



Munich Personal RePEc Archive

How a fast lane may replace a congestion toll

Fosgerau, Mogens

Technical University of Denmark

2011

Online at <https://mpra.ub.uni-muenchen.de/42271/>
MPRA Paper No. 42271, posted 30 Oct 2012 18:57 UTC

How a fast lane may replace a congestion toll

Mogens Fosgerau^{*†}

March 10, 2011

Abstract

This paper considers a congested bottleneck. A fast lane reserves a more than proportional share of capacity to a designated group of travellers. Travellers are otherwise identical and other travellers can use the reserved capacity when it would otherwise be idle. The paper shows that such a fast lane is always Pareto improving under Nash equilibrium in arrival times at the bottleneck and inelastic demand. It can replicate the arrival schedule and queueing outcomes of a toll that optimally charges a constant toll during part of the demand peak. Within some bounds, the fast lane scheme is still welfare improving when demand is elastic.

Keywords: congestion; tolling; bottleneck; scheduling; fast lane

1 Introduction

This paper considers a fast lane scheme as a means to regulate congestion in a regularly occurring demand peak. The fast lane scheme plays explicitly on the dynamics of congestion, which makes the [Vickrey \(1969\)](#) bottleneck model an appropriate framework. The elements of the basic bottleneck model are a description of the queueing technology in the bottleneck, a continuum of

^{*}Technical University of Denmark (mf@transport.dtu.dk), Centre for Transport Studies, Sweden and Ecole Normale Supérieure de Cachan, France.

[†]This research is supported by the Danish Social Science Research Council. In addition, support is gratefully acknowledged from PREDIT and ADEME grants to project "Tarification des Transports Individuels et Collectifs à Paris: Dynamique de l'acceptabilité". The opinions and conclusions expressed in this paper are those of the authors and do not necessarily represent those of the Ministère de l'Écologie, de l'Énergie, du Développement durable et de la Mer. This paper has been screened to make sure that no confidential information has been disclosed by the authors.

identical travellers with scheduling preferences who have to pass the bottleneck, and the concept of Nash equilibrium in arrival times at the bottleneck.

The fast lane scheme allocates the bottleneck capacity to different classes of travellers. The scheme is the following.

A set of travellers is assigned to a priority group. Not all travellers can be given priority. A more than proportional share of capacity is reserved for the priority group. When the reserved capacity is not used, it is available for the nonprioritized travellers.

This is similar to, e.g., the check-in in airports with separate queues and servers for economy and business class passengers. Whenever the business class server is idle, it may serve passengers from the economy class queue. Another example is the HOV or HOT lanes as found on US motorways. Yet another example is the flows at different motorway on-ramps that could be given different priority (Shen and Zhang (2010) consider such a scheme). Even though such a scheme it is called a fast lane scheme in this paper, the definition encompasses many other policies that do not involve the allocation of road lanes for different classes of vehicles; it is more general than allocation of lanes.

The paper compares the fast lane scheme to tolling. Like Arnott et al. (1990), this paper considers a coarse toll, which is a constant toll that applies only during part of the peak.¹ Arnott et al. (1990) found that Nash equilibrium under a coarse toll comprises a point mass in the arrival schedule at the time when the toll is lifted. This is an undesirable feature of their model as such point masses are physically implausible. The problem is avoided in this paper by a reformulation of the queueing technology. In this paper, the congestion technology is such that travellers who choose not to pay the toll can queue at the same time as travellers who are paying the toll pass the bottleneck. This is also true of the examples of fast lanes given above. In this case, a point mass in the arrival schedule does not arise. The analysis below uses the reformulated queueing technology and repeats the Arnott et al. (1990) analysis of a coarse toll under this assumption.

The first main result of this paper is that the fast lane scheme is always Pareto improving when demand is not price sensitive. There are no restrictions on how large the group of prioritized travellers should be as long as it is fixed exogenously. Prioritized travellers experience a strict utility gain while the properties of Nash equilibrium imply that nonprioritized travellers do not lose. It is significant that this occurs even when travellers are homogenous and there are no toll payments. This robustness is very desirable since it

¹Arnott et al. (1990) applied also a base toll level. This paper considers the coarse toll only under inelastic demand and so the base toll level does not matter.

means that a regulator needs little information to implement the scheme and be certain to achieve a welfare improvement. In fact, the regulator can monitor traffic in real time and assign capacity accordingly. This is consistent with the way the fast lane scheme is formulated in the model. With price sensitive demand, the fast lane scheme is still welfare improving if the price elasticity of demand is not too high and the share of prioritized travellers is not too large.

The second main result of this paper is that the fast lane scheme can reproduce the equilibrium arrival pattern of the optimal coarse toll when demand is not price sensitive.² In fact, the scheme can reproduce the equilibrium arrival pattern of any coarse toll, provided that the tolling interval is the same as the arrival interval that a prioritized group would endogenously select. This is significant since the fast lane scheme has a number of advantages over tolling. First, the fast lane scheme is always welfare improving and can be adjusted in real time. In order to set the right coarse toll it is necessary to know exactly when to start and when to end the tolling interval. Mistakes will reduce the welfare gain from tolling and can even lead to a welfare loss. Second, it is plausible that system costs can be a lot lower for a fast lane scheme than for a toll as the fast lane does not involve any payment. Finally, as there is no payment, a fast lane scheme may be more acceptable to travellers than tolling. Within the simple theoretical model presented here, prioritized travellers would be strictly better off under the fast lane scheme than under no policy, while the remaining travellers would be indifferent. In contrast, all travellers would be indifferent between tolling and no policy as toll revenues are not returned to travellers. This property of fast lanes may explain why fast lanes have been introduced while there is generally a reluctance to introduce tolls.

A notable feature of the present paper is the formulation of scheduling utility which generalizes those employed by [Vickrey \(1969, 1973\)](#), [Arnott et al. \(1990, 1993\)](#), and many others. Here, scheduling utility is taken to be a strictly concave function of times at which the trip starts and ends. Travellers prefer to depart later and to arrive earlier. For any fixed travel time there is a unique preferred departure time. These assumptions are sufficient for the results of this paper. The paper establishes that the socially optimal fast lane scheme achieves more than half the welfare gain of the

²Using $\alpha - \beta - \gamma$ scheduling preferences, [Knockaert et al. \(2010\)](#) show that the coarse charge user equilibrium can be obtained by barring a certain group from travelling during the charging period and letting the remainder of drivers travel in that period without paying any charge. The present fast lane scheme does not have to designate specific time periods for specific groups of travellers. Moreover, the result is shown to hold for quite general scheduling preferences.

socially optimal continuously time varying toll. This generalizes the parallel result by [Arnott et al. \(1990\)](#) for the coarse toll under their first-in-first-out congestion technology to the present formulation of scheduling utility combined with parallel queueing.

Section 2 presents the model setup. Section 3 considers the Nash equilibrium under no policy. Section 4 analyses the coarse toll. Section 5 analyses the fast lane scheme. Section 6 extends to the case of price sensitive demand. Section 7 concludes.

2 The bottleneck model of congestion

Travellers depart from some origin to arrive at the back of the queue at a bottleneck after a travel time of d_0 minutes. The queue has no physical extension. Travellers have to pass through the bottleneck to reach their destination. Travel time after exiting the bottleneck is d_1 minutes. The bottleneck serves the queue at the capacity rate of ψ travellers per minute. Travellers and capacity may be split such that only a time-varying part of capacity $\psi_i(s)$ is available to travellers of class i . There is a separate queue for each class of travellers such that it is possible for some classes to queue while others pass the bottleneck. The class i queue at the bottleneck has length $Q_i(a) \geq 0$ at time a . Travellers arrive at the bottleneck at the rate $\rho_i(a)$ during some interval $[a_0, a_1]$ with the total number of class i travellers having arrived at time a being $R_i(a) = \int_{a_0}^a \rho_i(s) ds$. The queue evolves according to

$$Q_i(a) = R_i(a) - \int_{a_0}^a (\psi_i(s) 1_{\{Q_i(s) > 0\}} + \min(\psi_i(s), \rho_i(s)) 1_{\{Q_i(s) = 0\}}) ds.$$

A traveller who arrives at the bottleneck at time a , exits the bottleneck at time $t_i(a)$ given implicitly by $Q_i(a) = \int_a^{t_i(a)} \psi_i(s) ds$. When there is only one class such that capacity is fixed at ψ and if $Q(s) > 0$ for all $s \in \text{int}[a_0, a]$, then $t(a) = R(a)/\psi + a_0$.

The preferences of travellers are expressed by a utility function, which is a function of the departure time from the origin, the arrival time at the destination and a toll associated with the trip. This is written $u(a - d_0, t + d_1) - \tau$, where u is called the scheduling utility, a is the arrival time at the bottleneck, d_0 is the travel time from the origin to the back of the queue at the bottleneck, t is the time of exit from the bottleneck, d_1 is the travel time from the bottleneck to the destination and τ is a toll associated with the trip. Tolls are not redistributed to travellers. The scheduling utility is strictly concave, increasing in a and decreasing in t . Each traveller chooses arrival time at

the bottleneck to maximize utility. The function $v(a) = u(a, a)$ is assumed to attain maximum, which says there is an optimal time at which travellers would prefer to go if travel was instantaneous. The distances d_0 and d_1 are normalized to zero at no loss of generality. Travellers are identical except for an observable characteristic which does not affect utility and which cannot be influenced by travellers. The characteristic is observable by the road operator and can be used to form groups of any size.

The model is closed by assuming Nash equilibrium, which holds that no traveller is able to strictly increase utility by changing arrival time at the bottleneck, taking the arrival schedule of all other travellers as given. With identical travellers this implies that all travellers of class i achieve the same utility, denoted by Δ_i . The welfare function is the total consumer surplus plus total toll revenues plus a constant to be explained below. Denoting demand for class i by n_i and the total toll payment by T , the welfare function is

$$W = \sum_i \int_{-\Delta_i}^0 n_i(c) dc + T$$

The constant in the welfare function occurs because the upper integral limit is fixed at 0 in order to accommodate situations with a price elasticity of demand of zero.

3 No policy Nash equilibrium for one class

The analysis of no policy Nash equilibrium is carried out allowing for variable capacity $\psi(\cdot)$ as it will allow results to be reused below. In Nash equilibrium, travellers will arrive during an interval $[a_0, a_1]$ determined by the equations $N = \int_{a_0}^{a_1} \psi(s) ds$ and $v(a_0) = v(a_1)$. This means that there is no queue at these endpoints, $Q(a_0) = Q(a_1) = 0$. There is a strictly positive queue at any time in the interior of this interval. The arguments for this are easy and will not be repeated here (de Palma and Fosgerau, 2009). This implies that the expression for the queue length simplifies to

$$Q(a) = R(a) - \int_{a_0}^a \psi(s) ds$$

and so the implicit expression for the time of exit becomes

$$R(a) = \int_{a_0}^{t(a)} \psi(s) ds.$$

Therefore, by differentiation, $\rho(a) = \psi(t(a))t'(a)$. Differentiation of the identity $u(a, t(a)) = \Delta$ given by the equilibrium condition leads to

$$\begin{aligned} 0 &= u_1(a, t(a)) + u_2(a, t(a))t'(a) \\ &= u_1(a, t(a)) + u_2(a, t(a))\frac{\rho(a)}{\psi(t(a))}, \end{aligned}$$

such that the equilibrium arrival rate may be expressed as

$$\frac{\rho(a)}{\psi(t(a))} = -\frac{u_1(a, t(a))}{u_2(a, t(a))}.$$

The welfare function has the value $W = \Delta N = v(a_0)N$.

4 Coarse tolling

This section considers a coarse toll, defined to be a toll that is zero except in an interval during the peak where it takes a fixed value τ . Denote this interval by $[b_0, b_1]$. The toll applies at the front of the queue, when the traveller passes the bottleneck. The toll splits travellers into two classes according as they choose to travel at a time when the toll applies or not. There are separate queues for these two classes.³

The coarse toll should be small enough and the tolling interval be brief enough that it does not lead to unused capacity. Let therefore the toll be such that $\Delta \leq v(b_0) - \tau$. Otherwise, a traveller arriving at the bottleneck at or just after time b_0 will be worse off than in Nash equilibrium, even when there is no queue at time b_0 . Hence nobody will travel at time b_0 and there will be unused capacity. Similarly, the toll must have $\Delta \leq u(b_1, b_1) - \tau$. By concavity of scheduling utility, these conditions imply that $\Delta \leq v(a) - \tau$ for all $a \in [b_0, b_1]$. Together the conditions ensure that there is not unused capacity under the toll. Another consequence is that the arrival interval $[a_0, a_1]$ is unchanged relative to Nash equilibrium. This in turn implies that the equilibrium utility is still Δ .

The Nash equilibrium that results under a coarse toll that does not lead to unused capacity is illustrated in figure ???. Arrivals start at time a_0 and follow the no policy Nash equilibrium schedule ρ until just before the first arrival time where the traveller must pay the toll. Nash equilibrium requires that utility is Δ also for the first traveller to pay the toll. Therefore there

³The Stockholm congestion charge varies in steps over the peak. Anecdotes from Stockholm tell about cars waiting at the roadside for time to pass such that they will reach the toll gate after the toll has taken a step down.

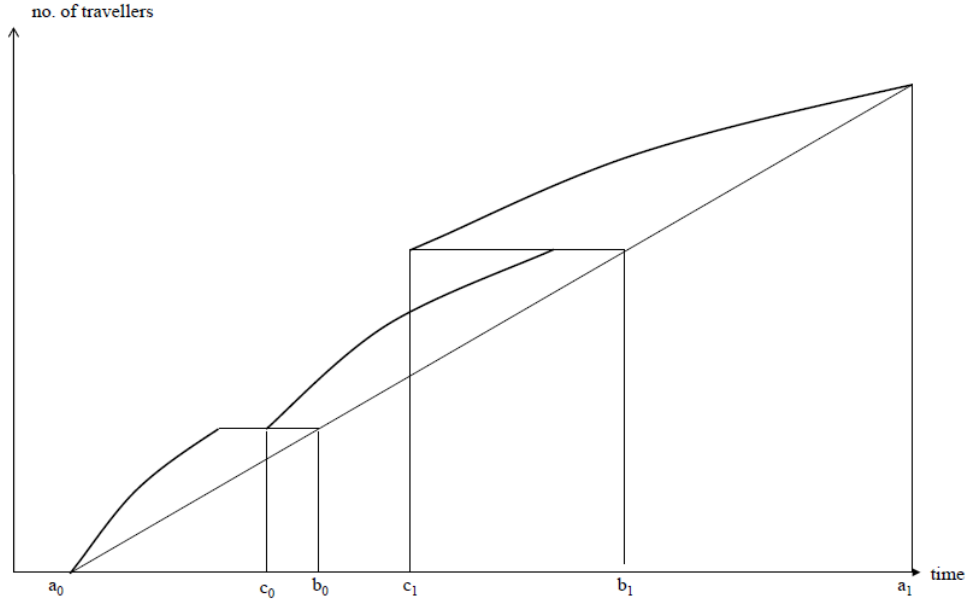


Figure 1: Arrival schedule under coarse toll

must be an interval with no new arrivals at the bottleneck. Arrivals resume at the time c_0 when $\Delta = u(c_0, b_0) - \tau$. Then arrivals continue until the last time when the traveller must pay the toll. Later travellers who will not pay the toll begin queueing in their own queue at time c_1 to follow the arrival schedule of the no policy equilibrium.

The optimal coarse tolling interval starts early enough and ends late enough that

$$\Delta = u(b_0, b_0) - \tau = u(b_1, b_1) - \tau. \quad (1)$$

This is illustrated in figure 2.

The welfare maximizing toll finds τ such that total scheduling utility is maximal. This is equivalent to maximizing toll revenue since the utility of every traveller is Δ in equilibrium. The toll revenue is $\tau\psi(b_1 - b_0)$. Maximizing toll revenue requires the first order condition

$$\psi(b_1 - b_0) + \tau\psi' \frac{\partial(b_1 - b_0)}{\partial\tau} = 0.$$

The endpoints of the tolling interval are such that

$$1 = \frac{\partial v(b_0)}{\partial b_0} \frac{\partial b_0}{\partial\tau} = \frac{\partial v(b_1)}{\partial b_1} \frac{\partial b_1}{\partial\tau}.$$

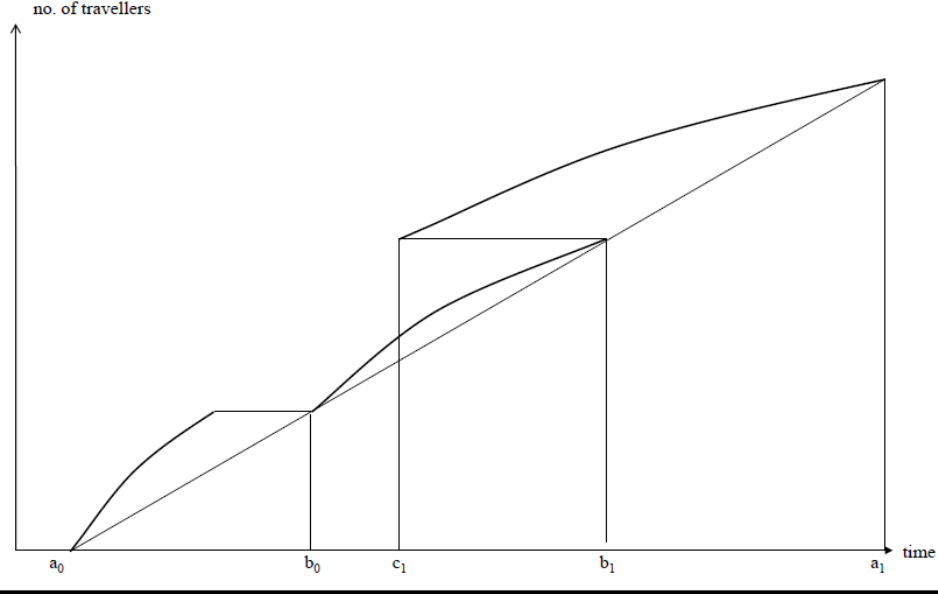


Figure 2: Arrival schedule under coarse toll with optimal tolling interval given size of toll

Hence the optimal toll satisfies

$$\tau = -\frac{(b_1 - b_0)}{\frac{1}{\frac{\partial v(b_1)}{\partial b_1}} - \frac{1}{\frac{\partial v(b_0)}{\partial b_0}}} > 0. \quad (2)$$

The following statement uses the theorem-proof format as the proof is fairly long. Let W_e be the welfare measure for the no policy Nash equilibrium. Index similarly the optimal coarse toll equilibrium with c and the social optimum with a time varying toll by so . The theorem states that the coarse toll captures at least half the welfare gain of the socially optimal time varying toll.

Theorem 1

$$\frac{W_{so} - W_e}{2} \leq W_c - W_e.$$

Proof. The function v is shown on figure 3. Time is on the horizontal axis and utility in monetary terms is on the vertical axis. All travellers receive utility $v(a_0) = v(a_1)$ in no policy Nash equilibrium. Under the socially optimal time varying toll travellers arrive at the bottleneck during $[a_0, a_1]$ at the rate ψ and there is no queueing. Hence $W_{so} = \int_{a_0}^{a_1} v(s) ds$ and so $W_{so} - W_e$

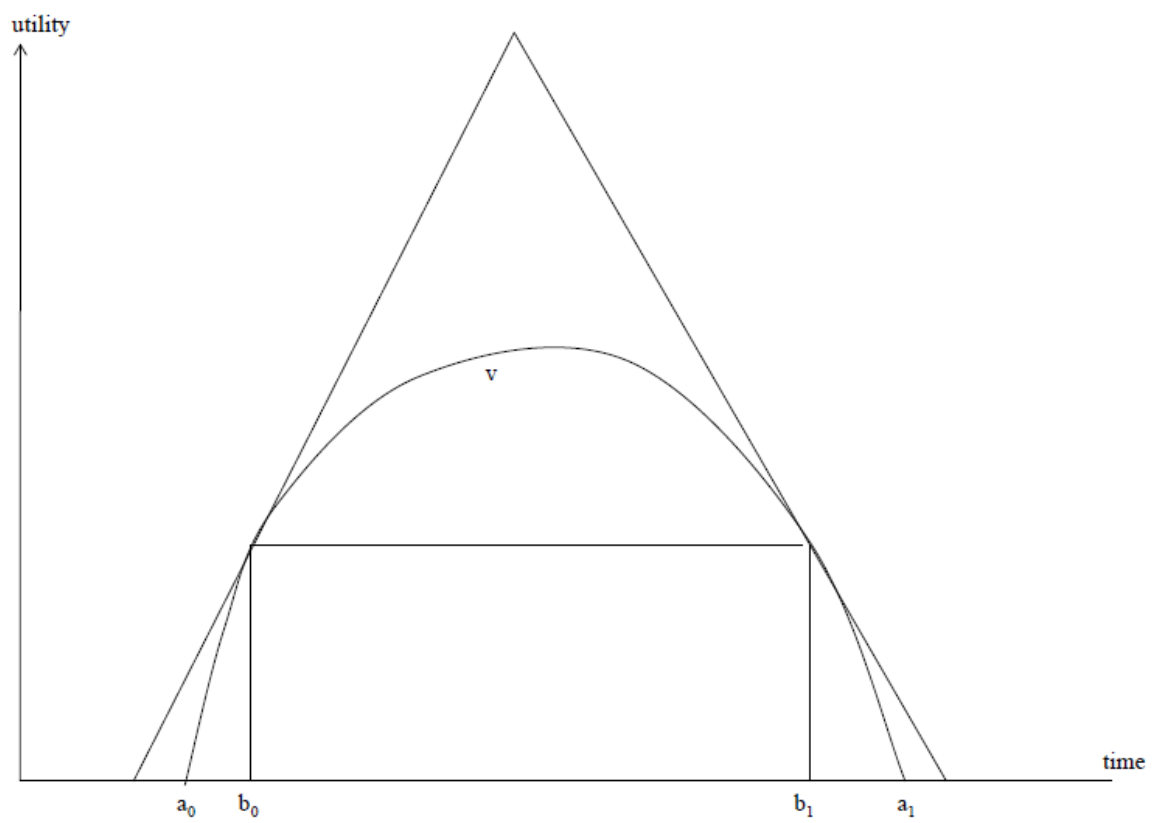


Figure 3: The welfare gain from the optimal coarse toll

corresponds to the area under v and over the horizontal line intersecting v at a_0 and a_1 .

Introducing a bit of shorthand notation, the optimal coarse toll is given in (2) by

$$\tau = -\frac{b_1 - b_0}{\frac{1}{v'_1} - \frac{1}{v'_0}}.$$

Consider the triangle, shown in figure 3, which has the horizontal line as baseline and edges that are tangent to v at $(b_0, v(b_0))$ and $(b_1, v(b_1))$. This triangle is strictly greater than $W_{so} - W_e$ since v is strictly concave. The baseline of the triangle is $\left[b_0 - \frac{\tau}{v'_0}, b_1 - \frac{\tau}{v'_1}\right]$. Parametrizing the edges of the triangle, the edges cross at the point x on the baseline given by the equation $v'_0 \left(x - b_0 + \frac{\tau}{v'_0}\right) = v'_1 \left(x - b_1 + \frac{\tau}{v'_1}\right)$ with solution $x = \frac{v'_1 b_1 - v'_0 b_0}{v'_1 - v'_0}$. The height of the triangle is then

$$\begin{aligned} v'_0 \left(x - b_0 + \frac{\tau}{v'_0}\right) &= v'_0 \left(\frac{v'_1 b_1 - v'_0 b_0}{v'_1 - v'_0} - b_0 + \frac{\tau}{v'_0}\right) \\ &= v'_0 \left(\frac{v'_1 b_1 - v'_0 b_0 - b_0 (v'_1 - v'_0)}{v'_1 - v'_0}\right) + \tau \\ &= \frac{v'_0 v'_1}{v'_1 - v'_0} (b_1 - b_0) + \tau = 2\tau. \end{aligned}$$

It follows that the area of the triangle is $\tau \left(b_1 - b_0 - \frac{\tau}{v'_1} + \frac{\tau}{v'_0}\right) = 2\tau (b_1 - b_0)$, such that $W_{so} - W_e \leq 2\tau (b_1 - b_0)$. The welfare gain from the coarse toll is equal to the toll revenue $W_c - W_e = \tau (b_1 - b_0)$. ■

The statement of the Theorem holds with equality, i.e. $W_{so} - W_e = 2(W_c - W_e)$, when scheduling preferences are piecewise linear $u(a, t) = -\alpha \cdot (t - a) - \beta \cdot \min(t, 0) + \gamma \cdot \max(t, 0)$. Hence the bound in the Theorem cannot be improved in general.

5 Fast laning

This section considers a fast lane scheme. The suggested scheme is the following. Based on the observable characteristic of travellers, an operator assigns travellers to two classes with N_1 and N_2 travellers. The preferences of travellers in the two groups are the same since the observable characteristic is unrelated to preferences. A more than proportional share of capacity ψ_1 is

reserved for class 1 travellers where

$$0 < \frac{N_1}{N} < \frac{\psi_1}{\psi} \leq 1. \quad (3)$$

Class 2 travellers are allowed to use all of capacity ψ when there are no class 1 travellers queueing. Otherwise they can only use the remaining capacity $\psi - \psi_1$.

This section establishes the two main results of this paper. The first result is that such a scheme is always Pareto improving. This is in contrast to a coarse toll, which may reduce welfare by causing idle capacity during $[a_0, a_1]$ if the toll is too high or lasts for too long. The second result is that the scheme can achieve exactly the same travel time and queueing outcomes as a coarse toll satisfying the partial optimality conditions that the tolling interval equalizes the utility of the travellers at the endpoints (1).

To establish that a fast lane scheme is always Pareto improving, consider first the behavior of class 1. Since $N_1/\psi_1 < N/\psi$, they are able to pass the bottleneck in a shorter interval than $[a_0, a_1]$. They form a Nash equilibrium during an interval $[b_0, b_1]$ with $a_0 < b_0, b_1 - b_0 = N_1/\psi_1, b_1 < a_1$ and $v(b_0) = v(b_1)$. These travellers therefore achieve a strict utility gain relative to the basic Nash equilibrium, $v(v_1) > \Delta$.

The remaining class 2 travellers, form a Nash equilibrium using the remaining capacity. The remaining capacity is $\psi - \psi_1$ during $[b_0, b_1]$ and ψ otherwise. All class 2 travellers are able to pass the bottleneck during $[a_0, a_1]$. They therefore form Nash equilibrium with variable capacity as described in section 3 and the resulting equilibrium utility is Δ . This establishes that any fast lane scheme satisfying (3) will be Pareto improving.

The next point is to establish that the fast lane scheme can achieve exactly the same travel time and queueing outcomes as a coarse toll satisfying the partial optimality conditions. So consider a toll τ during $[b_0, b_1]$ satisfying (1). This is the kind of situation depicted in figure 2. Note that toll paying travellers occupy all the bottleneck capacity during $[b_0, b_1]$ and that these travellers do not queue at times b_0 and b_1 . Hence their behavior is reproduced under a scheme that assigns $\psi(b_1 - b_0)$ travellers into class 1 and gives them all the capacity $\psi_1 = \psi$, while the remaining travellers are assigned to class 2. The behavior of the remaining travellers is similarly reproduced by this scheme, since their arrival schedule is Nash equilibrium for them when capacity is ψ except during $[b_0, b_1]$ where it is zero.

6 Fast lane with price sensitive demand

The no policy equilibrium scheduling utility depends only on the time it takes all travellers to pass the bottleneck $\frac{N}{\psi}$. Introduce a cost function $c\left(\frac{N}{\psi}\right) = -\Delta$ equal to the negative of the equilibrium scheduling utility. By the assumptions made, c is strictly increasing and strictly convex. Assume that demand is price sensitive with demand given as a decreasing function $n(c)$ of scheduling cost. Combine the functions n and c into a new function as $f\left(\frac{N}{\psi}\right) = n\left(c\left(\frac{N}{\psi}\right)\right)$. Then $N = f\left(\frac{N}{\psi}\right)$ in no policy Nash equilibrium.

Consider now a fast lane scheme that divides the population of potential travellers into two classes with proportions α and $1 - \alpha$. The allocation of travellers to classes is based on the observable characteristic which cannot be influenced by travellers. The division of the population of potential travellers into classes gives rise to two demand curves $\alpha n(\cdot)$ and $(1 - \alpha)n(\cdot)$.

Let N_1, N_2 be the equilibrium number of travellers in the two classes, where group 1 is the priority class and capacity is assigned such that $0 < \frac{N_1}{N_1 + N_2} < \frac{\psi_1}{\psi} \leq 1$. As shown, the equilibrium cost for class 1 travellers is $c(N_1/\psi_1)$, such that the equilibrium number of class 1 travellers is $N_1 = \alpha f\left(\frac{N_1}{\psi_1}\right)$. Similarly, the equilibrium cost for class 2 travellers is $c(N/\psi)$ such that equilibrium requires that $N_2 = (1 - \alpha) f\left(\frac{N_1 + N_2}{\psi}\right)$.

It has already been shown that such a fast lane scheme is welfare improving when demand is not price sensitive. It follows by continuity that the scheme is also welfare improving when the price elasticity of demand is sufficiently small. The remainder of this section shows that the scheme is welfare improving when the share α is sufficiently small.

The no policy Nash equilibrium results when $\psi_1 = \alpha\psi$. The following derivatives will be useful. They show that increasing α while holding ψ_1 constant increases the number of class 1 travellers and decreases the number of class 2 travellers.

$$\begin{aligned} \frac{\partial N_1}{\partial \alpha} &= f\left(\frac{N_1}{\psi_1}\right) + \frac{\alpha}{\psi_1} f'\left(\frac{N_1}{\psi_1}\right) \frac{\partial N_1}{\partial \alpha} = \frac{f\left(\frac{N_1}{\psi_1}\right)}{1 - \frac{\alpha}{\psi_1} f'\left(\frac{N_1}{\psi_1}\right)} > 0 \\ \frac{\partial N_2}{\partial \alpha} &= -f\left(\frac{N_1 + N_2}{\psi}\right) + \frac{1 - \alpha}{\psi} f'\left(\frac{N_1 + N_2}{\psi}\right) \left(\frac{\partial N_1}{\partial \alpha} + \frac{\partial N_2}{\partial \alpha}\right) \\ &= -\frac{f\left(\frac{N_1 + N_2}{\psi}\right) - \frac{1 - \alpha}{\psi} f'\left(\frac{N_1 + N_2}{\psi}\right) \frac{\partial N_1}{\partial \alpha}}{1 - \frac{1 - \alpha}{\psi} f'\left(\frac{N_1 + N_2}{\psi}\right)} < 0 \end{aligned}$$

The expression for $\frac{\partial N_2}{\partial \alpha}$ is composed of two negative terms. The first is the direct effect on N_2 of changing α , the second term is the effect through the cost. The second term is negative indicating that the cost is increasing. This indicates that unprioritized travellers lose as α increases such that larger α does not lead to a Pareto improvement. The following derivation shows that the total number of travellers strictly decreases as α increases.

$$\begin{aligned} \frac{\partial N_1}{\partial \alpha} + \frac{\partial N_2}{\partial \alpha} &= \frac{\frac{\partial N_2}{\partial \alpha} + f\left(\frac{N_1+N_2}{\psi}\right)}{\frac{1-\alpha}{\psi} f'\left(\frac{N_1+N_2}{\psi}\right)} \\ &= -\frac{\frac{\partial N_1}{\partial \alpha} + f\left(\frac{N_1+N_2}{\psi}\right)}{1 - \frac{1-\alpha}{\psi} f'\left(\frac{N_1+N_2}{\psi}\right)} < 0 \end{aligned} \quad (4)$$

The welfare function with variable demand is defined as the consumer surplus

$$W = \alpha \int_{c\left(\frac{N_1}{\psi_1}\right)}^{\infty} n(c) dc + (1 - \alpha) \int_{c\left(\frac{N_1+N_2}{\psi}\right)}^{\infty} n(c) dc.$$

Differentiation with respect to α yields

$$\frac{\partial W}{\partial \alpha} = \int_{c\left(\frac{N_1}{\psi_1}\right)}^{c\left(\frac{N_1+N_2}{\psi}\right)} n(c) dc - \frac{N_1}{\psi_1} c' \left(\frac{N_1}{\psi_1} \right) \frac{\partial N_1}{\partial \alpha} - \frac{N_2}{\psi} c' \left(\frac{N_1+N_2}{\psi} \right) \left(\frac{\partial N_1}{\partial \alpha} + \frac{\partial N_2}{\partial \alpha} \right).$$

Evaluating at $\alpha = 0$ yields

$$\frac{\partial W}{\partial \alpha} = -\frac{N_2}{\psi} c' \left(\frac{N_1+N_2}{\psi} \right) \left(\frac{\partial N_1}{\partial \alpha} + \frac{\partial N_2}{\partial \alpha} \right).$$

This is strictly positive by (4). Continuity implies that there is always some $\alpha > 0$ such that the fast lane scheme is welfare improving. Further simplifications do not seem to be available.

7 Conclusion

It is straightforward but tedious to generalize the results of this paper to tolls with more steps and fast laning schemes with more user classes.⁴ The

⁴Laih (1994) showed that it is straightforward to extend the coarse toll to a multistep toll. It is similarly straightforward to extend a fast lane scheme in this way. Laih (1994) did not recognize that it was necessary to reformulate the queueing technology in order to

general conclusion remains that fast laning can achieve the same benefits as step tolls when demand is not price sensitive. It is also straightforward to see that a sequence of step tolls, and hence a sequence of fast laning schemes, can be constructed that approach the optimal time varying toll. In the limit, the step toll would become the optimal continuously varying toll while the fast lane scheme would become equivalent to allocating a specific time slot to every traveler.

A potentially useful feature of the fast lane scheme is its robustness. As long as demand is not too elastic, or as long as the share of prioritized travellers is not too large, then any fast lane scheme satisfying (3) is welfare improving. If demand is not price sensitive, then any such fast lane scheme is Pareto improving. An interesting direction for further inquiry is how this robustness can be utilized. Is it the case that the fast lane scheme retains its favorable properties when some element of stochasticity is introduced into the model?

References

- Arnott, R. A., de Palma, A. and Lindsey, R. (1990) Economics of a bottleneck *Journal of Urban Economics* **27**(1), 111–130.
- Arnott, R. A., de Palma, A. and Lindsey, R. (1993) A structural model of peak-period congestion: A traffic bottleneck with elastic demand *American Economic Review* **83**(1), 161–179.
- de Palma, A. and Fosgerau, M. (2009) Random queues and risk averse users *Working Paper* .
- Knockaert, J., Verhoef, E. T. and Rouwendal, J. (2010) Bottleneck Congestion: Differentiating the Coarse Charge *Working Paper* .
- Laih, C.-H. (1994) Queueing at a bottleneck with single- and multi-step tolls *Transportation Research Part A* **28**(3), 197–208.
- Laih, C. H. (2004) Effects of the optimal step toll scheme on equilibrium commuter behaviour *Applied Economics* **36**(1), 59–81.
- Shen, W. and Zhang, H. M. (2010) Pareto-improving ramp metering strategies for reducing congestion in the morning commute *Transportation Research Part A* **44**(9), 676–696.

obtain his results. This was rectified in [Laih \(2004\)](#).

Vickrey, W. S. (1969) Congestion theory and transport investment *American Economic Review* **59**(2), 251–261.

Vickrey, W. S. (1973) Pricing, metering, and efficiently using urban transportation facilities *Highway Research Record* **476**, 36–48.