ROAD PRICING AND PUBLIC TRANSPORT

Kenneth A. Small
University of California at Irvine

September 29, 2003

Kenneth A. Small
Dept. of Economics
University of California
Irvine, CA 92697-5100
Phone: +1-949-824-5658
Email: ksmall@uci.edu

Forthcoming in:
Georgina Santos, ed., *Road Pricing: Theory and Evidence*, Elsevier.

ACKNOWLEDGEMENTS

This work was supported by the University of California Energy Institute. The author is grateful to Helen Wei for research assistance, and to Herbert Mohring, Chris Nash, Ian Parry, Stef Proost, Martin Richards, Deborah Salon, Ian Savage, Kurt Van Dender, Clifford Winston, participants at the June 2003 STELLA conference in Santa Barbara, two anonymous referees, and the editor for helpful comments. All results and opinions, as well as any errors, are the author's responsibility.

Abstract

Substantial benefits may arise from road pricing through its effects on the speed and service frequency of public transport. These effects are examined using a stylized model of local bus transport in a city center, a model requiring only a few parameters to obtain quantitative estimates. The model highlights four considerations: the cost savings to users and operators due to reduced road congestion; the service improvements made feasible by increased ridership; the potential pass-through of operator cost savings (even after paying for service improvements) as fare reductions; and the resulting multiplier effects on ridership and service offerings. The model is applied to central London using data from the first few months of the congestion charging program implemented in February 2003. Simulation results suggest significant effects, even if the pricing revenues had not been used to augment the public transport budget as they were in London: a ridership increase of 11 percent, a service increase of 7 percent, and user cost savings equivalent to 38 percent of the fare. Net benefits from these effects are equal to 39 percent of initial operator costs, suggesting the importance of fully accounting for these effects when evaluating congestion pricing proposals. These effects (but not the net benefits) are even larger in cities with more typical values for bus subsidies and initial modal share.

6. Road pricing and public transport

Kenneth A. Small

6.1 INTRODUCTION

More than three decades ago, Thomas Lisco (1970) argued that urban public transport in the United States suffers from poor quality or more precisely, a poorly chosen package of fare and level of service. His prescription was to raise fares and aim for a more high-class service that would appeal to increasingly affluent urban Americans.

Events have gone rather differently. Demand has sometimes been found to be quite sensitive to fares, for example in the comprehensive study of Boston's public transport system by Gómez-Ibáñez (1996). The main source of greater public transport funding has therefore been through sharp increases in federal and local subsidies. This seems to have caused little increase in public transport usage, not enough even to counteract other trends adverse to public transport. It has also caused some severe and widely documented inefficiencies, from high wages to uneconomical capital investments (Pickrell, 1985; Winston and Shirley, 1998).

Meanwhile, other approaches to reducing urban congestion have led to frustration. Demand for driving under peak conditions has outstripped the financial and environmental resources that communities are willing to devote to road building. Carpooling incentives, telecommuting, and other measures have had negligible effects on motor-vehicle use: the actual modal shifts have been extremely small and any improvements in congestion have been overwhelmed by "latent demand" for peak travel, i.e. demand that was previously deterred by congestion itself (Giuliano and Small, 1995; Small and Gómez-Ibáñez, 1999).

Road pricing is therefore a welcome newcomer to the menu of congestion-relief policies under serious consideration and trial. However, a potential benefit of road pricing that it is not widely appreciated is the dramatic effects it could have on public transport operations. One reason for this neglect is that a crucial link in the chain of effects is little known or understood outside of the community of professional transportation economists. Roughly, the entire chain can be described as follows:

- Raising the monetary price of car travel induces some modal shifting to public transport;
- Reduced congestion makes operating on-street public transport (bus or streetcar) faster and cheaper to operate;
- Increased route coverage and/or service frequency to handle the demand further enhances the service quality as perceived by the user;
- Higher costs of automobile commuting cause land near major business centres to become more valuable, hence to be developed at higher residential and commercial densities. This further enhances the market potential of public transport by increasing its density of demand in just those areas where it is already most efficient.

It is the third of these effects that is so little understood. Mohring (1972) showed that the dependence of service quality on frequency and route coverage is a form of economies of scale. This means that any modal shift toward public transport touches off a “virtuous circle” of further cost savings and/or service improvements, hence possibly further modal shifts. Viton (1983) and Kain (1994) suggest that the resulting equilibrium might involve a far higher modal share for public transport than the one prevailing before road pricing was undertaken, even not accounting for any changes in land use.

The purpose of this paper is to explore the potential quantitative importance of this “virtuous circle” in the short run, i.e. ignoring land use changes. It does so within the context of congestion pricing covering a major downtown area such as Singapore or London. The focus is on the likely impacts of road pricing on costs and service quality of urban bus transport, and on the second-round effects of these changes on the behaviour of public transport operators and potential users. The paper then considers the contribution of these impacts to the overall benefits from congestion pricing. The analysis is restricted to bus transport, partly because that is where the greatest benefits are likely and partly because the impacts of road pricing on underground service depend heavily on capacity constraints that are highly city and route specific.

It turns out that a reasonably sophisticated model of bus transport yields remarkably simple rules of thumb for making these predictions. These rules take as given the direct effects of road tolls on public transport ridership and average road speed. Using some preliminary numbers from London, it appears that the favourable impacts on public bus transport, and the associated contribution to net benefits of road pricing, are indeed considerable.

6.2 THEORETICAL MODEL

6.2.1 Model Structure

The model, adapted from Nash (1988) and Jansson (1997), both simplifies and extends the basic framework described by Mohring (1972). It simplifies Mohring's paper by ignoring the time taken by passenger boarding and alighting and the operator's decision as to where to place bus stops. It extends Mohring's paper, as do Nash and Jansson, by incorporating the spacing of routes as an operator decision. There are other, more minor, differences as well. The model is partial equilibrium, and so does not include the effects of changes in transport prices on other distortions in the economy.

Bus service takes place within a well-defined service area, which for convenience can be thought of as a rectangle. It has length ℓ (km) and a width described by the walk time $2w$ required to traverse it. Within this area, the operator can choose to operate N evenly spaced parallel routes, each of length ℓ. (Equivalently, the service area can be a circle of radius ℓ, whose perimeter could be walked in time $2w$, with evenly spaced radial routes converging at the centre and with walk access to those routes taking place along concentric circles). The bus operator can adjust the spacing of routes by choosing N and the average frequency of service by choosing vehicle-kilometres of service per hour, M. Passengers make trips at rate Q per hour, originating at locations evenly dispersed throughout the service area. The average trip length is m km. These and other symbols are summarised for convenience in Table 6.1. Conditions are assumed to be constant throughout the time period in which road pricing is in effect.

Table 6.1 Symbols used in the theoretical model

Bus service environment (exogenous):	
ℓ	Length of service area and of each route (km)
w	Time to walk half the width of service area if rectangular, or half its perimeter if circular (hr)
h	Duration of period of road pricing (hr/day)
m	Average trip length (km)
S	Average bus speed (km/hr)
Service and price variables chosen by bus operator	
N	Number of routes
M	Hourly bus-kilometres of service offered (km/hr)
f	Fare ($ per trip)
Bus ridership and finance:	
Q	Number of passenger-trips per hour
n	Average number of passengers on a bus = mQ/M
R	Revenue ($ per hour of service period)
Costs and cost parameters	
C	Cost to operator ($ per hour of service period)
U	Cost of user time to all users ($ per hour of service period)
Σ	Subsidy ($ per hour of service period)
ATC	Average total cost: $(C+U)/Q$ ($/trip)
a_1	Operator cost per bus-hour, independent of bus size
a_2	Increase in operator cost per bus-hour caused by increasing bus size to accommodate one additional passenger on average
α_w	Value of walking time ($/hr)
α_x	Value of waiting time ($/hr)
α_v	Value of in-vehicle time ($/hr)
Demand parameters	
ε_f	Demand elasticity with respect to fare
ε_M	Demand elasticity with respect to bus-km of service

Operator cost C is assumed proportional to the number of bus-hours of service, with the proportionality constant a linear function of bus size. Bus-hours of service are simply bus-kilometres of service M (which is controlled by the operator) divided by average bus speed S (which is not); this speed incorporates vehicle stops and passenger boarding and alighting time, so is less than the average speed of cars in the same locations.[1] Bus size must be adjusted in order to achieve a predetermined load factor (average percentage of passenger spaces filled); hence it is proportional to the average number of passengers on a bus at any point in time, n. Therefore agency cost is of the form:

$$C = (a_1 + a_2 n) \cdot M / S \tag{6.1}$$

where a_1 and a_2 are proportionality constants. All aggregate cost and service variables, such as C and M, are expressed per hour of a well-defined service period, which lasts h hours per day. This period may be taken to be that during which road pricing is in effect. Equation (6.1) is consistent with empirical evidence that bus-miles are supplied with few if any scale economies or diseconomies (Small, 1992:57).

The average number of passengers per bus, n, is simply the total passenger-km travelled per hour, mQ, divided by the bus-km of service offered per hour, M. Hence (6.1) can be rewritten:

$$C = (a_1 M / S) + (a_2 mQ / S) \equiv C_1 + C_2 \tag{6.2}$$

User cost U arises from three kinds of time spent. Walk time for the average user is equal to half the maximum walk time of a home from a bus line, that distance being w/N. Waiting time for the average user is half the headway between buses.[2] Headway is defined as the time interval between buses, which is the route-km, ℓN, divided by veh-km per hour, M. Finally, average in-vehicle time is equal to m/S, the average trip length divided by speed. The user's perceived opportunity costs for these three types of time (in \$/hour) are α_w, α_x, and α_v, respectively. Therefore aggregate user costs are:

[1] Kain (1994) gives evidence that bus operating speed is about 60 per cent of automobile speed in US cities.

[2] This assumption involves two biases in opposite directions. If headways are large, users will learn the schedule in order to shorten their waiting times. On the other hand, if they are trying to coordinate their trip with some other scheduled events, as is common, they incur costs of schedule mismatch, which are omitted here and which probably grow more than linearly with headway.

$$U = \tfrac{1}{2}(\alpha_w w / N) \cdot Q + \tfrac{1}{2}(\alpha_x \ell N / M) \cdot Q + (\alpha_v m / S) \cdot Q \equiv U_w + U_x + U_v \equiv U_o + U_v \quad (6.3)$$

Here the subscripts *w*, *x*, and *v* stand for "walk," "transfer," and "in-vehicle". In the last equality in (6.3), *U* is divided into out-of-vehicle and in-vehicle cost.

Ignoring boarding and alighting times causes the model to overstate scale economies somewhat and therefore to understate the incremental agency and user costs imposed by additional riders. This should not be too serious because of two other features of the model: the assumption of constant load factor, and the use of an assumed bus speed (and change in bus speed) that does incorporate those times along with all other reasons for buses being slower than surrounding traffic. As shall be seen, the model predicts that only one-third of any ridership increase is handled through increasing the number of passengers per bus, and therefore the effect of boarding and alighting on bus speed should be modest.

6.2.2 Optimal Operator Decisions and Cost Functions

Amazingly, this is all that is needed to compute the optimal responses, in percentage terms, of service levels, operating costs, and user costs to specified percentage increases in speed *S* and aggregate ridership *Q*. This can be done by simply letting the operator minimise the total cost of serving *Q* passengers. As is shown later, the operating agency will follow such a procedure if it is maximising aggregate consumer surplus less aggregate total cost, even if subject to a specified total subsidy. The following derivation is similar to that of Jansson (1997) except that he allows for boarding and alighting time.

The quantity *C*+*U* is therefore minimised with respect to the service variables *N* and *M*. This yields two first-order conditions:

$$(a_1 / S) - \tfrac{1}{2}(\alpha_x \ell N Q / M^2) = 0 \qquad (6.4)$$

$$-\tfrac{1}{2}(\alpha_w w Q / N^2) + \tfrac{1}{2}(\alpha_x \ell Q / M) = 0 \qquad (6.5)$$

Solving (6.4) for *M* yields the well-known "square-root rule" for adjusting optimal service offerings to changes in demand, when route structure is fixed (Mohring, 1972):

$$M = \left(\frac{\alpha_x \ell SN}{2a_1} \right)^{1/2} \cdot Q^{½} \tag{6.6}$$

However, route structure is not fixed. Solving (6.5) for optimal route spacing yields:

$$N = \left(\frac{\alpha_w w}{\alpha_x \ell} \cdot M \right)^{1/2} \tag{6.7}$$

Solving (6.6) and (6.7) for *N* and *M* yields a result different from the square-root rule:

$$N^* = (\alpha_w w)^{2/3} \cdot \left(\frac{S}{2a_1 \alpha_x \ell} \right)^{1/3} Q^{1/3} \tag{6.8}$$

$$M^* = (\alpha_x \ell \alpha_w w)^{1/3} \cdot \left(\frac{S}{2a_1} \right)^{2/3} \cdot Q^{2/3} \tag{6.9}$$

Optimal service provision M^* is seen to vary with output to the two-thirds power, not the one-half power; the reason for the difference is that when *N* can also be optimised, additional bus-kilometres can be used to reduce both walking and waiting time and thus are more beneficial than when *N* is held fixed. As shall be seen shortly, this is done at the expense of operating cost savings, which are still taken as *Q* expands but less rapidly than when route structure is fixed. It should also be noted that bus size, which is proportional to $n \equiv mQ/M$, varies with $Q^{1/3}$; thus two-thirds of a ridership increase is handled via additional buses and only one-third via larger buses.

The total cost function can be computed by substituting these solutions for optimal *N* and *M* into (6.2) and (6.3). Dividing each component by output yields the following expressions for the five average cost components corresponding to the five separate additive terms in (6.2) and (6.3):

$$C_2 / Q = a_2 m / S$$

$$U_v / Q = \alpha_v m / S \tag{6.10}$$

$$C_1 / Q = U_w / Q = U_x / Q = (\alpha_0 / 2)^{2/3} \cdot (a_1 / S)^{1/3} \cdot (Q / w\ell)^{-1/3}$$

where $\alpha_0 \equiv (\alpha_w \alpha_x)^{1/2}$ is the geometric mean of the values of walking and waiting time. Thus the first two components of average cost are independent of *Q*, indicating constant returns to scale. The other three components are equated to each other at the optimum; this generalises the result of Mohring (1972) that, under suitable simplifications, average operator cost is equated to average waiting cost. Furthermore, these latter three average cost components decline with passenger density, indicating scale economies. Average agency cost, $(C_1+C_2)/Q$, also declines with Q, consistent with statistical evidence for scale economies in final output (Small, 1992:57).

Combining all five cost components, the average total cost is:

$$ATC = \left(a_2 + \alpha_v\right)\cdot \frac{m}{S} + 3c_1 Q^{-1/3} \tag{6.11}$$

where $c_1 = \left(\alpha_0 / 2\right)^{2/3} \cdot \left(a_1 w\ell / S\right)^{1/3}$.

6.2.3 Revenues, Subsidies, and Constrained Welfare Maximisation

It is shown in this section why the agency might choose to minimise total cost even when operating under a fixed subsidy. Consider the broader problem of choosing service levels M and N and fare f in order to maximise social welfare, subject to a budget constraint involving a fixed aggregate subsidy Σ. The inverse demand curve for trips, $\pi(Q)$, is defined as the full price, including fare plus average user cost, at which ridership would be Q. Let $V(Q)$ be some measure of the value of public transport travel to consumers, such as the area under $\pi(Q)$. Subtracting total cost yields the following expressions for social welfare:

$$W = V(Q) - C - U \tag{6.12}$$

Revenues are equal to

$$R = f \cdot Q = \left[\pi(Q) - (U/Q)\right]\cdot Q = \pi(Q)\cdot Q - U$$

The agency's budget constraint, assumed for simplicity to be binding, is $C–R=\Sigma$ or

$$C + U - \pi(Q)\cdot Q = \Sigma \tag{6.13}$$

The quantity $C+U$ occurs as a unit in both (6.12) and (6.13), and furthermore the other terms in (6.12) and (6.13) do not contain N or M. Thus when a Lagrangian function is defined for this constrained maximisation problem, with multiplier λ on the constraint, the Lagrangian function divides into two parts; one part, $-(1+\lambda)\cdot(C+U)$, contains all the dependence on M and N. The optimality conditions for M and N are therefore identical to those derived previously for minimising $C+U$. Of course, the actual *amount* of service will be smaller if ridership Q is below the optimal level, but that amount will still minimise total cost given Q.

A similar derivation shows that a profit-maximizing private operator, also, will chose service characteristics to minimise total cost for given ridership level. Maximizing revenue less agency cost means maximizing $f \cdot Q - C = \pi(Q) \cdot Q - (C+U)$, which requires minimizing $(C+U)$ for given Q. This result, however, depends on the assumed uniformity of values of time; if such values differ across the population, a monopolist private operator would provide service levels weighted more heavily by the desires of those market segments with the highest demand elasticities.[3]

What if public transport providers have other objectives and so do not choose service levels to minimise the sum of agency and user costs? Several studies in the US, Australia, and UK, including London, have found that too much service is provided, even when accounting for the benefits to road users of enticing people from private vehicles.[4] In such a case, the model here does not apply precisely. Yet most of the benefits identified by the model are still measured accurately to first order in patronage changes. For example, if, for a given ridership, service levels are inefficiently high, then the ridership increase caused directly by pricing the competing auto mode can be handled at zero marginal cost (again abstracting from delays due to boarding and alighting of additional passengers). This produces a favorable fiscal effect on the operator, which is captured by the model here. What is not captured is the possibility that the cost savings might be dissipated in inefficient further service increases.

---

[3] For an analogous argument in the context of pricing by a monopolist road operator, see Edelson (1971).

[4] See Glaister (1987) for the UK, Dodgson (1986) for Australia, Winston and Shirley (1998) for a US aggregate analysis, Gómez-Ibáñez (1996) for Boston, and Savage and Schupp (1997) for Chicago. Glaister's estimates for the London bus system suggest that an optimal allocation of the agency's budget at that time would involve a 28 per cent reduction in fare and 31 per cent reduction in service levels (Savage and Schupp, 1997, Table 1). In contrast, Jansson (1980) and Larsen (1996) find that in the cases they study, more frequent service should be provided but with smaller vehicles, at higher fares if necessary to maintain budget balance.

6.3 IMPLICATIONS OF THE MODEL FOR ROAD PRICING

In this section, the results of small changes in bus operating speed dS and in ridership $dQ^{(1)}$ that are brought about directly by the reduction of automobile trips resulting from road pricing are considered. Some of these trips may be suppressed entirely or diverted to time periods outside the analysis; it does not matter for the present analysis because it is assumed that dS and $dQ^{(1)}$ are determined by some other model. It is also assumed that the adjustments in the public transport system described by the model here have only negligible effects on traffic speeds, so that dS captures all the needed information. On the other hand, $dQ^{(1)}$ is just the first-round effect of the fare increase, and does not include subsequent changes in ridership calculated here.

These assumptions are most appropriate when bus transport constitutes a small part of total trip-making and contributes only a small proportion of the traffic causing congestion. In other cases, of which London is surely one, they can be regarded as providing only a first approximation to the situation. A more complete model would need to describe the effect of public transport cost and quality on automobile traffic and the relationship between automobile traffic and travel speed.

The model of the previous section can easily be used to describe the results of finite as well as infinitesimal changes in S and Q. However, the results of doing so are less illuminating. Furthermore it is not clear that the improved accuracy is worth the trouble, given the model's aggregated and simplified nature.

6.3.1 Direct Effects of Improved Speed on Costs

Even without any adjustments by the operator, the improved speeds brought about by road pricing have favourable effects on both the agency and users, due to reduction in operating costs and in-vehicle time. To put it differently, speed S affects both agency cost C and user cost U directly, even in the absence of any changes in N, M, or Q. It can be easily seen from (6.2) and (6.3) that the affected components are C_1, C_2, and U_v, and that each responds to small changes in S with elasticity -1. If it is further assumed that the savings in agency cost are passed on through lower fares, then the budget constraint $f=(C-\Sigma)/Q$ implies that the fare responds with elasticity equal to that of cost multiplied by $C/(fQ)\equiv C/R$. These results can be summarised as follows:

$$\eta_{C,S} = -1; \quad \eta_{U_v,S} = -1; \quad \eta_{f,S} = -C/R \tag{6.14}$$

It should be noted that the favourable effect on bus in-vehicle time is matched by a similar effect on automobile in-vehicle time; thus rising speed does not necessarily create a net inducement for modal shift. Rather, the effect described here counteracts what otherwise would be a positive effect of improved speed on automobile usage. Since the latter effect is normally already included in the full price of automobile travel, it is important to recognise this competing improvement in the quality of bus service.

6.3.2 Direct Effects of Increased Bus Ridership on Costs

The ridership boost $dQ^{(1)}$ also affects cost, through the scale economies described earlier. These effects could be calculated either with or without optimal adjustments by the operator; either method would give the same total change in cost if the initial operation were optimised (perhaps subject to a subsidy constraint). Here, the effects are calculated assuming optimal operator adjustment.

Equation (6.9) shows that optimal bus service provision M^* is adjusted proportionally to $Q^{2/3}$, i.e. the elasticity of M^* with respect to Q is 2/3. Equation (6.10) shows (after multiplying by Q) that agency costs C_1 and C_2 respond to Q with elasticities 2/3 and 1, respectively. Therefore combined agency cost C responds with elasticity $\eta_{C,Q} = (2/3)\cdot(C_1/C) + 1\cdot(C_2/C) = (2/3)+(C_2/3C)$. Equation (6.10) also shows that user costs U_o and U_v respond to Q with elasticities 2/3 and 1, respectively (recalling $U_o=U_w+U_x$). This can be summarised as follows:

$$\eta_{M,Q} = \frac{2}{3}; \quad \eta_{C,Q} = \frac{2}{3} + \frac{C_2}{3C}; \quad \eta_{U_o,Q} = \frac{2}{3}; \quad \eta_{U_v,Q} = 1 \tag{6.15}$$

6.3.3 Effects on Agency Budget Balance

If fare were kept constant, then the bus subsidy could be reduced by an amount $dR–dC$, where $dR/R=dQ^{(1)}/Q$ and dC/C is determined by the elasticities of C in equations (6.14) and (6.15). Combining these results,

$$\frac{dR-dC}{C}=\frac{dS}{S}-\left(\frac{2}{3}+\frac{C_2}{3C}-\frac{R}{C}\right)\cdot\frac{dQ^{(1)}}{Q} \tag{6.16}$$

It can be seen that there are counteracting effects on the budget balance. Improved speed reduces total cost (first term), while increased ridership increases total cost but adds new revenue (second term). The overall balance is positive provided that

$$\frac{dS/S}{dQ^{(1)}/Q}>\frac{2}{3}+\frac{C_2}{3C}-\frac{R}{C}\,.$$

This condition is very likely to be met. Size-related cost C_2 is likely to be no greater than non-size-related cost (largely driver time), so that $C_2/(3C)\leq 1/6$; and the cost-recovery ratio is likely to be not too much smaller than 1/3. Therefore, the condition requires only a quite modest speed improvement, around half the ridership increase, and it could even be met with no speed improvement if the cost-recovery ratio is more than 1/2.

Alternatively, the argument behind (6.16) can be recast as the amount of fare reduction made possible if the total subsidy is unchanged. Again writing fare as $(C-\Sigma)/Q$ and using (6.14) and (6.15), the following expression is obtained:

$$\begin{aligned}\frac{df^{(1)}}{f}&=\frac{dC}{C}\cdot\frac{C}{R}-\frac{dQ^{(1)}}{Q}\\&=\frac{C}{R}\cdot\left[-\frac{dS}{S}+\left(\frac{2}{3}+\frac{C_2}{3C}-\frac{R}{C}\right)\cdot\frac{dQ^{(1)}}{Q}\right].\end{aligned} \tag{6.17}$$

Henceforth it is assumed that budget changes are passed through to passengers, i.e. that (6.17) applies.

6.3.4 Second-Round Effects on Ridership

Both the fare reduction computed in (6.17) and the service improvement computed in (6.15) will create a "second-round" increase in bus ridership, $dQ^{(2)}$. It is convenient to write this in terms of the elasticities of bus ridership with respect to fare (ε_f) and with respect to vehicle-km of service (ε_M). Using (6.15) to compute change in M yields:

$$\frac{dQ^{(2)}}{Q} = \varepsilon_f \frac{df^{(1)}}{f} + \frac{2}{3}\varepsilon_M \frac{dQ^{(1)}}{Q}$$

where $df^{(1)}/f$ is given by (6.17). Substituting (6.17) and grouping terms yields

$$\frac{dQ^{(2)}}{Q} = -\varepsilon_f \cdot \frac{C}{R}\frac{dS}{S} + \mu \cdot \frac{dQ^{(1)}}{Q} \tag{6.18}$$

where

$$\mu = \frac{2}{3}\varepsilon_M + \frac{C}{R}\cdot\left(\frac{2}{3} + \frac{C_2}{3C} - \frac{R}{C}\right)\cdot \varepsilon_f \tag{6.19}$$

is a factor leading to a multiplier on ridership changes. It should be noted that the first term of (6.19) is positive, while the second term could be positive or negative but, as argued before, is probably positive.

Subsequent rounds of fare and service changes result in further ridership changes following the iterative formula $dQ^{(n)} = \mu \cdot dQ^{(n-1)}$ for $n \geq 3$. Combining them, the ultimate change in ridership resulting from the direct changes in speed dS and ridership $dQ^{(1)}$ is

$$\frac{dQ^{SQ}}{Q} = \frac{1}{1-\mu}\cdot\left[-\varepsilon_f \cdot \frac{C}{R}\cdot\frac{dS}{S} + \frac{dQ^{(1)}}{Q}\right] > 0 \tag{6.20}$$

provided that $|\mu|<1$. The corresponding fare and service changes are obtained analogously from (6.17) and the first of equations (6.15):

$$\frac{df^{SQ}}{f} = \frac{C}{R}\cdot\left[-\frac{dS}{S} + \left(\frac{2}{3} + \frac{C_2}{3C} - \frac{R}{C}\right)\cdot\frac{dQ^{SQ}}{Q}\right] \tag{6.21}$$

and

$$\frac{dM^{SQ}}{M} = \frac{2}{3}\frac{dQ^{SQ}}{Q} \tag{6.22}$$

A more elegant way of deriving (6.20) is in fact to simultaneously solve (6.21), (6.22), and the equation expressing the sources of change in ridership.[5]

A crucial simplifying assumption made here is that the feedback effects of these second- and subsequent-round ridership increases on traffic speed can be ignored. Accounting for such changes would further magnify the favourable results found in the simulations presented later.

6.3.5 Effects of Using Road Pricing Revenues to Increase Bus Service

In some cases, such as London in 2003, a road pricing policy may include a provision to use all or part of the revenues for improved public transport service. Suppose, then, that the bus transport operator receives an additional exogenous increment in its subsidy, $d\Sigma$, designated for increased service M. Solving (6.2) for M, substituting Σ-R for C, and varying Σ while holding Q, S, and R constant, gives an elasticity of bus-kilometres M with respect to subsidy Σ of $\eta_{M,\Sigma} = \Sigma / C_1$. (It is larger than Σ/C because increasing service with Q constant permits use of smaller buses, which stretches the subsidy still further). This induces a first-round increase in ridership governed by elasticity ε_M, and subsequent rounds governed by the multiplier μ as before. Combining all these rounds, the ultimate service and ridership increase arising from the increased subsidy are:

$$\frac{dM^{\Sigma}}{M} = \frac{1}{1-\mu} \cdot \frac{d\Sigma}{C_1} \tag{6.23}$$

and

$$\frac{dQ^{\Sigma}}{Q} = \frac{\varepsilon_M}{1-\mu} \cdot \frac{d\Sigma}{C_1}\,. \tag{6.24}$$

[5] That equation is $dQ^{SQ}/Q = dQ^{(1)}/Q + \varepsilon_f \cdot df^{SQ}/f + \varepsilon_M \cdot dM^{SQ}/M$. The three equations are solved for the three unknowns dQ^{SQ}/Q, df^{SQ}/f, and dM^{SQ}/M.

Equations (6.20) and (6.24) are additive if all three first-round effects (i.e. the changes in speed, ridership, and subsidy) occur simultaneously, as for example is the case in London.

6.4 BENEFITS FROM ROAD PRICING

In this section, the model is used to assess the benefits from road pricing that occur due to the impacts on the public transport sector. That is, the results indicate what *additional* benefits occur beyond those computed just from the analysis of private vehicles alone.[6]

First, consider the benefits dB^S from improved speed, dS. They are simply the savings in agency costs (assumed here, for simplicity, not to be used to decrease fares) and in user costs. Using the elasticities shown in equation (6.14), these benefits, to a first-order approximation in speed change, are:

$$dB^S = (C + U_v) \cdot (dS / S) \tag{6.25}$$

Next, consider the benefits dB^Q from dQ^{SQ}, i.e. from the increase in bus ridership excluding that resulting from subsidy $d\Sigma$. There are three components. First is increased revenues to the bus operating agency. To understand this component, it should be noted that some of the increase in ridership comes from mode switches from private vehicles. The lost consumer surplus to mode switchers, already included as part of the assessment of the costs and benefits to private vehicles, takes as given the full price for bus encountered by those switchers. Part of that full price is average user cost on the bus mode for the switchers, which is a real cost and so is already accounted for. Another part, however, is the bus fare paid, which is a transfer to the public transport operating agency and therefore is a benefit in the public transport sector. Given constant fare, this component of benefits may be written as:

$$dR = R \cdot \left(dQ^{SQ} / Q\right)$$

[6] The benefits from private vehicles alone consist of the cost savings for continuing users less the lost consumer surplus for prior users. Equivalently, they consist of toll revenues plus changes in consumer surplus (negative in most cases) for all original private vehicle users.

The second component consists of the negative of the extra costs dC incurred by the bus operator to handle this increased ridership. From equation (6.15) and using $C \equiv C_1 + C_2$, it is

$$-dC = -\eta_{C,Q} \cdot C \cdot \left(dQ^{SQ}/Q\right) = \left(-C + \frac{1}{3}C_1\right) \cdot \left(dQ^{SQ}/Q\right)$$

It is worth noting that if the optimised marginal agency cost, $[C-(1/3)C_1]/Q$, exceeds the fare, then the sum of the first two components of benefits, $dR-dC$, is negative; this would indicate that the bus agency's finances are harmed by the additional ridership.[7]

The third component of benefits consists of cost savings to existing users from improved service. This may be written, again using (6.15), as

$$\begin{aligned} -Qd\left(U/Q\right) &= -dU + U \cdot \left(dQ^{SQ}/Q\right) \\ &= -\left(\eta_{U_o,Q} \cdot U_o + \eta_{U_v,Q} \cdot U_v\right) \cdot \left(dQ^{SQ}/Q\right) + \left(U_o + U_v\right) \cdot \left(dQ^{SQ}/Q\right) \\ &= \left(-\frac{2}{3} \cdot U_o - 1 \cdot U_v + \left(U_o + U_v\right)\right) \cdot \left(dQ^{SQ}/Q\right) \\ &= \frac{2}{3} \cdot C_1 \cdot \left(dQ^{SQ}/Q\right) \end{aligned}$$

where the last equality follows because, from (6.10), $U_o \equiv U_w + U_x = 2C_1$.

Combining the three components of dB^Q and simplifying, it is found that

$$dB^Q = dR - dC - Qd(U/Q) = \left(R - C_2\right) \cdot \left(dQ^{SQ}/Q\right) \tag{6.26}$$

Equation (6.26) has a simple interpretation. Assume that the public transport operator was already choosing service level optimally, subject to a given fare. The envelope theorem then implies that the change in total costs from additional passengers dQ^{SQ} can be computed by assuming they are handled with no change in bus schedules or routes. In that case, the agency accommodates passengers just by increasing bus size. Agency profit then changes by the increased revenue (proportional to R) less the cost of increased bus size (proportional to C_2). Under the assumptions made above, the welfare effects outside the agency are already captured in the analysis of car traffic (in the case of $dQ^{(1)}$) or may be assumed zero (in the case of $dQ^{(2)}$,

[7] The optimised marginal agency cost can alternatively be computed from (6.2) as:
$\dfrac{d(C_1 + C_2)}{dQ} = \dfrac{2}{3}\dfrac{C_1}{Q} + \dfrac{C_2}{Q} = (1/Q) \cdot \left(C - \dfrac{1}{3}C_1\right)$.

$dQ^{(3)}$, etc.) provided the subsequent-round increases in ridership are drawn from activities in which price equals marginal social cost.[8] Thus the increase in bus ridership may be treated, to first order in small changes, as a purely fiscal matter and so the theory can accommodate either positive or negative welfare effects from increasing bus ridership.

Combining the benefits from speed and ridership increases, writing dB^{SQ} for the total, and expressing the result as a fraction of agency cost gives

$$\frac{dB^{SQ}}{C} = \left(1 + \frac{U_v}{C}\right) \cdot \frac{dS}{S} + \left(\frac{R}{C} - \frac{C_2}{C}\right) \cdot \frac{dQ^{SQ}}{Q} \tag{6.27}$$

Remarkably, only one new parameter, U_v/C, is required to compute equation (6.27). As a very crude approximation, U_v/C could be expected to be close to the average value of user time, multiplied by number of passengers per bus, divided by the average driver wage rate (inflated for breaks and other non-productive time); hence it is greater than one under most circumstances, and could be quite large when each bus carries many passengers.

There is no comparable benefit from the use of road-toll revenues for public transport subsidies, because the revenues are already accounted for in the standard analysis of road pricing's effects on road users. That analysis assumes revenues are used efficiently; to first order in the changes, prior optimisation implies that this assumption can be met by spending the road-toll revenues on service improvements. Here the assumption that existing service levels are efficient is obviously critical.

6.5 NUMERICAL EXAMPLE: LONDON

It is interesting to calculate the consequences of road pricing for bus transport using the model described here and some realistic parameters. In this section, results are first computed using parameters representative of London's dramatic implementation of road pricing in February

---

[8] The quantity changes are assumed to be small enough so that each new bus rider is marginal, i.e. he or she can be assumed to be indifferent between riding the bus and doing whatever he or she did previously.

2003. They are then compared with results from some alternative sets of parameters likely to be more representative of other bus systems.

6.5.1 Background

In the London road pricing scheme, all vehicles travelling within central London on weekdays between 7:00 AM and 6:30 PM are charged £5, or approximately US$8, once for the entire day. Taxis and motorcycles are exempt, and residents of the central area receive a large discount. Net revenues are targeted to improvements in public transport. In addition, prior to implementation, a major program of improvements to bus transport was undertaken.[9]

6.5.2 Parameters for the Numerical Example

Table 6.2 presents the assumed parameters; justification is given in the Appendix. Available fiscal data refer to the entire London bus system, rather than just the central area; but since all parameters are ratios, this should be reasonably accurate.

Table 6.2 Parameters for numerical exercise

R/C	0.80	dS/S	0.09
C_2/C	0.45	$dQ^{(1)}/Q$	0.06
U_v/C	2.90	$d\Sigma/C$	0.07
ε_f	–0.25		
ε_M	0.30		

Source: See Appendix.

London is unusual in two respects, leading to the consideration of some alternative scenarios as well. First, it has a high cost-recovery ratio, *R/C,* estimated here to be 0.80. An alternative scenario takes this ratio to be 0.32.

[9] For further information see the Transport for London web site on congestion charging: www.tfl.gov.uk/tfl/cc_intro.shtml.

Second, the direct modal diversion $dQ^{(1)}/Q$ is rather small, estimated here at 6 per cent. The reason is that public transport usage in central London is already very high, with only 12 per cent of people entering the area during the morning peak traveling by car, according to Transport for London's monitoring study (Transport for London, 2003c:130, Figure 6.1). Thus the impact on bus is modest, even though the system was designed so that bus would absorb proportionally much more of the mode shift than the underground.[10] To represent cities with more typical initial mode share for bus transport, an alternative scenario is also considered in which $dQ^{(1)}/Q$ is five times as large.

As shall be seen, the model predicts that the incentive of higher speeds, fare reductions, and service improvements brought about by the program will cause a much greater increase in ridership than this direct modal diversion. It appears, in fact, that Transport for London anticipated this outcome by putting in place a major program of bus service improvements, leading to a 10 per cent ridership increase, before the start of congestion charging. Perhaps these improvements would have been undertaken regardless of the congestion charging, but it is equally reasonable to view them as part of a comprehensive package aimed at making congestion charging successful. The prior ridership increase therefore may be regarded, at least in part, as reflecting some of the later-round ridership increases $dQ^{SQ}+dQ^{\Sigma}-dQ^{(1)}$ predicted by the model, and the cost of providing those prior service increases may be regarded as being covered in part by the cost savings and revenue increases estimated in the model. (As it happens, the model does predict these later-round ridership increases to be almost exactly 10 percent.)

[10] The underground was already operating very close to capacity. Part of the strategy, encouraged by public transport pricing policies, was to shift many existing shorter trips from rail to bus, so that rail could accommodate some of the longer trips that shifted to public transport due to road pricing. The author is indebted to Chris Nash for this observation.

6.5.3 Results

Table 6.3 presents results of the base case, as well as four alternative scenarios: the two mentioned in Section 6.5.2, a third in which none of the road pricing revenues are spent on public transport, and a fourth in which all three changes are made simultaneously.

Table 6.3 Results of London Numerical Example

Change in:	*As fraction of:*	Base Case	R/C x0.4	$dQ^{(1)}/Q$ x5	$d\Sigma$=0	All three changes
Bus-km of service	Bus-km of service	0.231	0.180	0.430	0.073	0.208
Fare	Fare	-0.110	-0.276	-0.104	-0.110	-0.265
Ridership	Ridership	0.157	0.142	0.455	0.109	0.312
Average cost:						
In-vehicle time cost	Fare	-0.326	-0.816	-0.326	-0.326	-0.816
Out-of-vehicle time cost	Fare	-0.159	-0.310	-0.295	-0.050	-0.357
Total average user cost	Fare	-0.485	-1.125	-0.622	-0.376	-1.173
Average agency cost	Avg agency cost	-0.049	-0.069	-0.104	-0.110	-0.147
Benefits:						
From speed change	Agency cost	0.351	0.351	0.351	0.351	0.351
From ridership increase	Agency cost	0.038	-0.014	0.143	0.038	-0.041
Total benefits	Agency cost	0.389	0.337	0.494	0.389	0.310

Source: Generated by author; see text.

Several results stand out. First, the increase in bus speed is very important. The reduction in user in-vehicle time is equivalent, from the user's point of view, to a 33 per cent reduction in the fare (base case, row 4). This in turn is responsible for the expansion of the assumed direct ridership increase (six percent) into the 11 percent increase that would occur even without the increased subsidy (column labeled $d\Sigma$=0, row 3).

Second, the cost savings to the agency from higher bus speed are a very significant offset to the extra costs incurred as ridership expands. In the absence of a change in subsidy (column $d\Sigma$=0), average agency cost falls by 11 per cent, which happens to just offset the ridership increase. Holding aggregate subsidy constant, this enables the fare to be reduced by 11 per cent (column labeled $d\Sigma$=0, row 2).

Third, the complete package as envisioned in London produces a ridership increase much greater than that assumed from direct modal diversion. In the base case, the 6 per cent direct increase becomes a 16 per cent total increase; of the 10 percentage-point difference, half is due to the speed change and passed-through fare reductions, while the other half is due to the additional service funded by the road pricing revenues.[11]

Fourth, the net benefits are a significant fraction of initial aggregate agency costs (39 per cent in the base case). Most of these net benefits arise from the speed change. They exclude any advantages from spending road pricing revenues because, as noted earlier, those revenues are assumed to be already accounted for in the evaluation of the road pricing scheme's effects on private vehicles.

Fifth, the service improvements are quite substantial, amounting to a 23 per cent increase in bus-km of service in the base case and 7 per cent even if the road pricing revenues were spent elsewhere. These improvements have a favourable effect on average out-of-vehicle time costs (walking and waiting), which decline by 16 per cent of the initial fare in the base case and 5 per cent when subsidies are kept constant. The combination of in- and out-of-vehicle time changes is equivalent to a 48 per cent fare reduction in the base case.

Some of these results are even larger when the cost-recovery ratio is lower (second column), due mainly to the initial public transport revenue base being smaller. The reduction in average user cost is now 112 per cent of the fare, which from the user's perspective adds to the fare reduction itself of 28 per cent. With fare so far below average cost, the ridership increase is actually a net disbenefit; yet it is a very small one because the extra agency cost incurred is mostly offset by cost savings to users. Meanwhile the benefits from increased bus speed remain enormous.

[11] The first half, 4.9 percentage points, is the difference between the fourth entry in row 3 and $dQ^{(1)}/Q$. The other half, 4.8 percentage points, is the difference between the first and fourth entries in row 3.

Many results are larger when the direct modal diversion is larger (third column). In that case, service increases by 43 per cent, ridership by 45 per cent, average user cost declines by 62 per cent of the fare, and total benefits are 49 per cent of agency costs.

Recent research has disclosed that the benefits of congestion pricing may be much reduced by adverse effects on pre-existing tax distortions when general-equilibrium effects are included, unless the revenues are used to reduce other taxes (Parry and Bento, 2001; Van Dender, 2003). The main reason is that the cost of living is raised, hence the real wage is lowered, adding to the distortion from taxes on labour income. Such effects, not incorporated in this model, would reduce the direct benefits of congestion pricing and magnify the benefits through public transport. They also constitute an argument for considering the third alternative scenario, in which road-pricing revenues are not spent on public transport.

As already noted, the values for cost-recovery ratio and modal diversion shown in the alternative scenarios are more typical of bus transport in many large cities than are the base-case values. Arguably, then, the most useful scenario to consider for potential sites outside London is that shown in the last column, where all three alternative scenarios apply simultaneously. These results portray a system of public transport that is greatly transformed. Service is up by 21 per cent, ridership by 31 per cent, fares are down by 27 per cent, average agency cost is reduced by 15 per cent, and average user costs are reduced by 117 per cent of the fare. Net benefits are nearly one-third of agency costs.

6.6 CONCLUSIONS

Road pricing in a large city can dramatically change the role of public transport, at least of those modes that share the streets with private vehicles. Even without spending any of the road pricing revenues on public transport improvements, the combination of increased traffic speed and increased ridership permits large increases in service and ridership, reductions in user costs, and savings in average agency costs sufficient to pay for the increased service even while reducing fares. Net benefits may be very large, on the order of 30 to 40 per cent of agency costs, if traffic speed can be increased by just nine percent.

These basic results can be derived from a reasonably simple aggregate model requiring knowledge of remarkably few parameters. Two primary mechanisms drive the results: cost savings to both agencies and users from faster traffic, and optimisation of service levels in response to increased ridership. Of course, results would be quite different if pricing were applied to radial expressways with unpriced parallel arterials, as is being considered elsewhere in the UK. That case could be quite unfavorable for public transport if the bus routes become more congested as a result of diverted car traffic.

A number of model limitations are worth bearing in mind. First, in some cases the shift in demand to public transport could require expensive capital investments, especially to rail transport. Second, the implied targeting of public transport service toward places where ridership is greater may upset the precarious balance of power between suburban and inner-city communities that sometimes holds together regional public transport systems in the US. (This could, however, turn out to be a blessing in disguise because these regional systems have diverted a lot of subsidies to locations where public transport service is not economical.) Third, operators may not adjust optimally to changes and may waste subsidy funds. Finally, it would be desirable to examine the public transport system at a disaggregated level in order to describe in greater detail what an optimal response would look like.

The welfare benefits arising from the public transport sector should, in principle, affect the design of road pricing itself. The direction of this effect is uncertain: while the effects on public transport reduce the level of toll required to achieve a given reduction in congestion (Kain 1994:531), they also increase the benefits of that reduction. Attempts to take public transport into account in road pricing optimisation include Parry and Bento (2002) and Van Dender (2003). The model in this paper is somewhat richer than theirs in the way it predicts adaptation by the public transport agency to changes in its environment. There is certainly room for research, which combines the model features considered here with the general-equilibrium effects considered by those authors.

An important extension would be to consider two additional competing modes of urban travel: shared-ride taxi and single-destination shuttle services. Both have scale economies similar to those described here for public transport (Schroeter, 1983; Yang and Wong, 1998). Shared-ride taxi service may be a cheaper alternative to conventional public transport in low-density areas, and thus may be part of a comprehensive proposal to enhance the financial viability of public transport. Shuttle services have been successful at carrying passengers to and from airports and could probably be expanded to other destinations. Both will be affected significantly by road pricing policies in much the same way as public transport.

REFERENCES

Department for Transport (2002). *Transport Statistics Great Britain: 2002 Edition*. London: The Stationary Office.

Dodgson, J.S. (1986). Benefits of Changes in Urban Public Transport Subsidies in the Major Australian Cities. *Economic Record 62*, 224-235.

Edelson, N.M. (1971). Congestion Tolls Under Monopoly. *American Economic Review 61*, 873-882.

Giuliano, G. and Small, K. A. (1995). Alternative Strategies for Coping with Traffic Congestion. In H. Giersch (Ed.), *Urban Agglomeration and Economic Growth* (pp. 199-225). Heidelberg: Springer-Verlag.

Glaister, S. (1986). Bus Deregulation, Competition and Vehicle Size. *Journal of Transport Economics and Policy 20*, 217-244.

Glaister, S. (1987). Allocation of Urban Public Transport Subsidy. In S. Glaister (Ed.), *Transport Subsidy* (pp. 27-39). Newbury, UK: Policy Journals.

Gómez-Ibáñez, J. A. (1996). Big City Transit Ridership, Deficits and Politics: Avoiding Reality in Boston. *Journal of the American Planning Association 62*, 30-50.

Goodwin, P.B. (1992). A Review of New Demand Elasticities with Special Reference to Short and Long Run Effects of Price Changes. *Journal of Transport Economics and Policy 26*, 155-169.

Jansson, J.O. (1980). A Simple Bus Line Model for Optimisation of Service Frequency and Bus Size. *Journal of Transport Economics and Policy 14*, 53-80.

Jansson, J.O. (1997). Theory and Practice of Transport Infrastructure and Public Transport Pricing. In G. de Rus G. and C. Nash (Eds.), *Recent Developments in Transport Economics* (pp. 74-134). Aldershot, UK: Ashgate.

Kain, J. F. (1994). Impacts of Congestion Pricing on Transit and Carpool Demand and Supply. In National Research Council, *Curbing Gridlock: Peak-Period Fees to Relieve Traffic Congestion*, Transportation Research Board Special Report 242, Vol. 2 (pp. 502-553). Washington, D.C.: National Academy Press.

Kain, J. F. and Liu, Z. (1999). Secrets of Success: Assessing the Large Increases in Transit Ridership Achieved by Houston and San Diego Transit Providers. *Transportation Research A 33*, 601-624.

Larsen, O. I. (1996). Efficient public transit subsidies. In Hensher, D.A., King, J. and Oum, T. A. (Eds), *Proceedings of the 7th World Conference on Transport Research*, *Vol 3* (pp. 225-236). Oxford: Pergamon Press.

Lisco, T. E. (1970). Mass Transportation: Cinderella in Our Cities. *The Public Interest 18*, 52-74.

Lago, A. M., Mayworm, P. D., and McEnroe, J. M. (1981). Further Evidence on Aggregate and Disaggregate Transit Fare Elasticities. *Transportation Research Record 799*, 42-47.

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

Mohring, H. (1983). Minibuses in Urban Transportation. *Journal of Urban Economics 14*, 293-317.

Nash, C. (1988). Integration of Public Transport: an Economic Assessment. In J. S. Dodgson and N. Topham (Eds.), *Bus Deregulation and Privatisation* (pp. 97-118). Aldershot, UK: Avebury.

Parry, I. and Bento, A. (2001). Revenue Recycling and the Welfare Effects of Road Pricing. *Scandinavian Journal of Economics 103*, 645-671.

Parry, I. and Bento, A. (2002). Estimating the Welfare Effect of Congestion Taxes: The Critical Importance of Other Distortions within the Transport System. *Journal of Urban Economics 51*, 339-365.

Pickrell, D. (1985). Rising Deficits and the Uses of Transit Subsidies in the United States. *Journal of Transport Economics and Policy 19*, 281-298.

Pratt, R., Park, G., Texas Transportation Institute, Cambridge Systematics, Parsons Brinkerhoff Quade & Douglas, SG Associates, and McCollom Management Consulting (2000). *Traveler Response to Transportation System Changes: Interim Handbook.* http://gulliver.trb.org/publications/tcrp/tcrp_webdoc_12.pdf, accessed April 2002.

Savage, I. and Schupp, A. (1997). Evaluating Transit Subsidies in Chicago. *Journal of Public Transportation 1*, 93-117.

Schroeter, J. (1983). A Model of Taxi Service under Fare Structure and Fleet Size Regulation. *Bell Journal of Economics 14*, 81-96.

Small, K.A. (1992). *Urban Transportation Economics*, Fundamentals of Pure and Applied Economics, Vol. 51. Chur, Switzerland: Harwood Academic Publishers.

Small, K. A. and Gómez-Ibáñez, J. A. (1999). Urban Transportation. In P. Cheshire and E. S. Mills (Eds.), *Handbook of Regional and Urban Economics, Vol. 3: Applied Urban Economics* (pp. 1937-1999). Amsterdam: North-Holland.

Transport for London (2001). *Transport Statistics for London 2001*. London: Transport for London.

Transport for London (2003a). Central London Congestion Charging Scheme: Three Months On. London: Transport for London. www.tfl.gov.uk/tfl/pdfdocs/congestion_charging/cc-three-month-report.pdf, accessed 1 Sept 2003.

Transport for London (2003b). Congestion Charging Delivers Better Buses. London: Transport for London Press Release, 18 March. www.tfl.gov.uk/buses/press_666.shtml, accessed 25 March 2003.

Transport for London (2003c). Impacts Monitoring Programme: First Annual Report. London: Transport for London. www.tfl.gov.uk/tfl/cc_monitoring.shtml, accessed 1 Sept 2003.

Transport for London (2003d). Public Transport Improvements – After. London: Transport for London. www.tfl.gov.uk/tfl/cc_improve_pt_after.shtml, accessed 18 April 2003.

Van Dender, K. (2003). Transport Taxes with Multiple Trip Purposes. *Scandinavian Journal of Economics 105*, 295-310.

Viton, P A. (1983). Pareto-Optimal Urban Transportation Equilibria. In T. E. Keeler (Ed.), *Research in Transportation Economics, Vol. 1* (pp.75-101). Greenwich, Connecticut: JAI Press.

Wardman, M. (2001). A Review of British Evidence on Time and Service Quality Valuations. *Transportation Research E 36*, 107-128.

White, P. (1995). Deregulation of Local Bus Services in Great Britain: An Introductory Review. *Transport Reviews 15*, 185-209.

Winston, C. and Shirley, C. (1998). *Alternate Route: Toward Efficient Urban Transportation*. Washington D.C.: Brookings Institution Press.

Yang, H. and Wong, S. C. (1998). A Network Model of Urban Taxi Services, *Transportation Research B 32*, 235-246.

APPENDIX

Parameters for the Numerical Exercise

The first set of parameters describes the base scenario for the London bus system in 2000-01. Data are lacking for the central area specifically, so the ratios apply to the entire London region.

- Cost recovery ratio (R/C): Statistics from Transport for London (TfL) show that for 2000-01, the London bus system's receipts and costs were nearly identical, with receipts £656 million and costs £643 million (TfL 2001:38, Table 21). However these receipts include £118 million in "concessionary fare reimbursements" (Department for Transport 2002, Table 5.6), which represent subsidies rather than revenues. Therefore R/C = (643-118)/656 = 0.80.

- Ratio of user in-vehicle time costs to agency cost (U_v/C): The value of in-vehicle time for bus users is taken to be 40 per cent of the London area wage rate, which is estimated here at £530/40 from weekly earnings in year 2000 (TfL 2001:7, Table 2b). Thus α_v=£5.30/hr. The number of passenger-hours in buses per year is calculated at 4,709 million passenger-km in 2000-01 (TfL 2001:37, Table 20a) divided by average bus speed for the inner area of approximately 13.2 km/hr.[12] This yields U_v=£1,891 million for 2000-01. Again using bus system costs for that year of £643 million (as in the previous paragraph), U_v/C = 1,891/643 = 2.9.

- Size-related agency cost as fraction of agency cost (C_2/C): White (1995:197) quotes a 1990 study estimating that a minibus with 16-20 seats has operating cost per bus-km about 65 per cent that of a full-size bus in the UK. Assuming this applies to 20 seats and a full-size bus is 86 seats, extrapolation implies that $C_2/C\approx 0.45$.[13,14]

---

[12] Bus speed in 2002 for the central area is reported to be 11 km/hr in TfL (2003c:107). This number is been inflated here by 20 per cent, which is the ratio of traffic speeds in Inner London to Central London (averaged over morning peak, daytime between peaks, and evening peak), from TfL (2001:23, Table 12).

[13] A similar result is reported by Nash (1988, Table 5.1), who compiles figures from Glaister (1986) for the parameters of the linear cost function (6.1); they imply that if an 86-passenger bus is taken as representative of the current situation, $C_2/C\approx 0.50$. (Glaister's figures are based on cost per bus-km, not

The next parameters are demand elasticities:

- Fare elasticity (ε_f): Some typical estimates include –0.5 for a selection of studies, –0.35 for London bus, –0.25 for Chicago peak bus, and –0.32 for Houston and San Diego.[15] The value measured for London bus is most appropriate here, but must be reduced further because the fare changes considered here apply only to the central area whereas many trips begin or end outside that area. The value chosen is therefore ε_f=–0.25.

- Service elasticity (ε_M): some estimates include headway elasticities of 0.47 (Pratt et al. 2000:9-26, summarizing Lago *et al*. 1981); and service elasticities of 0.32 and 0.65 for Dallas and San Diego (Pratt et al. 2000, p. 9-14), 0.5 or higher in several US cities (Kain 1994:543), and 0.71 for Houston and San Diego (Kain and Liu 1999). With such a large initial share, London's bus system could be expected to have a somewhat smaller elasticity than other cities with perhaps an elasticity of 0.40; this is reduced further, to the value ε_M=0.30, to reflect the fact that many trips are only partly within the central area where the increased bus-miles are assumed to occur.

(..continued)
per bus-hr; this is appropriate here because any speed differential between a small and large bus is part of the cost advantage of the former.) Jansson (1980:68) implies that C_2/C=0.24–0.30. By contrast, Mohring (1983) calculates a ratio for Minneapolis that implies $C_2/C\approx1.0$; much of the cost savings from smaller buses in that calculation, however, are due to assuming non-union wages for minibus operators, which should not enter the calculation here.

[14] It is interesting to check the equality of non-size-related agency cost, C_1, and waiting time costs, U_x, according to equation (6.10); if the theory applies, this should give the same result for (C_2/C) as the calculation in the text. The equality can be written as $1\text{-}(C_2/C) = U_x/C = (U_x/U_v)\cdot(U_v/C) = (\alpha_x/\alpha_v)\cdot(T_x/T_v)\cdot(U_v/C)$, where T_x and T_v are the average waiting and in-vehicle times, respectively. The ratio α_x/α_v is estimated at 1.6 for Britain by Wardman (2001). U_v/C =2.90 from the previous bullet. Ideally the ratio T_x/T_v should apply to the central area. T_v can be estimated as m/S =19 min, since m=3.48 km (calculated from TfL 2001:37, Table 20a) and S=11 km/hr (TfL 2003c:107). Thus the equality holds if T_x=2.25 min, i.e. headway is 4.5 min. This is about 10 per cent higher than the actual average waiting time of 4.1 min reported by users for all public transport in the central area, including underground (TfL 2001:17, Table 8a).

[15] See Goodwin (1992), Pratt *et al.* (2000, p. 9-14), Savage and Schupp (1997, Table 2) and Kain and Liu (1999).

The third set of parameters describes the very early results of the congestion pricing implementation in London which started in February 2003.

- Change in bus speed (*dS/S*): According to TfL (2003b), central area bus speeds during the morning peak rose 15 per cent (from 10.4 to 12 km/hr) by the end of the second week of implementation, while speeds outside the central area were barely changed. This figure may exaggerate the effect of congestion charging because bus speed had already risen to 11.6 km/hr just before congestion charging began, at least partly due to some street improvements (TfL 2003c:107); on the other hand, it ignores some advantages of reliability improvements that were also documented (TfL 2003c:108-111). Here a more conservative figure is taken by assuming that the speed before congestion charging was 11 km/hr, the average figure for 2002 (TfL 2003c:107); i.e. the increase is from 11 to 12 km/hr, or 9 per cent.

- Modal diversion ($dQ^{(1)}/Q$): During the first few months, modal diversion to all public transport for trips crossing the charging zone boundary was estimated at 110,000 passenger or approximately 3 per cent of current public transport ridership (TfL 2003a:6). The breakdown between bus and rail is not reported, but a rough estimate is possible. Before congestion charging, daily bus and underground ridership to and within the charging zone during the charging period appears to have been about 710 thousand for bus and 2,012 thousand for the underground.[16] Together these figures would imply that other public modes, mainly National Rail with no underground connection, accounted for 944 thousand trips or about one-fourth of all public transport trips, which is similar to its modal share for trips entering central London during the three-hour morning peak period in 2001 (TfL 2003c:130, Figure 6.1). TfL estimates that underground ridership to the central area rose only one per cent and that the change in National Rail ridership, while not yet measured,

(..continued)

[16] For bus, there were 193,000 entering and 162,000 leaving the charging zone in autumn 2002 (TfL 2003c:100); adding 100% for intra-zone trips yields 710,000. For underground, 547,000 passengers arrived within or just outside the charging zone in spring 2002 (TfL 2003c:113); assuming the fraction of these that exited later in the day is the same as for bus (162/193), and adding another 100% for mid-day trips, yields 2,012,000.

was “likely to be comparable” (TfL 2003a:7); thus .01x(2,012+944) = 30 thousand diverted trips are to rail, leaving 80 thousand diverted to bus. This is an 11 per cent increase in bus ridership caused by congestion charging, and may be compared to the 14 per cent increase claimed by TfL (2003a:7) for the single peak hour of 8-9 a.m. However, to some extent these estimates are based on year-to-year comparisons, which could be misleading because substantial upgrades to bus service were implemented during 2002; thus some of the increased bus share is probably due to bus improvements (part of dQ^{SQ} or dQ^{Σ} in this paper) rather than to the direct effects of the road charge. For this reason, the figure assumed here for $dQ^{(1)}/Q$ is 6 per cent, just over half the amount calculated above.

- Subsidy increase from road pricing revenues ($d\Sigma/C$): Expected net toll revenue available for public transport in 2003 is £130 per year (TfL 2003d). It is assumed here to be allocated to bus and underground in proportion to passenger-km travelled on those two modes; thus a fraction 0.387 goes to bus service, based on TfL (2001:37,39). These are then divided by estimated agency costs for 2003 of £687.5 million; the latter is obtained by inflating the £643 million for 2000-01 by two years growth at rate 3.4 per cent per year (based on trends between 1998-99 and 2000-01, from TfL 2001:38, Table 21). The result is approximately $d\Sigma/C$=0.07.

Computational formulae for Table 6.3

Bus-km of service (fractional change): $(dM^{SQ}+dM^{\Sigma})/M$, from (6.22) and (6.23).

Fare (fractional change): df^{SQ}/f, from (6.21).

Ridership (fractional change): $(dQ^{SQ}+dQ^{\Sigma})/Q$, from (6.20) and (6.24).

Average in-vehicle time cost (change as fraction of fare):

(..continued)

$$\frac{d(U_v/Q)}{f} = -1\cdot\frac{dS}{S}\cdot\frac{U_v}{fQ} = -\frac{dS}{S}\cdot\frac{U_v}{C}\cdot\frac{C}{R}\,.$$

Average out-of-vehicle time cost (change as fraction of fare):

Having already computed dM/M, the change in average cost U_o/Q can be computed from the first two terms in (6.3). These show that the two components of U_o/Q are proportional to $(1/N)$ and (N/M), respectively; but substituting (6.7) for N shows that both terms are in fact proportional to $M^{-1/2}$. Thus:

$$\frac{d(U_o/Q)}{f} = -\frac{1}{2}\cdot\frac{dM}{M}\cdot\frac{U_o/Q}{f} = -\frac{1}{2}\cdot\frac{dM}{M}\cdot\frac{2C_1}{R} = -\frac{dM}{M}\cdot\left(1-\frac{C_2}{C}\right)\cdot\frac{C}{R}\,,$$

where the second equality depends on the equality of U_w, U_x, and C_1 as described by (6.10). (Equivalently, one may assume N does not change, in which case U_x is inversely proportional to M while U_w is unchanged, leading to the same result.)

Average agency cost (fractional change):

Differentiating (6.2) with respect to M, S, and Q, we find

$$\frac{dC}{C} = -\frac{dS}{S} + \frac{C_1}{C}\cdot\frac{dM}{M} + \frac{C_2}{C}\cdot\frac{dQ}{Q}\,.$$

Then *average* agency cost satisfies:

$$\frac{d(C/Q)}{C/Q} = \frac{dC}{C} - \frac{dQ}{Q} = -\frac{dS}{S} + \left(1-\frac{C_2}{C}\right)\cdot\left(\frac{dM}{M} - \frac{dQ}{Q}\right).$$

Benefits: from (6.25), (6.26), and (6.27), divided by agency cost.