The Welfare Effects of Congestion Tolls with Heterogeneous Commuters

By Richard Arnott, André de Palma and Robin Lindsey*

1. Introduction

Economists have been recommending congestion tolls as a means of regulating urban traffic congestion for many years (Vickrey, 1963; UK Ministry of Transport, 1964 — the Smeed Report). Until recently, however, the opposition to congestion tolling from politicians and the public appeared overwhelming. An overriding objection was the cost of collection; the queuing cost at tollbooths would have exceeded the benefit from the toll. The famous Hong Kong experiment persuasively demonstrated that electronic road pricing (ERP) is technically feasible. But it was not introduced in Hong Kong because of political opposition. A major concern was that the form of ERP employed would have permitted government monitoring of individuals’ travel (Borins, 1986). That concern has now been addressed through the development of ‘smart cards’, which are completely anonymous and permit instantaneous payment of tolls by making deductions from a prepaid credit balance.

As a result, the focus of discussion on urban road pricing has been shifting from technological considerations towards political acceptability, as Evans (1992) has recently observed. Until recently, it seemed that no amount of reasoned argument would persuade most drivers to give up their ‘right’ to free travel on urban roads, even under conditions of extreme congestion. But late developments suggest that this pessimism was excessive. Three cities in Norway — Bergen, Oslo and Trondheim — have successfully implemented cordon tolls around the CBD for inbound traffic, based on the need to supplement grants from the central government to finance expansion of the urban road network (Ramjerdi, 1992). Singapore is moving beyond its Area Licensing System toward an ERP system using smart-card technology. Private toll roads, already common in Southern Europe, are being constructed in the US. A recent public attitude survey in London found that “Road charges would be acceptable to the majority if the revenues were reinvested in improving roads and public transport, or reducing taxes” (Bayliss, 1992, p.7).

* Richard Arnott is a member of the Department of Economics, Boston College, Massachusetts; André de Palma is a member of the Département COMIN, Université de Genève; Robin Lindsey is a member of the Department of Economics, University of Alberta. Financial support from the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged. André de Palma would also like to thank the Fond National de la Recherche Suisse for support. Finally, thanks are due to Ken Small for helpful comments on an earlier version of the paper, and to the editor and an anonymous referee for suggestions.
These developments are very encouraging since they suggest that urban road pricing will be politically viable when the majority is persuaded that pricing is in its self-interest. This in turn suggests that economists and transport planners should design tolling schemes that benefit all, or almost all, urban travellers. The design of such tolls is the motivation of this paper. What we accomplish is more limited. We consider a population of heterogeneous commuters, who differ in terms of unit costs of travel time and schedule delay (time early and time late) as well as desired arrival time at work, and who commute by car along a common congested urban traffic corridor. We ask who benefits and who is hurt by an optimal tolling system when (i) toll revenues are not returned in cash or services to drivers; (ii) toll revenues are rebated as an equal lump-sum payment to all drivers; and (iii) toll revenues are used to finance expansion of the road. Answers to these questions should provide insight into the design of tolling schemes, including the disposition of toll revenue.

The welfare effects of tolls have been examined in other papers, both theoretical (Layard, 1977; Wigan, 1977; Glazer, 1981; and Evans, 1992) and empirical (Kulash, 1974; Segal and Steinmeier, 1980; and Small, 1983). The majority have concluded that the incidence of congestion tolls is probably regressive, and that distributional effects can dominate efficiency gains. Foster (1974, 1975) reached the tentative conclusion that road pricing is progressive; a dissenting opinion was voiced by Richardson (1974).

The results of these papers must be interpreted carefully because different authors use different definitions of progressivity (see Section 3.3 for a discussion). Moreover, they all employ static models of congestion. They therefore ignore important dynamic features of congestion, notably the fact that tolls can influence drivers’ choice of when to travel, as well as their choice of route and decision whether to travel. Furthermore, different types of individuals will tend to travel at different times during the rush hour and will therefore be differently affected by tolls. For example, an assembly-line worker required to start work at 8.30 am will have to pay a large toll since he travels in a highly congested period. If his value of travel time is low, he may derive little benefit from the toll. In contrast, a professional with a high value of travel time may benefit greatly from the reduction in travel time induced by the toll.

The first dynamic model of vehicle congestion with endogenous driver behaviour was developed by Vickrey (1969) for the morning rush hour. In his model, each commuter decides when to depart from home so as to minimise the sum of travel time and schedule delay. Congestion takes the form of queuing behind a bottleneck. Henderson (1974, 1977, 1981) was the first to incorporate driver heterogeneity into a dynamic model of rush-hour congestion. He treated two groups, differing in the ratio of unit travel time to unit schedule delay costs. Unlike Vickrey, he assumed a form of flow congestion in which a driver’s travel time depends on the density of vehicles she encounters on entering the road. A more general analysis, using the bottleneck model, was provided by Newell (1987), who dealt with continuous distributions of commuters differing in work start time, costs of travel

---

1 Differences in commuting behaviour across socioeconomic groups have been documented in a number of empirical studies (for example, Ott et al., 1980; Small, 1982; and Moore et al., 1984).
time and costs of schedule delay. In previous work (Arnott, de Palma and Lindsey, 1988, 1992) we also incorporated heterogeneous commuters into Vickrey’s model to examine such issues as cost-benefit analysis of road investments and route choice.

The only study we know of to examine the welfare effects of tolls in a dynamic setting is that of Cohen (1987). Like Vickrey, Cohen treats queuing congestion and assumes that desired arrival times of drivers are distributed uniformly over a one-hour interval. Like Henderson, he considers two groups of commuters differing in unit travel time and schedule delay costs, but with the same relative cost of late to early arrival. In this paper we extend Cohen’s analysis in several directions. First, we allow for differences across commuters in the relative costs of late to early arrival. Second, we consider at a general level differences in commuters’ desired arrival times. Third, we analyse the welfare effects of tolls when toll revenue is returned as a lump sum to drivers. Finally, we consider the welfare effects of road investments, both exogenous and those financed through toll revenue.

Evans (1992) has recently addressed some of these same issues. Our model differs from his mainly in that the timing of trips is endogenous while trip demand is inelastic. Several features of Evans’ model have been criticised by Hills (1993). Because our model is dynamic it avoids problems with the treatment of time and the definition of demand and supply. However we consider only a single origin and destination connected by one route, clearly an idealisation of road networks on which road pricing is likely to be implemented. In his reply to Hills (1993), Evans (1993, p.102) argues that models need not be complex to provide useful insights; nor need they be representative of specific networks. We agree. It is not our intention here to provide either a comprehensive or a quantitative assessment of the welfare effects of road pricing, but rather to identify relevant factors that have not previously been documented and may not have been recognised as important.

Since our model has a single route it ignores any effect road pricing may have on route choice (reassignment in Hills’ terminology). In an earlier paper (Arnott et al., 1990b) we modelled two routes running in parallel along a commuting corridor and showed that the efficiency gains from rescheduling trips are likely to dominate gains from reassigning traffic between routes. Hence, for at least some types of simple networks, it may not be crucial to model route choice in studying road pricing. In addition, we showed in Arnott et al. (1993) that the benefits from road pricing in retiming trips may well exceed the benefits from altering the number of trips. Rescheduling of trips can thus yield comparable gains to suppression of trips as a behavioural adaptation to road pricing. In this respect, Evans’ paper which focuses on numbers of trips, and ours which focuses on their timing, are complementary.²

A final remark is worth making on the congestion technology. We limit attention to the bottleneck queuing model, whereas Evans also considers two flow-congestion models and finds that they yield different quantitative estimates of the efficiency gains from pricing. In defence of the bottleneck model we can note that it provides a good fit to traffic

² Generalising the model to incorporate elastic demand is not conceptually difficult, but makes the algebra more tedious. Furthermore, empirical studies (McFadden, 1974; Pucher and Rothenberg, 1976; and Small, 1983) have found that the demand for commuting trips by car, at least in most US cities, is highly inelastic.
flow data on some roads (Small, 1992, pp.69-74) as well as yielding qualitatively sensible results (Small, 1992, pp.88-94).

The paper is organised as follows. In Section 2 we review briefly the Vickrey model, as treated in Cohen (1987) and Arnott et al. (1988, 1990a). We then examine the welfare effects of tolls, capacity expansions, and capacity expansions funded by tolls, for different dimensions of commuter heterogeneity: first with respect to the ratio of unit travel time to unit schedule delay costs (Section 3), then with respect to the ratio of unit time late to time early costs (Section 4), and finally with respect to desired arrival time at work (Section 5). Section 6 contains some concluding remarks.

2. Review of the Model with Identical Commuters

In this section we briefly review the Vickrey model, as adapted in Arnott et al. (1990a). Details are found in Cohen (1987) and Arnott et al. (1988, 1990a) among others. Here we provide only the background necessary to understand later sections.

In the model, \( N \) identical commuters travel each morning, one per car, from home to work along a single road. \( N \) is assumed to be fixed; trip demand is completely inelastic. Travel is uncongested except at a bottleneck with a capacity of \( s \) cars per unit time. If the arrival rate at the bottleneck exceeds \( s \), a queue develops. Travel time is \( T(t) = T^f + T^v(t) \), where \( T^f \) is free-flow travel time, \( T^v(t) \) is waiting time at the bottleneck, and \( t \) is departure time from home. It is assumed without loss of generality that \( T^f = 0 \), so that an individual reaches the queue at the bottleneck as soon as he leaves home, and arrives at work immediately upon exiting the bottleneck. If \( Q(t) \) is the number of vehicles in the queue at time \( t \), then \( T^v(t) = \frac{Q(t)}{s} \). Individuals are assumed to have a preferred arrival time at work, \( t^* \). Their travel cost is given by the linear function

\[
C(t) = \alpha T^v(t) + \beta (\text{time early}) + \gamma (\text{time late}) + \text{toll}(t).
\]

To make the equilibrium departure rate of drivers finite, it is assumed that \( \alpha > \beta \).\(^3\) For individuals who arrive before \( t^* \), (time late) = 0, and for those who arrive after \( t^* \), (time early) = 0. Henceforth, we take ‘depart early’ to mean arrive at work early, and ‘depart late’ to mean arrive at work late.

2.1 No-toll equilibrium

In choosing when to leave home, individuals face a trade-off between trip duration and schedule delay. They are assumed to have full information about the departure time distribution. Equilibrium obtains when no driver can reduce her travel costs by altering her departure time. With identical individuals, this means that costs are constant over the rush hour.

\(^3\) If \( \alpha < \beta \) there is a multiplicity of equilibria in which a mass of individuals leaves at the beginning of the departure period, with positions of individuals in the mass determined randomly. This case is analysed in Arnott et al. (1985). We ignore it here because it is awkward to deal with multiple equilibria; because mass departures with random queue positions do not seem to be prevalent empirically; and because most empirical evidence supports \( \alpha > \beta \) (such as McFadden et al., undated, Small, 1982). Wilson (1988, p.203) does obtain contrary estimates, but his \( \hat{\alpha} - \hat{\beta} \) is not significantly less than zero.
The equilibrium is depicted in Figure 1. The number of vehicles in the queue, \( Q(t) \), is measured by the vertical distance between the cumulative departures and cumulative arrivals schedules. Travel time is measured by the horizontal distance. The queue builds up at a constant rate from \( t_q \) when the first individual leaves, until \( t_q' \), the departure time for which an individual arrives on time (at \( t^* \)). The queue then dissipates, again at a constant rate, reaching zero at \( t_q'' \) when the last person departs. Equilibrium values for \( t_q \) and \( t_q' \) are solved by equating the costs of the first and last individuals to depart, both of whom incur only schedule delay costs:

\[
\beta(t^*-t_q) = \gamma(t_q'-t^*). \tag{2}
\]

Since the bottleneck operates at capacity throughout the rush hour

\[
t_q' = t_q + \frac{N}{s}. \tag{3}
\]

Combining (2) and (3) one obtains

\[
t_q = t^* - \frac{\gamma}{\beta + \gamma} \frac{N}{s}, \tag{4a}
\]

\[
t_q' = t^* + \frac{\beta}{\beta + \gamma} \frac{N}{s}. \tag{4b}
\]

Using (2) it is easy to show that
\[ t_e = t^* - \frac{\delta}{\frac{N}{3}} \]

where \( \delta = \beta \gamma (\beta + \gamma) \). Total travel time, total time early and total time late are identified as areas in Figure 1. Total Travel Time Costs, \( TTC^* \), total Schedule Delay Costs, \( SDC^* \), and total Travel Costs (net of free-flow travel time costs), \( TC^* \), are:

\[ TTC^* = SDC^* = \frac{\delta}{2} \frac{N^2}{3} \]  \hspace{1cm} (5)

\[ TC^* = TTC^* + SDC^* = \delta \frac{N^2}{3} \]  \hspace{1cm} (6)

where the superscript \( e \) denotes no-toll equilibrium.

2.2 The social optimum

The social optimum is determined by minimising the sum of travel time and schedule delay costs. Since queuing is wasteful the departure rate is held at \( s \) throughout the rush hour. The time of first departure is chosen so that the first and last commuters incur equal costs, since otherwise costs could be reduced by moving an individual from the beginning of the rush hour to the end, or vice versa. Since this condition is also true of the no-toll equilibrium, the timing of the rush hour and the arrival distribution are the same as in equilibrium.\(^4\)

With variables corresponding to the social optimum denoted by a superscript \( o \), aggregate costs are:

\[ TTC^o = 0, \]  \hspace{1cm} (7)

\[ SDC^o = TC^o = \frac{\delta}{2} \frac{N^2}{3}. \]  \hspace{1cm} (8)

Comparing (8) with (6), one sees that total costs (net of free-flow travel time costs) are just half their level in the no-toll equilibrium.

2.3 Equilibrium with tolls

Since the number of commuters is fixed and there is only one route, a uniform (time-independent) toll has no effect on equilibrium; it merely extracts income from drivers. However, a time-varying toll does affect queuing. The social optimum can be decentralised by a toll that increases linearly from zero at time \( t_q \) to a maximum at \( t^* \), and subsequently decreases linearly to zero at \( t_q' \). The toll replaces queuing time as the rationing device for desired arrival time slots, leaving drivers' private costs inclusive of the toll unchanged,\(^5\) but overall welfare higher by the amount of toll revenue. The social optimum can also be decentralised by combining the above time-varying toll with a uniform toll. We assume hereafter that the uniform component is zero.

---

\(^4\) With flow, rather than queuing, congestion, the optimal distribution of arrivals differs from the no-toll equilibrium distribution.

\(^5\) This contrasts with static flow-congestion models, in which tolling makes drivers worse off (Glazer, 1981; Evans, 1992).
3. Individuals with Different $\alpha$ and $\beta$, and the Same $\gamma/\beta$

In this section, we allow the unit costs of travel time ($\alpha$) and schedule delay ($\beta$ and $\gamma$) to differ across commuters, but assume commuters have the same relative cost of late to early arrival, $\gamma/\beta$, and the same $r^*$. Cohen’s (1987) specification differs from ours in that he assumes $r^*$ is uniformly distributed in the driving population over a one-hour interval (see also Vickrey, 1969).

Suppose without loss of generality that there are $G$ groups of commuters indexed in order of increasing $\beta/\alpha$, so that group $G$ dislikes schedule delay most highly relative to travel time, while group 1 has the greatest relative aversion to travel time:

$$\frac{\beta_1}{\alpha_1} \leq \frac{\beta_2}{\alpha_2} \leq \ldots \leq \frac{\beta_G}{\alpha_G},$$

We assume $\alpha_i > \beta_i$, $\forall i$. Let $N_i$ be the number of individuals in group $i$, with

$$\sum_{i=1}^{G} N_i = N.$$

3.1 No-toll equilibrium

A derivation of the no-toll equilibrium is given in Arnott et al. (1987); descriptions are found in Cohen (1987) and Arnott et al. (1988). Figure 2 shows equilibrium for three groups. The intuition for the timing and shape of the queuing pattern is as follows. Since departures are continuous over the rush hour, $t_q - t_q^* = N/s$. The first and last individuals to depart are members of group 1, with the lowest relative cost of schedule delay to travel time. These individuals must have the same cost in equilibrium, and since neither faces a queue this is entirely schedule delay cost:

$$\beta_1 (r^* - t_q^*) = \gamma_1 (t_q^* - t_q^*).$$

The equilibrium values of $t_q$ and $t_q^*$ are therefore given by (4a) and (4b) above, with $\beta_1$ and $\gamma_1$ in place of $\beta$ and $\gamma$. Let $\delta_1 = \beta_1 \gamma_1 / (\beta_1 + \gamma_1)$, $i = 1, 2, 3$. Then the equilibrium cost per driver of the first group is

$$C_1^* = \delta_1 N_1.$$

The locus of points in $t-Q$ space on which group 1 drivers incur cost $C_1^*$ is labelled $C_1$ in Figure 2. Cohen (1987) refers to this as an ‘isocost’ contour. Along it group 1’s travel costs, given by (1), are constant:

$$C_1(t) = C_1^* = \begin{cases} 
\beta_1 (r^* - t) + (\alpha_1 - \beta_1)Q(t)/s \text{ for early departures}, \\
\gamma_1 (t - r^*) + (\alpha_1 + \gamma_1)Q(t)/s \text{ for late departures}. 
\end{cases}$$

$Q(t)$ evolves so as to satisfy this condition during the time intervals when group 1 is departing.

Group 2, with the second-lowest relative cost of schedule delay to travel time, is the second group to depart early and the second-last to depart late. The boundary times between group 1 and group 2 departures, $t_{12}$ and $t_{21}$, are determined simultaneously by the conditions that $C_2(t_{12}) = C_2(t_{21})$, and that the number of group 1 departures between $t_q$ and $t_{12}$, plus the number between $t_{21}$ and $t_q^*$, equals the group 1 population. These conditions yield

145
Figure 2
No-Toll Equilibrium for Three Groups of Commuters
with Different $\alpha$ and $\beta$ but the same $\gamma / \beta$ and $t^*$

$$C_2^* = \delta_2 \frac{N_2}{s} + \frac{\alpha_2}{\alpha_1} \delta_1 \frac{N_1}{s}.$$  \hspace{1cm} (10)

The departure interval for group 3 and its equilibrium cost are determined sequentially in a similar manner:

$$C_3^* = \delta_3 \frac{N_3}{s} + \left( \frac{\alpha_3}{\alpha_2} \right) \delta_2 \frac{N_2}{s} + \left( \frac{\alpha_2}{\alpha_1} \right) \delta_1 \frac{N_1}{s}.$$  \hspace{1cm} (11)

Figure 2 displays the equilibrium iso-cost contours for the three groups. To the left of the locus $O t^*$ with slope $-s$, along which individuals arrive on time, the slope of $C_i$ is $\beta_i s / (\alpha_i - \beta_i)$; to the right it is $-\gamma / s (\alpha_i + \gamma)$. The slopes are greater for higher-numbered groups, who are more willing to queue in order to reduce schedule delay.

The equilibrium queue, or Travel Equilibrium Frontier (TEF), is the upper envelope (heavily shaded in Figure 2) of the various groups' equilibrium iso-cost contours. Commuters may be viewed as self-selecting into departure times that minimise their cost on the TEF. The equilibrium departure interval of a group occurs where its equilibrium iso-cost curve coincides with the TEF. Each departure slot is allocated to that group which is willing to pay the most, in terms of queuing time, for it. Group 2, for example, departs between $t_{12}$ and $t_{23}$, and again between $t_{32}$ and $t_{21}$. It is evident that the higher a group's $\beta / \alpha$, the closer to the centre of the rush hour it travels. For early departures each group's departure interval is a connected set, so that there is no mixing of groups. There is also no mixing during late departures.

146
3.2 The social optimum

There is no queuing in the social optimum; thus, groups are allocated to departure time slots so as to minimise schedule delay costs. The group with the largest $\beta$ travels closest to $t^*$, the group with the second-highest $\beta$ departs on adjacent time intervals earlier and later, and so on. The first and last commuters to depart, who belong to the same group, must incur the same schedule delay costs, as is also true in equilibrium. Because all groups have the same $\gamma/\beta$, equations (4a) and (4b) yield the same values for $t_q$ and $t_q^*$ as in equilibrium. However, the order in which groups depart in the social optimum differs from the order in equilibrium unless the ranking of groups according to $\beta$ is the same as the ranking according to $\beta/\alpha$, which need not be the case. For example, assembly-line workers may have a relatively higher $\beta$ than stockbrokers because there is little for assembly-line workers to do before the shift starts, but have a lower absolute $\beta$ than stockbrokers, whose hourly value of time is significantly higher. If these are the only two groups of commuters, then assembly-line workers travel at the peak in the no-toll equilibrium, whereas stockbrokers travel at the peak in the social optimum. Thus, travel order in the no-toll equilibrium is determined by relative schedule delay and travel time costs, while in the social optimum it follows absolute schedule delay costs. If the orders differ, aggregate schedule delay costs in the no-toll equilibrium are not minimised because groups with high unit schedule delay costs (for example, stockbrokers) travel on the tails of the rush hour rather than in the middle.

The social optimum can be decentralised with a time-varying toll. A toll for three groups is drawn in Figure 3, where it is assumed that $\beta_2 < \beta_1 < \beta_3$, so that groups 1 and 2 are reversed in the departure sequence relative to the no-toll equilibrium. The toll coincides with the upper envelope of the toll-equilibrium isocost curves of each group. Since there is no queuing, departure and arrival times coincide, and each group’s curve peaks at $t^*$ (rather than before $t^*$ as in Figure 2). Group $i$’s curve rises with slope $\beta_i$ for $t < t^*$, since its schedule delay cost falls at this rate. For $t > t^*$, the isocost curve has slope $-\gamma_i$.

In the toll equilibrium, group $i$ departs early while the toll is increasing at rate $\beta_i$, and late while it is decreasing at rate $\gamma_i$. Since commuters in a given group prefer their existing departure interval to departure at any other times, the toll induces commuters to self-select into the efficient departure time intervals. Because the toll offsets the change in schedule delay costs of the group departing, there is no queuing.

Since calculation in the general case is tedious, for the remainder of the section we shall focus on two groups, 1 and 2. Substituting $N_3 = 0$ into (9) and (10) one has

$$C_1' = \delta_1 \frac{N_1 + N_2}{\delta}, \quad C_2' = \delta_2 \frac{N_2}{\delta} + \left( \frac{\alpha_2}{\alpha_1} \right) \delta_1 \frac{N_1}{\delta}. \quad (12)$$

Total costs are $TC = N_1 C_1' + N_2 C_2'$. Total schedule delay and total travel time costs are straightforward to compute. Let $f_2 = N_2/N$ be the fraction of commuters in group 2. Further, define $\alpha_r = \alpha_2/\alpha_1$ and $\beta_r = \beta_2/\beta_1$. Aggregate costs in the no-toll equilibrium work out to be:\footnote{See Amott et al. (1987).}
Figure 3
Optimal Time-Varying Toll for Three Groups of Commuters
with Different $\alpha$ and $\beta$ but the Same $\gamma/\beta$ and $t^*$

\[
SDC^* = \frac{\delta_1}{2} \frac{N^2}{s} \left[ 1 + (\beta_r - 1) \frac{\beta_r}{s^2} \right],
\]
(13)

\[
TTC^* = \frac{\delta_1}{2} \frac{N^2}{s} \left[ 1 - 2(1 - \alpha_r) \frac{\beta_r}{s^2} + (1 + \beta_r - 2\alpha_r) \frac{\beta_r^2}{s^2} \right],
\]
(14)

\[
TC^* = \frac{\delta_1}{2} \frac{N^2}{s} \left[ 1 + (\alpha_r - 1) \frac{\beta_r}{s^2} + (\beta_r - \alpha_r) \frac{\beta_r^2}{s^2} \right].
\]
(15)

In the social optimum, group 2 travels in the middle of the rush hour if $\beta_r > 1$, and on the tails if $\beta_r < 1$. Total costs can be derived geometrically from Figure 3 using a procedure similar to that outlined above. The result is

\[
SDC^* = TC^* = \begin{cases} 
\frac{\delta_1 N^2}{2s} \left[ 1 + (\beta_r - 1) \frac{\beta_r}{s^2} \right], & \text{if } \beta_r \geq 1; \\
\frac{\delta_1 N^2}{2s} \left[ 1 - 2(1 - \beta_r) \frac{\beta_r}{s^2} + (1 - \beta_r) \frac{\beta_r^2}{s^2} \right], & \text{if } \beta_r \leq 1.
\end{cases}
\]
(16a)

3.3 Welfare
3.3.1 Welfare effects of the optimal time-varying toll
In sub-section 2.3 we saw that if commuters are identical their welfare is unaffected by the optimal time-varying toll because it substitutes perfectly for queueing time. If commuters differ, intuition suggests that those who value congestion relief highly benefit
Table 1

Welfare Effects of the Optimal Time-Varying Toll on Two Groups of Commuters with Different $\alpha$ and $\beta$ but the Same $\gamma\beta$ and $t^*$

$$(\beta_1/\alpha_1 < \beta_2/\alpha_2)$$

(a) No rebate of toll revenue

<table>
<thead>
<tr>
<th></th>
<th>Order of departure preserved $(\beta_2 &gt; \beta_1)$</th>
<th>Order of departure reversed $(\beta_2 &lt; \beta_1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Group 2</td>
<td>$- \text{ if } \alpha_2 &lt; \alpha_1$</td>
<td>$- \text{ if } \alpha_2 &gt; \alpha_1$</td>
</tr>
</tbody>
</table>

(b) Equal per capita lump-sum rebate of toll revenue

<table>
<thead>
<tr>
<th></th>
<th>Order of departure preserved $(\beta_2 &gt; \beta_1)$</th>
<th>Order of departure reversed $(\beta_2 &lt; \beta_1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>$+$</td>
<td>$+$</td>
</tr>
<tr>
<td>Group 2</td>
<td>$+$</td>
<td>$+$ unless $f_2$ and $\alpha_2/\alpha_1$ both small</td>
</tr>
</tbody>
</table>

Note: $+$ means better off, $-$ means worse off, 0 means no change.

most from tolls. If there is no rebate of toll revenue, some may gain and others lose. With a uniform per capita lump-sum rebate, it is more likely that everyone benefits.

The toll considered is that of Figure 3. Table 1 shows that its welfare distributional effects depend not only on whether revenue is rebated, but also on whether departure order gets reversed by the toll.

Consider panel (a) of Table 1, where there is no toll rebate, and the case $\beta_2 > \beta_1$ for which the order of departure is unaffected. Group 1 is equally well off with the toll; instead of incurring a queuing cost of $\alpha_1 Q(t)$ in the no-toll equilibrium, a member departing at $t$ pays a toll of the same amount. Group 2’s costs, however, are affected. To see this, consider the group 2 individual who departs at $t_{12}$. He pays a toll equal to $\alpha_1 Q(t_{12})$ rather than face a queue of length $Q(t_{12})$, the cost of which is $\alpha_2 Q(t_{12})$ to him. Group 2 ends up worse off if $\alpha_2 < \alpha_1$, and better off if $\alpha_2 > \alpha_1$. If group 1 comprises professionals, and group 2 assembly-line workers, $\alpha_2 < \alpha_1$ is likely. But if group 1 comprises office workers, $\alpha_2 > \alpha_1$ may be true.

If $\beta_2 < \beta_1$, group 1 travels in the peak of the rush hour. Any member is free to depart at the beginning of the rush hour, and to incur the same travel cost as in the no-toll equilibrium. But the toll rises only at a rate $\beta_2$, while schedule delay costs for group 1 fall
at rate $\beta_1 > \beta_2$. Delaying departure therefore makes group 1 strictly better off. Group 2, however, is worse off. In the no-toll equilibrium, it ‘outbids’ group 1 for the choice departure time slots because of its relatively low value of travel time and hence tolerance to queuing. But with the toll, willingness to pay becomes the rationing device, and group 2 is forced to the tails of the rush hour.

Suppose now, as in panel (b) of Table 1, that toll revenues are returned as an equal lump sum to all drivers. While group 1 is always better off with the toll, group 2 is not. In particular, if group 2 is a minority ($f_2$ is small) and relatively insensitive to travel time costs ($\alpha_2/\alpha_1$ is small), its travel costs are so low when travelling near the peak in the no-toll equilibrium that it is worse off with the toll even with a rebate.

Overall, Table 1 suggests that the group with a lower $\beta/\alpha$ fares better with a toll than the group with a higher $\beta/\alpha$. Whether or not the toll is progressive or regressive depends on the definition of progressivity. There are three commonly-employed notions of progressivity — absolute progression, average progression, and marginal progression. A tax (a bad) is said to be absolute progressive if tax payable rises with income, average progressive if the average tax rate rises with income, and marginal progressive if the marginal tax rate rises with income. A benefit is said to be absolute progressive if the benefit falls with income, and so on.

Because our analysis makes no reference to income, we must employ an alternative definition. We shall say that a toll is progressive if the net benefit falls with the (relevant) unit shadow value of time. (If the shadow value varies positively with income, then our definition of progressivity is consistent with the standard definition.) According to this definition, if group 1 comprises highly-paid white-collar workers and group 2 blue-collar workers, so that $\alpha_1 > \alpha_2$, the toll considered here tends to be regressive.

This is broadly consistent with Evans (1992) who finds that individuals with high values of time fare better with tolling than those with low values. (This is also true if individuals with above average values of time have relatively inelastic trip demand.) Small (1983) also finds that, exclusive of redistributed revenues, tolls favour higher-income groups. Lower-income users are disadvantaged by having to share roads with those who value congestion relief highly. With revenue redistribution, Small finds that only the poorest individuals still end up worse off.

Segal and Steinmeier (1980) also find tolls to be regressive. They assume that toll revenue is rebated in proportion to income, rather than on a uniform per capita basis. A proportional rebate would reinforce our conclusion that tolls tend to be regressive. Henderson (1974) assumes that high-income workers have a higher $\beta$, higher $\alpha$ and higher $\beta/\alpha$ than low-income workers. High-income workers then travel in the middle of the rush hour in both the no-toll and with-toll equilibria. Low-income workers are unaffected by a toll without rebate, while high-income workers are better off.\(^7\)

\(^7\) Since the toll generates efficiency gains, if the groups were identifiable it would be possible to design a rebate scheme that would result in everyone being made better off. Though individuals’ $\alpha$, $\beta$ and $\gamma$ are not directly observable, if they were strongly correlated with income there would be some scope for differentiating the rebate via the income tax system. An alternative approach to designing a tax-cum-rebate scheme that benefits everyone would be to provide a uniform rebate along with a Pareto-improving time-varying toll.
3.3.2 Welfare effects of capacity expansion

In both the no-toll equilibrium and the social optimum, total costs are inversely proportional to capacity (namely, equations (15) and (16)). The same is true for individual costs in each group. Thus, individual benefits from capacity expansion are proportional to their travel costs. For conciseness we limit attention here to the no-toll equilibrium. Given equation (12) the costs of a group 2 commuter exceed those of a group 1 commuter if and only if:

$$f_2 \beta \alpha + (1-f_2) \alpha > 1.$$  \hspace{1cm} (17)

To see this, consider first the limit $f_2 \uparrow 1$. The first member of group 2 then departs immediately after the last early member of group 1, just after $t_e$. Both incur only a schedule delay cost: $\beta(t^* - t_d)$ for the group 1 member, $\beta_2(t^* - t_d)$ for the group 2 member. Group 2's cost thus exceeds group 1's iff $\beta > 1$. This is the limiting condition of (17) as $f_2 \uparrow 1$.

Now let $f_2 \downarrow 0$. Members of both groups then arrive at $t^*$, and incur only travel time costs. Group 2's cost exceeds group 1's iff $\alpha > 1$, which is the limiting condition of (17) as $f_2 \downarrow 0$.

The general case of (17) is simply a weighted average of the limit cases, with the fractions of individuals in each group as the weights. If group 1 comprises highly paid white-collar workers and group 2 blue-collar workers, then $\alpha < 1$. For downtown areas in which white-collar workers predominate ($f_2$ is small), condition (17) is likely to fail, so that road investments tend to benefit higher-paid workers more because their travel costs are higher and thus fall absolutely more with a capacity expansion.

Before concluding that road investments have regressive welfare effects, however, it is necessary to consider how capacity expansions are financed. One possibility is through general tax revenue; if the tax system is strongly progressive at the margin, the progressive tax effects might offset the regressivity of benefits. Investments could also be financed by a toll, a possibility that we now examine.

3.3.3 Welfare effects of a toll-financed capacity expansion

In the introduction we mentioned recent developments which suggest that tolling is more politically acceptable when toll revenues are used to finance transport improvements. Thus, it is of practical interest to consider the welfare effects of a toll when the revenue is spent on capacity expansion. Consider the case where departure order is unchanged by the toll. Let $s^b$ and $s^a$ denote capacity before and after expansion respectively. Then, given (9),

$$C_1^b - C_1^a = \delta_1 N_{1,5}^b - \delta_1 N_{1,5}^a.$$  

It is straightforward to show (Arnott et al., 1987, Appendices A and B) that with capacity $s^b$ and no toll

$$C_2^b = \delta_2 N_{5,5} [\alpha(1-f_2) + \beta, f_2],$$

whereas with capacity $s^a$ and the optimal toll without rebate

$$C_2^a = \delta_2 N_{5,5} [1-f_2 + \beta, f_2].$$
Thus
\[ C_2^e - C_2^o = \delta \frac{N}{N_e} [\alpha_r(1 - f_2) + \beta r_2 f_2] - \delta \frac{N}{N_e} [1 - f_2 + \beta r_2 f_2]. \]
Revenue raised from the optimal toll can be calculated as \( R = N_1 C_1^{e} + N_2 C_2^{e} - TC^{o} \) \((TC^{o})\) is given by (16). If unit capacity costs are \( k, s^e = s^2 = R/k \). Combining the results yields a complicated condition under which group 2 is made better off when the optimal toll is used to finance capacity expansion. If initial capacity is sufficiently suboptimal from group 2's perspective, the toll/capacity expansion scheme unambiguously benefits group 2, even if \( \alpha_2 < \alpha_1 \). Of course, group 1 always benefits because a toll with neither rebate nor capacity expansion leaves its costs unchanged.

If departure order is reversed by the toll, group 1 is again better off because it gains even without a toll rebate. Group 2, which is worse off with only a toll, gains if the cost of capacity expansion is sufficiently small.

4. Individuals Differing with Respect to \( \gamma \)

4.1 No-toll equilibrium and social optimum
In this section we allow the cost of late arrival, \( \gamma \), to differ across commuters. For simplicity, all other parameters are assumed to be the same for everyone. Let there be \( G \) groups of commuters, indexed in order of decreasing \( \gamma \) so that: \( \gamma_1 > \gamma_2 > ... > \gamma_G \). The derivation and intuition behind this case is similar to that for heterogeneity with respect to \( \alpha \) and \( \beta \); accordingly we will be brief.

4.1.1 No-toll equilibrium
In the no-toll equilibrium, groups with the lowest \( \gamma \) depart late, in strict sequence of decreasing \( \gamma \), with group \( G \) departing last. Groups with higher \( \gamma \) depart early. Their order of departure is indeterminate.

4.1.2 The social optimum
In the social optimum, the departure rate is held at \( s \) to prevent queuing. To minimise schedule delay costs, groups depart in strict sequence of decreasing \( \gamma \) during late departure, just as in the no-toll equilibrium. The timing of the rush hour is also the same as in equilibrium.\(^8\) The optimum can be decentralised with a time-varying toll. Since schedule delay costs are minimised in the no-toll equilibrium, toll benefits equal travel time saved.

4.2 Welfare

4.2.1 Welfare effects of the optimal time-varying toll
Since the order of departure and the timing of the rush hour are unaffected by imposing an optimal time-varying toll, and since toll revenue equals travel time costs without a toll, a toll without rebate is welfare neutral. With a rebate, all commuters are better off. These results apply to any number of groups.

\(^8\) This result, which is not intuitive, is proved in Arnott et al. (1987, Appendix D).
4.2.2 *Welfare effects of capacity expansion*

The welfare effects of capacity expansion are not as immediate as those for the toll, and we simplify as in Section 3 by considering two groups. Group 1, with relatively high penalties for late arrival at work, can be thought of as assembly-line workers. Group 2, with the smaller \( \gamma \), can be thought of as clerks or office workers, provided they work largely independently of each other.

The welfare effects of capacity expansion depend on the relative sizes of the two groups. If group 2 is sufficiently large that only it departs late, then the equilibrium is the same as if all individuals belonged to group 2. (All of group 1 arrives early, and thus it is indistinguishable from group 2.) This is the case when

\[
\frac{f_2}{f} > \frac{\beta}{\beta + \gamma_2},
\]

(18)

where \( \frac{\beta}{\beta + \gamma_2} \) is the fraction of drivers who depart late when all belong to group 2 (see equation (4b)). Individuals in the two groups incur the same travel costs, and *road investment is welfare neutral*.

If condition (18) fails, so that some members of group 1 depart late, group 1 incurs the higher travel costs. To see this, note that the last member of group 1 to depart and the first member of group 2 have the same travel time and schedule delay, but the group 1 individual incurs a greater cost for time late. Group 1 thus benefits more from road investment. If group 1 comprises blue-collar workers and group 2 white-collar workers, *road investment favours the commuting 'poor'*.

4.2.3 *Welfare effects of a toll-financed capacity expansion*

The welfare effects of a toll-financed investment follow immediately from the results of the two preceding subsections. Because a toll without rebate leaves the welfare of both groups unchanged, whereas a capacity expansion benefits them both, a toll-financed capacity expansion leaves both groups better off. If condition (18) is met, then expansion by itself is welfare neutral, as is a toll-financed expansion. If (18) is not satisfied then group 1 (tentatively identified as the commuting 'poor') benefits more.

5. *Individuals with Different t*

5.1 *No-toll equilibrium and social optimum*

Suppose now that commuters differ only in their desired arrival times at work. Let \( N_i \) be the number of commuters with desired arrival time \( t^*_i, i = 1 \ldots G, \) with \( t^*_1 < t^*_2 < \ldots < t^*_G \), so that group 1 has the earliest desired arrival time, and so on. It is assumed that desired arrival times are sufficiently close that the bottleneck operates at capacity throughout the rush hour. Groups can comprise worker teams with different work hours, or teams with identical work hours but employed at varying distances from the bottleneck (teams with further to travel wish to pass through the bottleneck earlier).

---

9. The two-group case incorporates the solution for any number of groups as long as no more than two groups depart late.
Equilibrium for Two Groups of Commuters with Different Desired Arrival Times: Queue Double-Peaked

5.1.1 No-toll equilibrium
An example of equilibrium with two groups is shown in Figure 4. Group 1 wishes to arrive at $t_1^r$, and group 2 at $t_2^r$. The cumulative desired arrival time distribution, $W(t)$, is shown by the heavy line OFGIJKD. Cumulative departures are OABCD and cumulative arrivals OD. The equilibrium isocost curves of the two groups, $C_1$ and $C_2$, intersect at $t_{12}$ when group 1 stops departing and group 2 starts. The departure time for which an individual arrives at $t_1^r$ is denoted by $t_{n,1}$. For departures between $t_1$ and $t_{n,1}$, cumulative arrivals exceed cumulative desired arrivals, and the departure rate exceeds capacity. Between $t_{n,1}$ and $t_{12}$, cumulative arrivals lag desired arrivals, and the departure rate falls below capacity, and so on. The queue peaks at $t_{n,1}$ and $t_{n,2}$, and has a local minimum at $t_{12}$. Aggregate travel time, early arrival time and late arrival time correspond to the areas indicated.

Figure 5 shows another possible equilibrium for two groups in which the queue has a single peak. All of group 2 arrives late. The equilibrium isocost curves $C_1$ and $C_2$ coincide over the interval $[t_{n,2}, t_q^-]$. Members of group 1 and group 2 may depart at any time in this interval; the order of departure is indeterminate. Because group 1 is indifferent about departing at any time in the interval $[t_q^-, t_q^+]$, the timing of the rush hour and the costs of group 1 are the same as if all travellers belonged to group 1. For group 2, late arrival (and total travel) costs are smaller than for the same number of group 1 individuals by the area EFGH. Per capita travel costs are smaller by $\gamma(t_q^--t_q^+)$. A single-peaked queue also arises when all of group 1 arrives early. The timing of the rush hour and the costs of group 2 are the same as if all travellers belonged to group 2.
Welfare Effects of Congestion Tolls with Heterogeneous Commuters

\[ \text{Figure 5} \]

*Equilibrium for Two Groups of Commuters with Different Desired Arrival Times: Single Peak with all Group 2 Arriving Late*

It can be shown (Arnott et al., 1987, Appendix E) that the queue is single-peaked if and only if

\[ s(t_2^* - t_1^*) \leq \left| \frac{\beta}{\beta + \gamma} N_1 - \frac{\gamma}{\beta + \gamma} N_2 \right|, \]  

(19)

and double-peaked if and only if

\[ \left| \frac{\beta}{\beta + \gamma} N_1 - \frac{\gamma}{\beta + \gamma} N_2 \right| < s(t_2^* - t_1^*) < \frac{\beta}{\beta + \gamma} N_1 + \frac{\gamma}{\beta + \gamma} N_2. \]  

(20)

The rush hours of the two groups are separate if and only if

\[ s(t_2^* - t_1^*) \geq \frac{\beta}{\beta + \gamma} N_1 + \frac{\gamma}{\beta + \gamma} N_2. \]  

(21)

Condition (21) holds if the difference in desired arrival times, \( t_2^* - t_1^* \), is greater than a weighted average of the amounts of time required for the two groups to arrive, \( N_1/s \) and \( N_2/s \).

5.1.2 The social optimum

Because groups share the same parameters except \( t^* \), having groups depart in strict sequence is optimal, although not necessarily the unique optimum. The timing of the rush
hour turns out to be the same as in equilibrium (see Arnott et al., 1987, Appendix E). So is the indeterminacy in the order of departure. To see this, note that if two individuals in different groups who are both arriving early remain early if interchanged in the departure sequence, or if they are both arriving late and remain late if interchanged, then total schedule delay costs are unaffected. These are the conditions under which indeterminacy arises in equilibrium.

5.2 Welfare

5.2.1 Welfare effects of the optimal time-varying toll
Since the timing of the rush hour and the order of departure in equilibrium are socially optimal, and $\alpha$, $\beta$ and $\gamma$ are the same for everyone, a toll simply replaces queuing time cost with money, leaving the costs of both groups unchanged. If toll revenue is rebated, both groups are better off. These results apply to any number of groups. Thus, tolls are welfare neutral.

5.2.2 Welfare effects of capacity expansion
As in Sections 3 and 4, the welfare effects of capacity expansion are not transparent, and so we shall again limit consideration to two groups. If the queue is single-peaked (condition (19) is satisfied) before and after expansion the two groups benefit equally because the difference in their travel costs is independent of capacity. If the queue is double-peaked (condition (20) is satisfied) before and after expansion, group 2 turns out to benefit more than group 1 if

$$k > \frac{\beta}{\beta + \gamma}$$

Group 1 benefits more if $f_2 < \beta/(\beta + \gamma)$, and the groups benefit equally if $f_2 = \beta/(\beta + \gamma)$. Because conditions (19) - (21) involve $s$, a capacity expansion may change the shape of the queuing pattern. A single-peaked queue may be transformed into one with two peaks, or two separate rush hours. Similarly, a double-peaked queue may evolve to two separate rush hours. The welfare rankings are readily generalised to allow for such changes:

1. If the queue is single-peaked after capacity expansion, then the two groups benefit equally.

2. If, after expansion, the queue is double-peaked or there are two separate rush hours, then group 2 benefits more than group 1 if and only if condition (22) holds.

Since $\gamma/\beta$ is typically large this suggests that individuals with later work start times (namely, group 2) benefit more if they comprise more than a modest fraction of the commuting population. Whether this means that the welfare effect of investment is progressive depends on how the income distribution of commuters varies with work hours, an empirical issue to be resolved on a city-specific basis.

5.2.3 Welfare effects of a toll-financed capacity expansion
As in Section 4, the welfare effects of a toll-investment package can be decomposed. The toll without rebate by itself is welfare-neutral. The welfare effect of investment depends on the shape of the queuing pattern after investment, as described in subsection 5.2.2. A toll-financed capacity expansion affects welfare in the same way.
### Table 2

**Distributional Effects of a Time-Varying Toll and Capacity Expansion on Two Groups**

<table>
<thead>
<tr>
<th>Variable parameters</th>
<th>Time-varying toll without rebate</th>
<th>Time-varying toll with lump-sum rebate</th>
<th>Capacity expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>None (homogeneous commuters)</td>
<td>Neutral</td>
<td>Everyone benefits equally</td>
<td>Everyone benefits equally</td>
</tr>
<tr>
<td>$\alpha$ and $\beta$ ($\gamma\beta$ fixed)</td>
<td>Typically favours drivers travelling in tails of rush hour</td>
<td>Typically favours drivers travelling in tails of rush hour</td>
<td>Typically favours drivers travelling in tails of rush hour</td>
</tr>
<tr>
<td>$\gamma\beta$</td>
<td>Neutral</td>
<td>Everyone benefits equally</td>
<td>Everyone benefits equally or favours drivers travelling at peak</td>
</tr>
<tr>
<td>$r^*$</td>
<td>Neutral</td>
<td>Everyone benefits equally</td>
<td>Everyone benefits equally or favours drivers travelling at peak</td>
</tr>
</tbody>
</table>

### 5.3 Generalisations

Table 2 summarises the qualitative results derived in the previous three sections. We have treated only cases in which groups differ in one or two parameters. In reality, groups generally differ in all parameters. Studying this hybrid case would generate little additional insight but is necessary for practical application. Analytical solution is feasible up to a point, but tedious, as Newell (1987) has shown.\(^\text{11}\) Furthermore, it is evident that the qualitative characteristics of the equilibrium and optimum are sensitive to the cost parameters of different commuter groups, and to the proportion of commuters in each group. Further empirical work along the lines of Small (1982, 1983) is needed to refine parameter estimates. Also, city-specific studies are required to measure the composition of the commuting population.

---

\(^{11}\) Whereas in all cases exposited here the equilibrium and optimal arrival time intervals coincide, this is not true in general, a complication that adds to the analytical complexity. Richter et al. (1990) have developed fixed point algorithms for computing equilibria numerically.
6. Concluding Remarks

We have analysed the welfare effects of tolls, capacity investments, and toll-financed capacity investments on commuters who differ in their cost of travel time, their preferred arrival time at work, and the costs they incur from early and late arrival. In the absence of tolling, departure times are rationed by queuing. The order in which different groups of drivers depart is based on the ratio of unit schedule delay costs to unit travel time costs, the relative costs of late to early arrival, and work start times.

A time-varying toll can always be designed that eliminates queuing. With such a toll, the order of departure is based on absolute unit schedule delay costs. Groups that value on-time arrival most incur the least schedule delay. The elimination of queuing benefits drivers with a higher unit value of travel time, and hurts those with lower values. If the toll changes the order of departure, it benefits drivers with higher unit schedule delay costs. The toll is neutral with respect to differences in relative late arrival to early arrival costs, and differences in work start times.

Overall, therefore, a toll without rebate benefits drivers with high unit travel time and schedule delay costs, and harms those with low values. Since unit travel time costs are typically strongly positively correlated with income, a toll without rebate tends to benefit the rich on average, and hurt the poor.

An equal per capita toll rebate is (average) progressive since a fixed payment results in a larger proportional increase in income for the poor than for the rich. The overall progressivity of a toll with a uniform rebate is therefore determined by the balance between the typical regressivity of the toll without rebate, and the progressivity of the rebate. Which dominates depends in a complex way on the distribution of the population by work start time, the shadow value of queuing time and the unit costs of schedule delay.

The distributional effects of capacity expansion in our model are simple. Because each group's travel costs (net of fixed travel time) are inversely proportional to capacity, a capacity expansion reduces everyone's costs by the same proportion. Hence those with higher absolute travel costs enjoy higher absolute benefits. In the case of a toll-financed capacity expansion, the analytics of the general case are complex but the intuition is clear. The effects are the same as those for a toll with an effective rebate proportional to (net of fixed) travel costs.

In the introduction we mentioned some encouraging recent developments suggesting that road pricing schemes may well be politically acceptable if they can be designed in such a way that a large majority of the population benefits. This points to the need for research on the design of such schemes, which are defined by a time-varying toll, a cash rebate formula, and a size of capacity expansion. Our analysis was positive. But it can be adapted for normative analysis to solve for the road pricing scheme that is optimal when considerations of acceptability are taken into account. The objective could be to maximise a social welfare function subject to a political acceptability constraint, or to maximise political acceptability.

A number of practical complications absent from our model will need to be considered
in the design of an actual road pricing scheme—public transport, car-pooling, discretionary travel, and differences in origins and destinations, among others.

To analyse car-pooling properly, it will be necessary to model which individuals travel together, and how differences in desired arrival time are resolved. Car-pooling can affect both the physical and psychic costs of travel time and schedule delay. For example, Small (1982) found that car-poolers value schedule delay more relative to travel time than do single drivers. Poorer workers tend to pool cars more than affluent workers, and to use public transport in greater proportions. Car-poolers and bus users benefit from tolls because they occupy less road space than solo drivers, and hence pay lower charges. Bus users also benefit more to the extent that congestion slows buses more than cars. These factors tend to make road pricing less regressive.

Non-commuting travel was ignored in our model. Since distance driven tends to rise with income, the rich would pay more with road tolls, particularly for non-work-related trips, if charges are levied at off-peak hours. However, the rich benefit more from road investments. And tolls tend to reduce travel by those with low values of time, and encourage travel by those who value congestion relief highly, making the welfare impact more likely to be regressive. The effect of demand elasticity has been considered in static models by Layard (1977), Glazer (1981), Niskanen (1987) and Evans (1992).

More sophisticated and detailed estimation of travel cost functions for different population groups will be needed, in order to determine accurately the distributional effects of a particular road pricing scheme. Finally, more research is needed on the determinants of the perceived benefits of alternative road pricing schemes, and hence of their political acceptability.

Economists are largely to blame for the political unreceptiveness to road pricing. They have done a poor job of educating politicians and the public about the theoretical virtues of road pricing. In addition, and at least as importantly, they have paid little attention to considerations of political acceptability and have been content to let others worry about the practical details of implementation. If they pursue the research programme sketched above with diligence, intelligence and sensitivity, they should be able to mount a far stronger case.

---

12 The linear travel cost function used here and elsewhere in the literature is not based on a production technology. A model incorporating workplace interactions is needed to explain why employers prefer employees to keep the same work hours, and to calculate the benefits of flexible time or staggered work hours. As Henderson (1981) noted, employers prefer their most productive and interactive workers to travel in the middle of the rush hour. Neglect of this consideration could lead to poor predictions as to which groups of commuters travel at the rush-hour peak.

13 We have limited consideration throughout to road pricing on public roads. Privately-operated toll roads are common in Europe and the Pacific Rim, and a number are at the planning or construction stages in the US. The efficiency and welfare distribution effects of private-sector road pricing have been the subject of some study, as in Vickrey (1968), Edelson (1971), Nichols et al. (1971), DeVany and Saving (1980), and Poole (1988), but have yet to be considered in a dynamic equilibrium framework.
References


Welfare Effects of Congestion Tolls with Heterogeneous Commuters


*Date of receipt of final manuscript: August 1993*