levels of \( P \) and \( M \). Finally, if \( F_{RM} = 0 \), the tangency would coincide with the point of maximum ridership. In this very restrictive case, ridership maximisation would be second-best efficient, given the budget constraint.

THE EFFICIENCY OF TRANSPORT SUBSIDIES

It is well known that if the objective of a public transport firm is to achieve an efficient allocation of resources, and if the transit fare is above the marginal social cost of transit rides because of economies of scale, then a lump-sum subsidy will lead to an increase in the efficiency of resource allocation (Frankena, 1980, chapter 5). Unfortunately, if the transport firm’s objective function departs from efficiency or maximisation of net social benefits, it no longer follows that a subsidy will necessarily increase efficiency, even if the fare is above the marginal social cost of public transport rides. Consequently, in practice one cannot justify a public transport subsidy on efficiency grounds without reference to empirical information about the transport firm’s objective function, demand function, and cost function.

This point is illustrated in Figure 3, which repeats Figure 2 with additional iso-profit and iso-net social benefit contours. We assume that the transport firm receives a lump-sum subsidy which shifts the firm’s constraint from the “break-even contour” to the “lump-sum subsidy contour”, which is tangent to a higher iso-net social benefit contour. If the firm maximises net social benefits, it will move from point \( E \) to point \( J \) as a result of the subsidy, and net social benefits will increase. However, if the firm maximises vehicle-miles of service it will move from \( P \) to \( S \), and if it maximises ridership it will move from \( Q \) to \( W \). In either case, given the explicit and implicit assumptions on which Figure 3 is based, net social benefits will decrease. However, under some other assumptions about the demand and cost functions, one would find that a subsidy to a transport firm which maximises ridership would increase net social benefits, even when ridership maximisation is not an efficient objective.

The net social benefits resulting from a subsidy with a given cost to taxpayers will depend not only on the objective function of the transport firm and the demand function, but also on the subsidy formula used. If the transport firm maximises net social benefits subject to its budget constraint, then a lump-sum subsidy will be most efficient. It will be more efficient than a subsidy based on vehicle miles of service or (unless \( F_{MR} = 0 \), so that ridership maximisation and efficiency coincide) a passenger subsidy.

There are, however, circumstances under which a subsidy based on vehicle miles of service, such as the cost subsidy analysed in Frankena (1981), would be more efficient than a lump-sum subsidy. For example, if the transport firm maximises ridership and if \( F_{RM} < 0 \), in the absence of subsidies the transport firm will choose a point along the budget constraint where \( F \) and \( M \) are below the second-best efficient levels. Since a subsidy based on vehicle miles of service would induce the firm to increase \( F \) and \( M \) more than would a lump-sum subsidy with the same cost to taxpayers (see Frankena, 1981), in this case a subsidy based on vehicle miles of service might produce higher net social benefits than a lump-sum subsidy.
SUMMARY AND A QUALIFICATION

The results of this paper are essentially negative. Without knowing the demand and cost functions one cannot determine whether ridership maximisation would lead to levels of fares and vehicle miles below, equal to, or above those which would be second-best efficient under the budget constraint, and one cannot determine whether a subsidy to a firm which maximises ridership would increase or reduce efficiency, even if fares exceed marginal social costs. Also, unless one has information about the transport firm's objective function, demand function, and cost function, one cannot
determine which subsidy formula would be most efficient. It follows that empirical research on these functions is important for the evaluation of subsidies.  

A final qualification is in order. The analysis in this paper has assumed that there are no relevant externalities such as unpriced road congestion. If some public transport riders were diverted from use of roads which were priced at less than marginal social cost, then the net social benefit of public transport service would be greater than that indicated by equation (5). The principal effect of this would be to shift $V_s = 0$ to the right in Figure 2. This in turn would shift the tangency of the break-even iso-profit contour and the iso-net social benefit contour in the direction of the point of maximum ridership (regardless of the sign of $F_{RM}$). However, while this would reduce the size of the efficiency loss resulting from maximisation of ridership, it would not have much effect on the qualitative problems raised.

REFERENCES


* Similar research is important for other non-profit organisations to which the basic model might apply.

76