Journal of Transport Economics and Policy, Volume 44, Part 3, September 2010, pp. 365–372

The Case for Subsidisation of Urban Public Transport and the Mohring Effect

Leonardo J. Basso and Sergio R. Jara-Díaz

Address for correspondence: Leonardo J. Basso, Department of Civil Engineering, Universidad de Chile, Casilla 228-3, Santiago, Chile (lbasso@ing.uchile.cl).

This research was partially funded by FONDECYT-Chile, Grant 1080140, and by the Millennium Institute 'Complex Engineering Systems', Grants ICM: P-05-004-F and CONICYT: FB016. We would like to thank José Muñoz for helpful comments.

Abstract

In this journal, van Reeven (2008) develops a model aimed at showing that scale economies on users' time costs would not provide a justification for public transport subsidies. He claims that a profit-maximising operator allowed to take the demand effects of its pricing into account would offer a frequency f^{π} at least as high as a welfare-maximising one f^{*}, and with no welfare losses. We show that his result depends crucially on a strong assumption of demand. Introducing a slight modification to make it more realistic, we show: (i) $f^{*} > f^{\pi}$, (ii) welfare losses emerge under profit-maximisation, (iii) subsidies are required for first-best operation. Thus, the Mohring effect is a valid argument for subsidisation.

Date of receipt of final manuscript: May 2009

1.0 Introduction

By minimising the sum of the costs of transit users (their time) and operators, Mohring (1972) and Jansson (1979) have shown that urban public transport operations should be subsidised in the first-best case. The novelty is that these subsidies appear not necessarily due to economies of scale on the operators' side but due to the existence of economies of scale on users' time costs. This result is sometimes known as the *Mohring effect*.

In a recent article in this journal, Peran van Reeven (2008) develops a model aimed at showing that the existence of economies of scale on users' time costs does not provide a justification for subsidies. The author claims that 'if an operator is allowed to take the demand effect of their pricing and frequency decisions into account' — something that does not appear in a cost minimisation framework — then a profit-maximising operator would offer a frequency f^{π} at least as high as a welfare-maximising one, f^{*}. Moreover, even though the private operator would charge a higher price, there would be no welfare losses. In this article, we show that his result depends crucially on a strong assumption on demand. Keeping van Reeven's formulation we introduce a slight modification of demand representation to make it more realistic, and then show that:

(i) $f^{*} > f^{\pi}$;
(ii) welfare losses emerge under a profit-maximising operator;
(iii) subsidies are required for first-best operation.

Thus, contrary to the author's assertion, the Mohring effect is a valid argument for subsidisation.

2.0 Demand Properties in the Original Model

Van Reevan (2008) considers X potential travellers who have preferred departure times uniformly distributed along the period of analysis. All the travellers share the same utility function, namely:

$$U = v - p - \tau, \tag{1}$$

where v is a reservation utility level, p is the transit fare, and τ are waiting costs. If travellers do not know the timetable, then these waiting costs are given by $\tau = t/2f$, where t is value of time, and f is frequency. The cost of providing the service is simply modelled as $c \cdot f$. The key assumption behind the formulation in (1) is that either all consumers travel, or none does, depending on whether the full price is larger or smaller than v.

Therefore, the demand model is essentially an inelastic one: for any given frequency, price can be increased without affecting the demand level at all, up to a point where any infinitesimal increase in price makes all consumers vanish. This is important because the author claims that what makes the difference between his model and Mohring's is that here, changes in price and frequency would affect demand and that this is taken into account by a private operator: in other words, that the demand for the transport operator would be a function $D(p, f)$. Yet the demand function is quite special: $D = X$ whenever $v \geqslant p + \tau$ while $D = 0$ otherwise.

Without going into detail, the results that follow from such demand specification can be obtained directly by invoking results from Spence (1975). Spence analysed market failure when a monopolist, in addition to price, sets some aspect of product quality; here this would obviously be frequency. Spence showed that if $p(D, f)$ is the inverse demand function and $\partial^2 p/\partial D\partial f < 0$, then for a given D the monopolist undersupplies quality, that is, the *quality rule* is such that $f^{\pi}(D) < f^{*}(D)$, where π denotes profit-maximisation and * denotes the first-best. If $\partial^2 p/\partial D\partial f > 0$, on the other hand, then $f^{\pi}(D) > f^{*}(D)$ and the monopolist oversupplies quality for a given D. Finally, if $\partial^2 p/\partial D\partial f = 0$, then $f^{\pi}(D) = f^{*}(D)$. It is then direct that if $\partial^2 p/\partial D\partial f = 0$, and if for a given quality a monopolist would induce no allocative inefficiency, that is, $D^{\pi}(f) = D^{*}(f)$, then the actual first-best and profit-maximising qualities will be the same. This rare case is exactly van Reeven's: in the range where there is a positive demand, it holds that $\partial^2 p/\partial D\partial f = 0$ because $\partial p/\partial D = 0$ and, by construction, $D^{\pi} = D^{*} = X$. Therefore, as long as the profit-maximising firm decides to operate, Spence's result applies and profit-maximising frequency and first-best frequency will coincide, although prices will certainly not. Yet, since demand is essentially inelastic, the price increase is obviously nothing but a transfer, and therefore there are no welfare losses.

3.0 A Slightly More Realistic Model and the Case for Subsidies

There are several ways in which one can make more realistic demand assumptions, for example — following transport economics practice — one could add an alternative mode (for example, car) and use a binomial Logit model. Here, however, in order to make our point, it is enough to make only a slight change to van Reeven's model, namely, that not all consumers are identical. For this, it is sufficient to assume that the reservation utility v is uniformly distributed in the interval $[\underline{v}, \bar{v}]$ and that this is independent of the distribution of preferred departure times. It is clear then that a consumer characterised by v will travel if $U(v) \geqslant 0$; thus, the critical traveller's v is given by $v^* = p + t/2f$ and everyone with $v \geqslant v^*$ travels. Given the assumption about the uniformity of the distribution, then the demand for transit services will be given by $X(\bar{v} - v^*)/(\bar{v} - \underline{v})$. Replacing v^* and letting $Y \equiv X/(\bar{v} - \underline{v})$ we obtain the demand and inverse demand functions, respectively:

$$D(p, f) = Y\left(\bar{v} - p - \frac{t}{2f}\right) \qquad p(D, f) = \bar{v} - \frac{D}{Y} - \frac{t}{2f}. \tag{2}$$

It is apparent that this demand is indeed sensitive to price in a smooth way and that, in fact, it is linear in the full price or generalised cost, $p + \tau$. Moreover, since $\partial^2 p/\partial D \partial f = 0$ it will still be true that frequency rules will coincide, that is, $f^\pi(D) = f^*(D)$, following Spence (1975).

In order to calculate first-best frequency, price, and traffic, we need to obtain an expression for consumer surplus, which is given by $CS = \int_0^D p(D, f)dD - pD$. Straightforward calculations lead to $CS = (1/2)(D^2/Y)$.[1] Then, using the inverse demand function from (2), social welfare can be written as:[2]

$$\underset{D,f}{\text{Max}} \; \frac{1}{2}\frac{D^2}{Y} + p(D, f) \cdot D - c \cdot f,$$

where the first term is consumer surplus, and the last two represent the firm's profits. First-order conditions lead to the following frequency and traffic rules:

$$f^*(D) = \sqrt{\frac{Dt}{2c}}, \tag{3}$$

$$D^*(f) = Y\left(\bar{v} - \frac{t}{2f}\right). \tag{4}$$

On the other hand, the frequency and traffic rules that would maximise profit $p(D, f) \cdot D - c \cdot f$ happen to be:

$$f^\pi(D) = \sqrt{\frac{Dt}{2c}}, \tag{5}$$

$$D^\pi(f) = \frac{Y}{2}\left(\bar{v} - \frac{t}{2f}\right). \tag{6}$$

[1]The same result is obtained if we calculate instead $CS = Y\int_{v^*}^{\bar{v}} (v - p - t/(2f))\, dv$.

[2]Obviously, results do not depend on whether we take D and f, or p and f as decision variables. We use D and f because it leads more directly to the results we want to show.

As advanced, the frequency rules are the same. But now the traffic rules are not the same: a profit-maximising firm will induce — through pricing — a contraction of traffic for any given frequency. Replacing (4) in (3) and (6) in (5), we obtain the equations for the actual (not dependent on *D*) first-best and profit-maximising frequencies:

$$f^* = \sqrt{\frac{\left(\bar{v} - \frac{t}{2f^*}\right) Yt}{2c}}, \tag{7}$$

$$f^\pi = \sqrt{\frac{1}{2}}\sqrt{\frac{\left(\bar{v} - \frac{t}{2f^\pi}\right) Yt}{2c}}. \tag{8}$$

Equations (7) and (8) define fixed points for f^* and f^π. In the Appendix we show that if the fixed point for f^π has a solution, then the fixed-point for f^* has only one stable solution, and it fulfils $f^* > f^\pi$. Therefore, it is not true that if a private operator is allowed to choose price freely, it would provide the first-best frequency; it would actually undersupply frequency. This shows that a very simple and reasonable generalisation of van Reeven's formulation yields a completely different result.

Next we deal with the need for subsidies. To analyse this it is sufficient to replace equations (6) and (4) in the inverse demand function in (2). We obtain:

$$p^* = 0 \qquad p^\pi = \frac{1}{2}\left(\bar{v} - \frac{t}{2f}\right).$$

Note that the first-best frequency requires a first-best price that does not cover costs. In fact, the first-best requires a subsidy equal to $c \cdot f^*$, just as Mohring (1972), Jansson (1979), and others claimed within cost minimisation contexts.[3]

We now also show that our results do not change if consumers know the timetable. According to van Reeven, in this case things improve because customers no longer have to wait. Yet it is still true that customers suffer a dissatisfaction caused by the fact that they cannot take the bus at the

[3]It is worth looking at the findings of Evans (1987) who compares theoretically different regimes for a bus line regarding welfare, frequency, and price. Among other things, he finds that for every level of demand, the monopolist charges a larger than optimal fare and offers a lower than optimal frequency, producing a large welfare loss. Recently, Jara-Díaz and Gschwender (2009) have shown that, as the subsidy given to the transit system decreases from its optimal level, frequency and bus size decrease as well. Furthermore, they show that if no subsidies are provided but the resulting price is so large that the authority decides to impose a smaller price, then frequency would be reduced even more while bus size would increase.

time they would like to: customers now wait at home instead of the bus stop, but wait nevertheless. This time cost, caused by the difference between actual and preferred departure times, is usually known as *schedule delay cost*.

If the desired departure times were distributed uniformly across customers, and we redefine t to be the subjective value of schedule delay, then the demand and inverse demand functions would be exactly as the ones used above — given by (2) — and hence the conclusions would not change: a simple reinterpretation of waiting costs as schedule delay costs would deliver the same conclusions. But what if desired departure times are not uniformly distributed? In that case, the derivation of the demand function would be more complex and obviously dependent on the assumption about the distribution. Yet, it is easy to foresee that the resulting demand functions would fit into the general framework of the following demand function: $D(p, f) = Y(\bar{v} - p - g(f))$, where g is a general aggregated schedule delay function, whose precise functional form depends on the distribution of preferred departure times. Yet, for any distribution, function g would fulfil $g > 0$, $g' < 0$ and $\lim_{f \to \infty} g(f) = 0$. In other words, a marginal increase in frequency would raise demand in some general way given by $-Y \cdot g'(f) > 0$.

Having recognised this, it is now easy to see how our previous results go through: first since it is still true that $\partial^2 p/\partial D \partial f = 0$, it will still be true that frequency rules will coincide, that is, $f^{\pi}(D) = f^{*}(D)$, following Spence (1975). Next, the optimisation problem would be identical as before, since the expression for consumer surplus does not change. Consequently, one would obtain the same frequency and traffic rules as the ones obtained before — equations (3) to (6) — only that now $g(f)$ replaces $t/2f$; and since $g(f)$ is also decreasing in f, the proof that $f^{*} > f^{\pi}$ is similar to the one provided in the Appendix for the case of $t/2f$. Finally, it is also easy to show, following the same procedure as before, that the first-best requires subsidies: the unknown and known timetable cases lead to the same qualitative results.

4.0 Final Comments

Van Reeven argues that the main problem with the Mohring–Jansson reasoning would be that they do not take price effects into consideration. His model, however, incorporates price in a way that a change in bus fare induces no demand effect. By using a demand model that does respond to price changes — something that here was obtained simply by assuming away that all users are identical — we have shown that their conclusions do not change: profit-maximising frequencies are smaller than optimal and subsidies are required for first-best operation.

Furthermore, incorporating consumers' own time — which they use in the consumption of scheduled transport services — and then minimising the resources used is actually no different from considering demand effects and then maximising social welfare. To see this, consider for simplicity the unknown timetable case and note that the cost that would be minimised in a Mohring–Jansson setting would be the sum of users costs plus operator costs, that is, $Dt/2f + cf$, which yields the same optimal frequency rule as in (3). The optimal price, on the other hand is the difference between total marginal cost and average user costs, both of which are easily shown to be $t/2f^*$ leading to $p^* = 0$ and to an optimal subsidy $c \cdot f^*$, as obtained above.

References

Evans, A. (1987): 'A Theoretical Comparison of Competition with other Economic Regimes for Bus Services', *Journal of Transport Economics and Policy*, 21, 7–36.

Jansson, J. O. (1979): 'Marginal Cost Pricing of Scheduled Transport Services', *Journal of Transport Economics and Policy*, 13, 268–94.

Jara-Díaz, S. R. and A. Gschwender (2009): 'The Effect of Financial Constraints on the Optimal Design of Public Transport Services', *Transportation*, 36(1), 65–75.

Mohring, H. (1972): 'Optimization and Scale Economies in Urban Bus Transportation', *American Economic Review*, 62, 591–604.

Spence, M. (1975): 'Monopoly, Quality and Regulation', *The Bell Journal of Economics*, 6, 417–29.

van Reeven, P. (2008): 'Subsidisation of Urban Public Transport and the Mohring Effect', *Journal of Transport Economics and Policy*, 42(2), 349–59.

Appendix

Equations (7) and (8) define fixed points for f^* and f^π of the form:

$$f^* = \Omega(f^*) \qquad f^\pi = \sqrt{\frac{1}{2}} \cdot \Omega(f^\pi) \qquad \Omega(f) = \sqrt{\frac{\left(\bar{v} - \frac{t}{2f}\right) Yt}{2c}}.$$

It is then evident that $\Omega(f) = 0$ when $f = t/(2\bar{v}) > 0$, that $\Omega(f)$ is real only when $f \geqslant t/(2\bar{v})$ and that $d\Omega(f)/df > 0$. Further, when $f \to \infty$ then $\Omega(f) \to \sqrt{(\bar{v}Yt/2c)}$. Next, the solutions for f^* are at the intersection between $\Omega(f)$ and the 45° line; the solutions for f^π are at the intersection between $\sqrt{(1/2)}\Omega(f)$ and the 45° line. Therefore, it follows that if f^π has one solution, then f^* will have two solutions as shown in Figure A1 but only the one that fulfils $f^\pi < f^*$ is stable. If f^π has two solutions, then one is unstable, f^* will also have two solutions, and the stable solutions fulfil $f^\pi < f^*$.

Figure A.1
Optimal Stable Solution for Frequency