# Congestion in a city with a central bottleneck

Fosgerau, Mogens and de Palma, André

Technical University of Denmark

2012

# Congestion in a city with a central bottleneck*

Mogens Fosgerau
Technical University of Denmark
Centre for Transport Studies, Sweden
Ecole Nationale Supérieure de Cachan, CES, France
mf@transport.dtu.dk

André de Palma
Ecole Nationale Supérieure de Cachan, CES, France
andre.depalma@ens-cachan.fr

November 19, 2011

**Abstract**

We consider dynamic congestion in an urban setting where trip origins are spatially distributed. All travelers must pass through a downtown bottleneck in order to reach their destination in the CBD. Each traveler chooses departure time to maximize general concave scheduling utility. We find that, at equilibrium, travelers sort according to their distance to the destination; the queue is always unimodal regardless of the spatial distribution of trip origins. We construct a welfare maximizing tolling regime, which eliminates congestion. All travelers located beyond a critical distance from the CBD gain from tolling, even when toll revenues are not redistributed, while nearby travelers lose. We discuss our results in the context of acceptability of tolling policies.

Key words: dynamic model; toll policy; spatial differentiation; acceptability

*We thank Richard Arnott as well as the editor Robert Helsey and two referees for very insightful comments. Mogens Fosgerau is funded by the Danish Council for Strategic Research. 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'Ecologie, de l'Energie, 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.

# 1 Introduction

This paper presents a model that integrates two prominent features of urban congestion, focusing on the exemplary case of the morning commute. The first feature is that congestion is a dynamic phenomenon in the sense that congestion at one time of day affects conditions later in the day through the persistence of queues. The second feature is that trip origins are spatially distributed. We analyze how these features interact in a city with a central bottleneck and provide results concerning optimal pricing.

The dynamics of congestion were analyzed in the seminal Vickrey (1969) bottleneck model (see also Arnott et al., 1993), which captures the essence of congestion dynamics in a simple and tractable way. Travellers are viewed as having scheduling preferences concerning the timing of trips that have to pass the bottleneck. The analysis concerns equilibrium in the traveller choice of departure time.

The Vickrey (1969) analysis of congestion, however, essentially ignores space. Using the notation of the current paper, travellers are depicted as travelling some distance $c$ (measured in time units) until they reach a bottleneck at time $a$. They exit the bottleneck to arrive at the destination at time $t$. They have scheduling preferences, always preferring to depart later and always preferring to arrive earlier. The Vickrey (1969) scheduling preferences can be expressed by the scheduling cost $\alpha \cdot c + \alpha \cdot (t - a) + D(t)$ , where $\alpha$ is the value of travel time, $t - a$ is the time spent in the bottleneck and $D(t) = \beta \cdot \max(0, t^* - t) + \gamma \cdot \max(0, t - t^*)$ is a convex function capturing the cost of being early or late relative to some preferred arrival time $t^*$. The Vickrey formulation of scheduling preferences is additively separable in trip duration and arrival time and it is linear in trip duration. So it is clear that the distance $c$ to the bottleneck does not matter for the Vickrey analysis of how travellers time their arrival at the bottleneck and the ensuing congestion.[1]

It is not generally true that the distance from trip origins to the destination is irrelevant for the timing of trips. Consider a traveller who always prefers to depart later and always prefers to arrive earlier. Faced by a fixed trip duration that is independent of the departure time, such a traveller will optimally time his trip such that his marginal utility of being at the origin at the departure time equals his marginal utility of being at the destination at the arrival time. If the marginal utility at the origin is decreasing and his marginal utility of being at the destination is increasing, then an increase in trip duration will cause him to depart earlier and

[1] The analysis of the bottleneck model has been developed and extended in various directions by Arnott, de Palma and Lindsey in a series of papers; notably Arnott et al. (1993). These authors use the above $\alpha - \beta - \gamma$ preferences or a version where the function $D(t)$ has a more general form. They always maintain linearity and additive separability of travel time and are hence unable to analyse the consequences of distance for congestion.

to arrive later. In this way the distance can matter for the timing of trips. This paper concerns travellers with such scheduling preferences.

Congestion can arise when there is a bottleneck and many individuals who want to pass the bottleneck at the same time. It is not a sufficient condition for congestion to arise that travellers have similar scheduling preferences. Trip origins must also be located with similar distances to the bottleneck. If trip origins are sufficiently dispersed, then congestion does not arise as there is no overlap in the times when travellers want to pass the bottleneck. Hence it is clear that the spatial distribution of travel demand is a fundamental determinant of urban congestion. This observation stands in contrast to the standard urban model, where congestion increases with population dispersion.

This paper is the first to allow for spatial heterogeneity in the bottleneck model in a meaningful way. In our model, heterogeneity is induced by the structure of the city. A number of earlier contributions have considered preference heterogeneity in the context of the bottleneck model (e.g., Vickrey, 1973; Arnott et al., 1994; van den Berg and Verhoef, 2011). These papers work in the context of linear separable Vickrey (1969) scheduling preferences and heterogeneity is introduced by varying $\alpha-\beta-\gamma$, while maintaining the ratio $\beta/\gamma$ fixed for reasons of analytical convenience. Generally speaking, this sort of heterogeneity can induce travellers to sort according to the degree of closeness to the center of the congestion peak; in a two group case, sorting has the form that one group occupies a central time interval while the other group occupies the early and late shoulders. In contrast, this paper finds that travelers sort according to their distance to the bottleneck; this occurs both under no tolling and under optimal tolling, and the result is derived under quite general assumptions concerning scheduling preferences.[2] Hendrickson and Kocur (1981), Smith (1984), Newell (1987), and Arnott et al. (1994) consider the case of travellers with scheduling preferences, such as $\alpha-\beta-\gamma$, that are additively separable in trip duration and arrival time and who differ in their preferred arrival time. In that case, travellers sort according to their preferred arrival time, which is similar to what we obtain here. Kuwahara (1990) extends this to a geometry consisting of two residential areas and a CBD with bottlenecks in between. Travellers within each group then still sort according to their preferred arrival time, but a strict sequence does not hold for the two groups together. The present case is more involved, as travellers have different distances to the CBD as well as strictly concave and non-separable scheduling preferences. We show that the optimal arrival time $a_*$, in the absence of congestion, is increasing in distance $c$, such that also here travellers sort according to their preferred arrival time.

Daganzo (2007) and Geroliminis and Daganzo (2008) show that several aspects of congestion in an urban area can be described as a form bottleneck con-

[2] Lindsey (2004) considers more general heterogeneity with a finite number of user classes.

gestion. A space average of traffic measurements show that the trip completion rate is a stable inverse u-shaped function of the number of vehicles present in the system. Cars that have not yet completed their trips remain in the urban area, such that it is possible to think of the system as a generalized sort of queue. The bottleneck model supposes a constant trip completion rate and a queueing system that maintains a first-in-first-out queue. See also Geroliminis and Levinson (2009).

Section 2 presents the model setup. The analysis of equilibrium in section 3 shows that equilibrium exists uniquely in the laissez-faire situation without tolling. Under laizzes-faire, travellers sort according to their distance to the bottleneck such that those who are closest to the bottleneck reach the destination first. However, in general there is not a monotonous relationship between distance and departure time; it is not necessarily the case that those who are located further away will depart earlier. Section 4 then concerns equilibrium under socially optimal tolling at the bottleneck. The toll can be taken to be zero for the first and last travellers and strictly positive for everybody else. The optimal toll exactly removes queueing. The sequence of arrivals at the destination is preserved from the laissez-faire equilibrium.[3] However, in contrast to the Vickrey analysis with homogenous travellers, arrivals at the destination occur earlier in social optimum than under laissez-faire. When the use of toll revenues does not affect the utility of travellers, then the toll just represents a loss for them. This is compensated to some extent by a gain in utility. Comparing social optimum to laissez-faire reveals that those who are located furthest away from the bottleneck will experience a net gain, while those who are located near the bottleneck will experience a net loss. Section 5 illustrates the model numerically. Section 6 concludes. All proofs are deferred to the appendix.

# 2 A spatial model

Consider a city in which a continuum of $N$ identical individuals are spatially distributed. They make one trip each and must all pass through a downtown bottleneck. The trip duration of an individual is the sum of the distance, measured in time units, to the bottleneck and the time spent in the bottleneck.

The bottleneck has a capacity of $\psi$ persons per time unit. The distribution of travel distances $c$ has cumulative distribution $F$ with density $f$ and support $C = [c_0, c_1]$. The situation is illustrated in Figure 1.

[3] Arnott et al. (1991) consider a variant of the standard bottleneck model in which drivers have to park, and parking spots are located at varying distances from the CBD. In the laissez-faire user equilibrium, drivers park in order of increasing distance from CBD. The optimal location-dependent parking fee reverses this pattern. This contrasts with the present setting where optimal tolling does not change the order of arrivals.

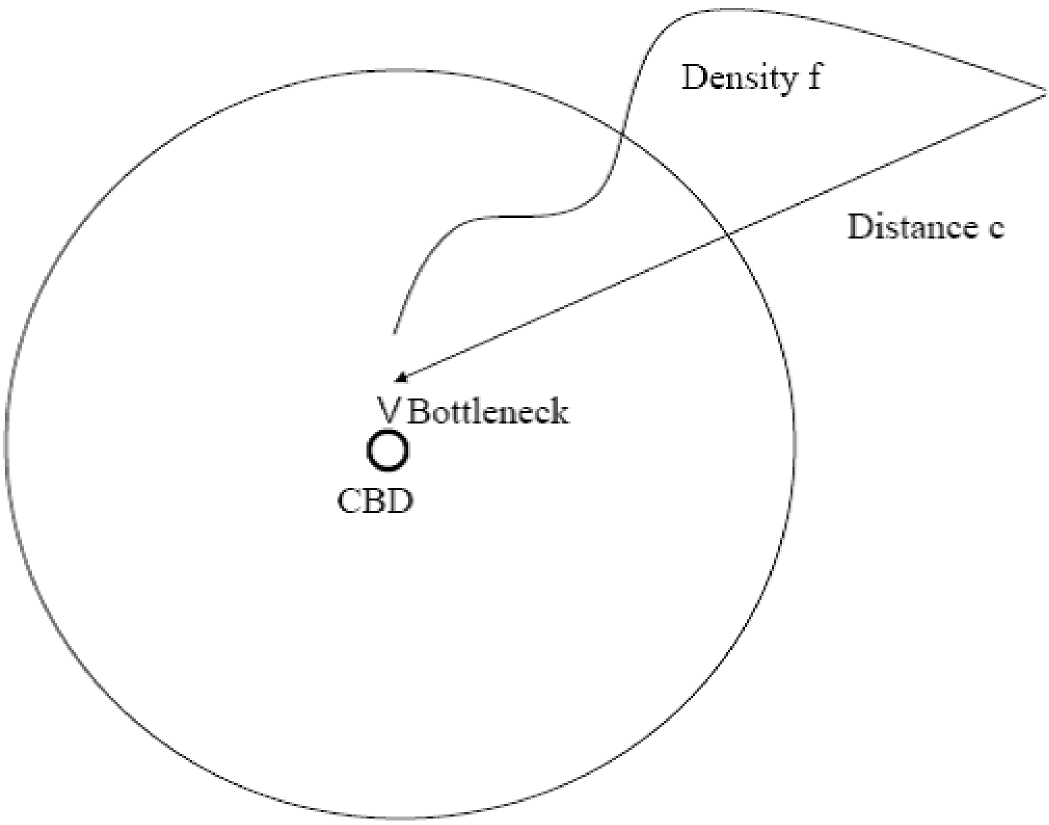


Figure 1: The representation of space

Let $a$ be the arrival time at the bottleneck and $c$ be the travel distance to the bottleneck. Then $d = a - c$ is the departure time from home. Individuals arrive at the bottleneck during some interval $[a_0, a_1]$ at the time varying rate $\rho(a)$. Cumulative arrivals are $R(a) = \int_{a_0}^{a} \rho(s)\, ds$. In the case when there is queue from time $a_0$ to time $a$, then an individual who arrives at entrance of the bottleneck at time $a$ will arrive at the destination at time $R(a)/\psi + a_0$, since it takes $R(a)/\psi$ time units for the first $R(a)$ travelers to pass the bottleneck (Arnott et al., 1993). The queue is vertical, meaning that its physical extension has no consequences.[4]

Individuals are identical with preferences concerning the timing and cost of the trip expressed by the utility $u(d,t) - \tau$, where $d$ is the departure time, $t > d$ is the arrival time and $\tau$ is the toll payment associated with the trip.[5] Toll revenues are not returned to travellers and their utility is not affected by the use of revenues. This assumption puts a focus on the direct impact of tolling on travellers. All other monetary trip costs are held constant and hence ignored. In brief we shall refer to $u(d,t)$ as the scheduling utility. We shall define a social welfare function to be sum of the individual utilities plus the toll revenue. Hence the welfare function reduces to the sum of the individual scheduling utilities.

Partial derivatives of $u(\cdot,\cdot)$ are assumed to exist up to second order and are denoted as $u_1, u_{12}$ etc. We assume that $u(\cdot,\cdot)$ is strictly concave and that $u_1 > 0 > u_2$. In addition, we shall refer to the following conditions regarding the derivatives of $u$.

**Condition 1** $\forall d \leq t : u_{11}(d,t) - \frac{u_1(d,t)}{u_2(d,t)} u_{12}(d,t) < 0.$

**Condition 2** $\forall d \leq t : u_{12}(d,t) - \frac{u_1(d,t)}{u_2(d,t)} u_{22}(d,t) < 0.$

**Condition 3** $\forall d \leq t : u_{11}(d,t) + u_{12}(d,t) < 0.$

**Condition 4** $\forall d \leq t : u_{12}(d,t) + u_{22}(d,t) < 0.$

Conditions 1 and 2 are used to prove that equilibrium is unique. Together they ensure that $u_1(a-c,a)/u_2(a-c,a)$ is decreasing as a function of $a$. Conditions 3 and 4 ensure that $u_1(a-c,a)$ and $u_2(a-c,a)$ are decreasing as functions of $a$. We assume that $u(a-c,a)$ attains a maximum as a function of $a$ for any $c \geq 0$. This simply assures that travellers have an optimal time of departure when trip

[4]See il Mun (1999) and Arnott and DePalma (2011) for models with horizontal queues.

[5]A reviewer noted that scheduling preferences may be partly motivated by referring to consumers who derive utility from consumption and leisure, where, e.g., leisure is produced before departure and income for consumption is earned at work after arrival. Scheduling preferences allow for the opportunity cost of time lost in transportation. They may also account for a benefit from coordination of activities (Fosgerau and Small, 2011).

duration is constant. We also assume that $u_1(a-c,a)+u_2(a-c,a)$ ranges over all of $\mathbb{R}$ as $a$ varies.

As noted above, the key to the analysis is the formulation of scheduling utility that is not additively separable in trip duration. Vickrey (1973) described a scheduling utility that is instead additively separable in departure time and arrival time and not linear in trip duration. Such a separable specification would have $u(d,t)=H(d)+W(t)$, and the above conditions on utility would be satisfied if $H'>0>W'$, $H''<0$ and $W''<0$; cross-derivatives are zero due to the additive separability. A possible interpretation of the Vickrey (1973) specification is that $H(d)$ is the utility that the traveller obtains from being at home until departure (at time $d$) and $W(t)$ is the utility the traveller obtains from being at work after arrival (at time $t$). Tseng and Verhoef (2008) provide empirical evidence that supports the Vickrey (1973) specification.

The customary $\alpha-\beta-\gamma$ specification is a limiting case of the Vickrey (1973) specification in which $H'=\alpha$, and $W'=\alpha-\beta$ for $t<t_0$ and $W'=\alpha+\gamma$ otherwise; the $\alpha-\beta-\gamma$ specification is piecewise linear and hence does not satisfy the present conditions.

de Palma and Fosgerau (2011) use a formulation, equivalent to the current, in which scheduling utility is taken to be a general concave and potentially nonseparable function of trip duration and arrival time.

Consider a single individual facing a fixed travel time $c$, such as in the case when there is no congestion. His scheduling utility as a function of the arrival time at the bottleneck $a$ is then $u(a-c,a)$. Denote the optimal arrival time at the destination by $a_*(c)=\arg\max_a u(a-c,a)$. Then we obtain the following result.

**Theorem 1** *The optimal arrival time $a_*(c)$ exists and is unique. Condition 3 implies that $a_*'(c)>0$. If also condition 4 holds, then $a_*'(c)<1$.*

This theorem has significant implications. Heterogeneity in the distance to the destination translates into heterogeneity in the optimal arrival time at the destination. That $a_*'(c)>0$ means that more distantly located individuals will prefer to arrive later in the absence of queue. In particular it means that individuals sort according to their distance from the destination. If $a_*'(c)<1$, then $d'(c)=a_*'(c)-1<0$, which means that more distantly located individuals will depart earlier from home. Fosgerau and Engelson (2011) presents empirical evidence indicating that commuters located further from the CBD tend both to depart earlier and to arrive later.

It is possible that the density of travellers at different distances is so low that there will not be queueing. Therefore a condition is introduced to guarantee that all travellers will be queueing.[6] The condition comes in three versions of different strengths.

**Condition 5** $\forall c \in C, c > c_0 : a_* (c) < \frac{F(c)}{\psi} + a_* (c_0)$.

**Condition 6** $\forall c \in C : a'_* (c) < \frac{f(c)}{\psi}$.

**Condition 7** $\forall c \in C : 1 < \frac{f(c)}{\psi}$.

Each result below refers to one of these conditions. Condition 5 ensures that if the first traveller arrives at his optimal arrival time and if travellers sort according to their distance to the bottleneck, then all other travellers will arrive later than their optimal arrival time. Given sorting, the next condition 6 implies that if one traveller hits his optimal arrival time, then all travellers located further away will arrive later than their optimal time; thus condition 6 implies condition 5 if travellers sort. Condition 7 is the strongest version, since together with $a'_* (c) < 1$ it implies condition 6 ; condition 7 states that the density of travellers in $C$ is everywhere greater than the bottleneck capacity. Condition 5 is weaker than condition 7 since it can be satisfied if $f$ places more than average mass at small values of $c$ and less than average mass at large values of $c$.

# 3 Laissez-faire

Consider equilibrium in pure strategies, where each individual takes the behaviour of all other individuals as given, and no individual has incentive to change his time of arrival at the bottleneck.[7] The following theorem provides some basic characteristics of equilibrium under the laissez-faire policy of no toll.

**Theorem 2** *Assume that condition 5 holds. Under laissez-faire, arrivals at the bottleneck take place during an interval* $[a_0, a_1]$, *where* $a_1 = a_0 + N/\psi$. *There is always queue during this interval. Furthermore,* $a_0 \leq a_* (c_0)$ *and* $a_1 \geq a_* (c_1)$.

So the first traveller arrives at the bottleneck earlier than he would prefer in the absence of congestion and the last traveller arrives later. The no residual queue property holds, i.e., the queue is exactly gone at the time the last traveller arrives at the bottleneck (see de Palma and Fosgerau, 2011). Otherwise the last traveller

[6] This condition serves to make analysis simpler, but it is not essential: analysis could be carried out just for those travelers who do queue, the complication is that the boundary between queueing and not queueing is affected, e.g., by tolling.

[7] We accept without proof that any Nash equilibrium is characterised by a differentiable function $a (c)$ describing the time of arrival at the bottleneck for a traveller located at distance $c$.

could simply postpone departure to reduce queueing while not affecting the arrival time. The next theorem states some properties of laissez-faire equilibrium.

**Theorem 3** *Assume that conditions 1, 2, 3 and 5 hold. Under laissez-fare, equilibrium exists and is unique. Individuals located at distance $c$ from the bottleneck arrive at the bottleneck at time $a(c)$, where $a(c)$ satisfies*

$$a'(c) = -\frac{u_2\left(a(c) - c, \frac{F(c)}{\psi} + a_0\right)}{u_1\left(a(c) - c, \frac{F(c)}{\psi} + a_0\right)} \frac{f(c)}{\psi} > 0. \tag{1}$$

*The arrival schedule at the destination is*

$$\frac{F(c)}{\psi} + a_0$$

*and the equilibrium scheduling utility is*

$$u\left(a(c) - c, \frac{F(c)}{\psi} + a_0\right).$$

The theorem first gives a differential equation for the arrival time $a(c)$ at the bottleneck as a function of distance. The derivative $a'$ is strictly positive which means that the travellers located at greater distances arrive at the bottleneck later. So under laissez-faire, travellers sort according to their distance to the bottleneck.

However, it is not the case that travellers located at greater distances also depart earlier. In general, it is not possible to sign the derivative of the departure time $d'(c) = a'(c) - 1$, since $a'(c)$ may be greater or smaller than $1$. In particular, if the density of distances is low in some region, then derivative of the departure time may change sign. The numerical exercise in section 5 below exhibits such a case where $d'(c)$ is positive at some distances and negative at other.

The sorting property lies behind the expression for the arrival schedule at the destination. Travellers arrive at the bottleneck in sequence sorted according to their distance to the bottleneck and the sequence is preserved by the bottleneck. The first traveller, located at $c_0$, arrives at the destination at time $a_0$. The traveller at $c$ arrives at the destination when the $F(c)$ travellers who are located closer have arrived. They take $F(c)/\psi$ time units to pass the bottleneck and so the traveller at $c$ arrives at time $F(c)/\psi + a_0$.

Given the initial condition $a(c_0) = a_0$, (1) describes $a(\cdot)$ uniquely. The proof of the theorem shows that the equilibrium condition $a(c_1) = a(c_0) + N/\psi$ has a unique solution.

The next theorem describes the evolution of the queue. We use the terminology that a unimodal function has one local maximum whereas a multimodal function has at least two.

**Theorem 4** *Assume that conditions 3, 4 and 7 hold and also the conclusions of Theorem 3. The equilibrium queue length $q(c) = a_0 + F(c)/\psi - a(c)$ is unimodal as a function of distance $c$.*

This theorem is interesting since it shows that the evolution of the queue is largely dissociated from the urban form; thus the distribution of distances from residences to the CBD may be multimodal, but the queue is always unimodal. The fundamental reason behind is that this aspect of the shape of the queue is determined by the equilibrium condition in conjunction with the properties of scheduling preferences. In particular, the proof of the theorem shows first that the derivative of the queue length has the same sign as $u_1 + u_2$, and then that $u_1 + u_2 < 0$ implies that $u_1 + u_2$ is decreasing. This guarantees that $q$ has only one local maximum, which hence must be global.

# 4 Socially optimal tolling

This section concerns the socially optimal toll, retaining the definition of equilibrium from the analysis of the laissez-faire equilibrium. Recall that the toll at time $a$ at the bottleneck is $\tau(a)$.

**Theorem 5** *A socially optimal toll exists when condition 6 holds. Arrivals at the bottleneck take place during an interval $[a_{\tau 0}, a_{\tau 1}]$ according to the schedule $a_\tau(c) = \frac{F(c)}{\psi} + a_{\tau 0}$ where $a_{\tau 0}$ is the unique solution to*

$$0 = \int_{a_{\tau 0}}^{a_{\tau 0}+N/\psi} \left(u_1(a-c,a) + u_2(a-c,a)\right) da.$$

*An optimal toll satisfies*

$$\tau'(a_\tau(c)) = u_1(a_\tau(c) - c, a_\tau(c)) + u_2(a_\tau(c) - c, a_\tau(c)),$$

*where $a_\tau$ is the arrival schedule. The optimal toll satisfies $\tau(a_{\tau 0}) = \tau(a_{\tau 1})$ and may be chosen such that $\tau(a_{\tau 0}) = 0$. A socially optimal toll removes exactly the queue. The sequence of arrivals at the destination is unchanged relative to laissez-faire. The arrival schedule at the bottleneck and at the destination is*

$$a_\tau(c) = \frac{F(c)}{\psi} + a_{\tau 0}.$$

*In general, $a_{\tau 0} \neq a_0$, such that the arrival interval is shifted relative to laissez-faire.*

Social welfare is improved by optimal tolling, since queueing is removed. The toll is a transfer from travellers. So it is of interest to examine whether travellers will be better or worse off under optimal tolling, when the use of revenues does not affect travellers. This is the concern of the following theorem.

**Theorem 6** *Assume conditions 3 and 6. Consider social optimum implemented by a toll with $\tau(a_{\tau 0}) = \tau(a_{\tau 1}) = 0$. Then the schedule of arrivals at the destination is earlier in social optimum than under laissez-faire: $a_{\tau 0} < a_0$. There exists a location $c$ with $a_\tau(c) > a(c)$.*

*There is an interval containing $c_0$ such that all travellers located in this interval are strictly worse off in social optimum than under laissez-faire. There is also an interval containing $c_1$ such that all travellers located in this interval are strictly better off in social optimum than under laissez-faire.*

Under laissez-faire, $a(c_0) < a_*(c_0)$ and $a(c_1) > a_*(c_1)$, and there is no queue for these travellers. The first traveller has a favourable location and can depart quite late and still be first in the queue. The last traveller must travel a longer distance before reaching the bottleneck and so prefers to arrive later at the bottleneck. But because capacity is limited, he arrives even later than that. Due to the longer distance to the bottleneck, the marginal utility of arriving earlier is greater for the last traveller than for the first traveller . This is the mechanism underlying the finding that the arrival interval in social optimum is earlier than under laissez-faire.[8]

The existence of a location $c$ with $a_\tau(c) > a(c)$ shows that even though arrivals at the bottleneck begin and end earlier in social optimum than under laissez-faire, there are travellers who arrive later at the bottleneck in social optimum than under laissez-faire; this is a consequence of the fact that there is no queue in the social optimum. The observation implies a restriction on how much earlier arrivals can shift in social optimum.

It is interesting to contrast the present model with that in Mirrlees (1972a). Maximising the same social welfare function generates inequality of equals (in

[8]There are other situations where optimal pricing causes the arrival interval to shift. Fosgerau and de Palma (2011) consider a parking fee in the bottleneck model where travellers are equipped with general scheduling preferences similar to those in this paper. The first traveller to arrive parks longer and therefore pays more than the last traveller to arrive; this asymmetry causes the arrival interval to shift later relative to the situation with no parking fee. Fosgerau and Small (2011) also consider bottleneck congestion but in a situation where scheduling preferences form endogenously. Here, depending on parameters, the arrival interval can shift in either direction as a consequence of optimal pricing.

terms of both preferences and endowments) in Mirrlees' standard urban model, whereas it reduces the inequality resulting under laissez-faire in the present setup with unequal endowments.[9]

# 5 Numerical illustration

In this section, the theoretical model is illustrated by a numerical simulation. The simulation assumes a continuum of individuals with mass 1. Their scheduling preferences are given by the Vickrey (1973) type of scheduling utility

$$u(d,t) = \int_{T_1}^{d} e^{-s} ds + \int_{t}^{T_2} e^{s} ds.$$

The constants $T_1$ and $T_2$ are arbitrary and are set to 0 at no consequences for the results. The capacity rate of the bottleneck is $0.5$ individuals per hour, such that all can pass the bottleneck in two hours. The distribution of distances to the bottleneck is bimodal, composed of two beta distributions, each with mass $1/2$. One has support on $[1, 1.5]$ and the other has support on $[1.5, 2]$. The density of distances is shown in Figure 2. The distribution of distances is taken to be bimodal, to show how such a feature carries through to the simulation results. Recall in particular Theorem 4, which states that the queue is unimodal, regardless of the spatial distribution of trip origins.

Given a value of the first arrival time at the bottleneck $a_0 = a(c_0)$, the simulation solves the differential equation (1) numerically to find $a(c)$. Then a search is carried out for the value of $a_0$ that solves $a(c_1) = a_0 + N/\psi$. The simulation results are shown in Figure 3. The figure is discussed in detail in the following. It shows various functions of the distance to the bottleneck. In addition to the arrival time at the bottleneck, the figure also shows the departure time from home, and the arrival time at the destination. The figure furthermore shows $a_*$, the preferred arrival time at the destination if there was no queue, and it shows the optimum arrival time at the destination.

It is instructive to begin by noting the preferred arrival time $a_*$. The specification of symmetric scheduling utility rates $\beta(t) = \gamma(-t)$ implies that $a_*(c) = c/2$.

Consider now the laissez-faire equilibrium. The first traveller to arrive at the bottleneck is the one located at the least distance $c_0$. He arrives at the bottleneck at time $a_0$, which is earlier than $a_*(c_0)$ and arrives at the destination at the same time $a_0$, since there is no queue for him. Similarly, the last traveller is the one located at the maximum distance $c_1$. He arrives at the bottleneck at time $a_1 > a_*(c_1)$ and the queue is exactly gone at this time.

[9] We thank a reviewer for contributing this observation.

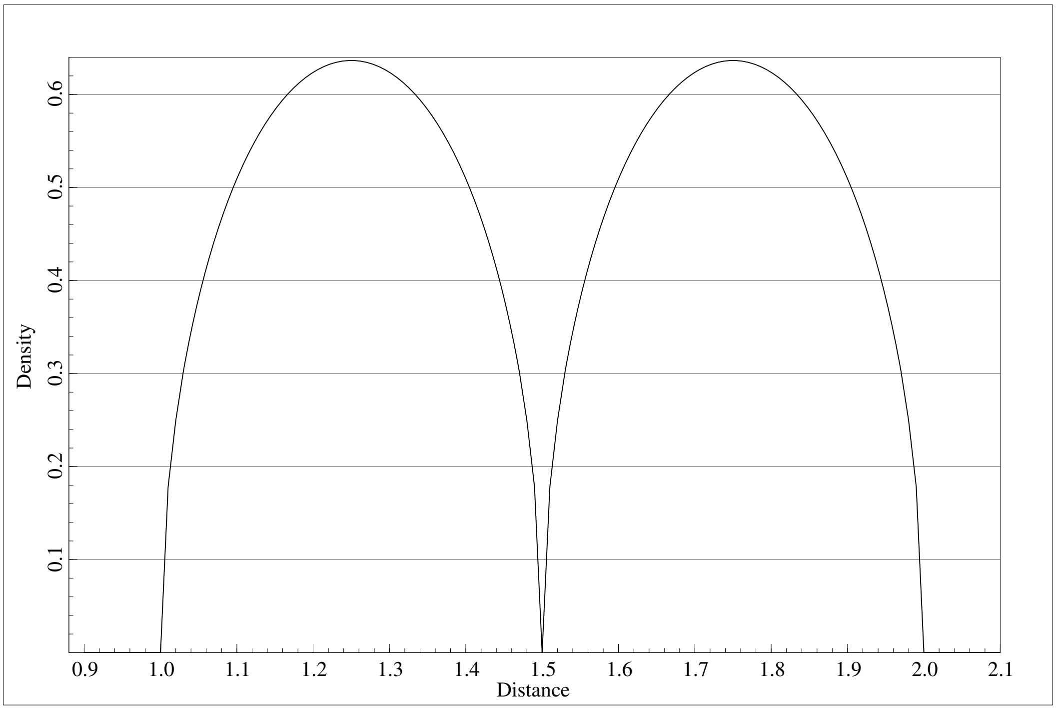


Figure 2: The density of distances to the bottleneck

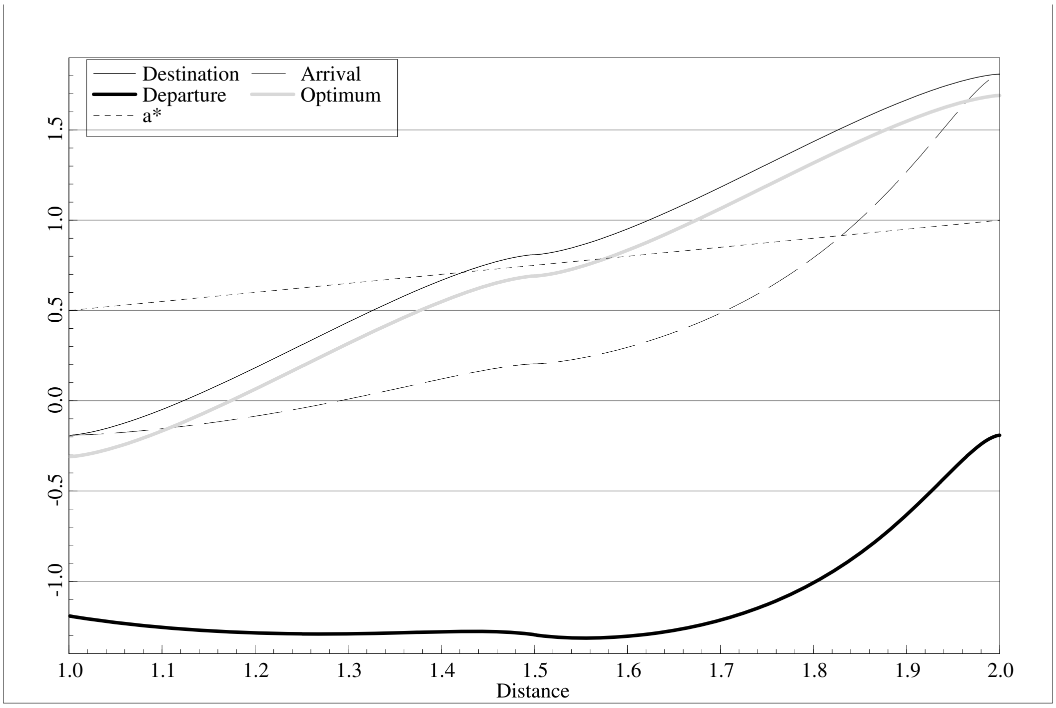


Figure 3: Departure and arrival times

Figure 4 shows the length of the queue as a function of the location of the travellers. This is not a concave function but it is unimodal as shown in Theorem 4.

Returning to Figure 3, consider next the departure time function. This is evidently not a monotone function. In this simulation, the traveller at distance 1.54 departs as the first at time about -1.2. Travellers located closer depart slightly later. This confirms the general finding that more distant travellers will not always depart earlier. For travellers located further away there is almost a monotonous relationship whereby more distant travellers depart later.

Consider now the social optimum. The arrival time at the destination has the same functional form as the laissez-faire arrival time at the destination $F\left(c\right)/\psi+a_{\tau 0}$, where $a_{\tau 0}$ is a constant. This happens because the bottleneck capacity is fixed and the sequence of arrivals at the bottleneck is unchanged in optimum relative to laissez-faire. In the simulation, $a_{\tau 0}$ is found numerically to maximize average scheduling utility. The gray curve on Figure 3 shows the optimum arrival time at the bottleneck as a function of distance. It is also the arrival time at the destination, since there is no queue in optimum. The simulation confirms the result from Theorem 6 that the optimum arrival time is earlier than the laissez-faire arrival time. In this case, the first traveller departs about 0.12 hours earlier in optimum

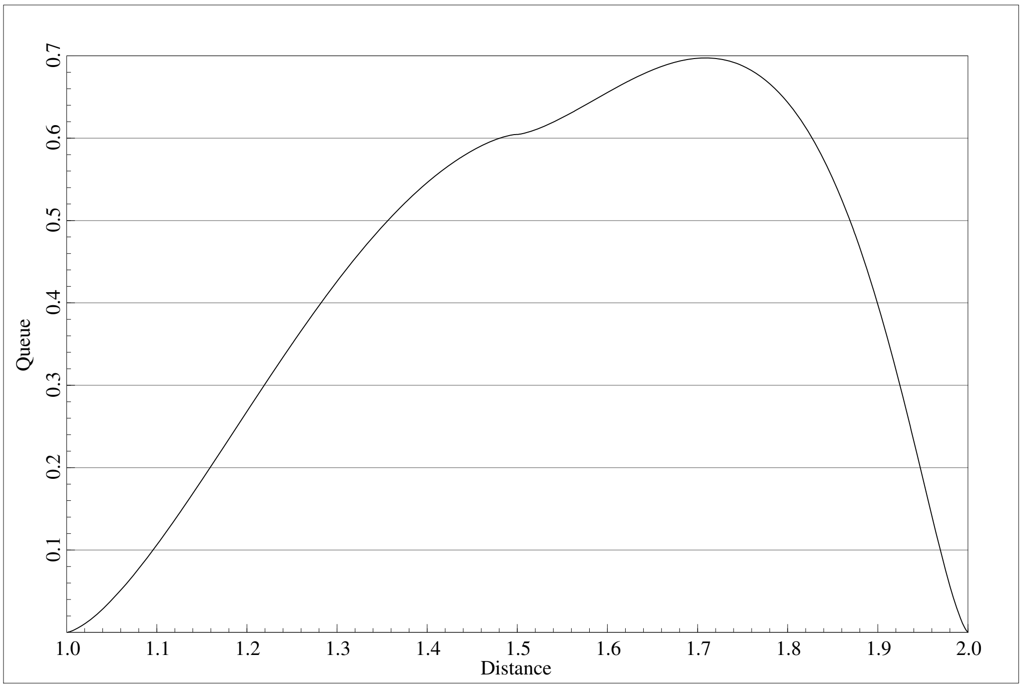


Figure 4: Length of the queue

than in laissez-faire. This means that the traveller at $c_0$ is worse off in equilibrium since he already arrives before his preferred arrival time. Conversely, the traveller at $c_1$ is better off.

Figure 5 shows the utilities achieved by individuals at different locations. The fat line shows the indirect utility of individuals under laissez-faire, consisting of scheduling utility only. It is decreasing in the distance to the bottleneck. The upper thin line shows the scheduling utility in social optimum. The difference between that and the laissez-faire utility, weighted by the density of individuals at different locations, is the welfare gain from tolling. The lower thin line shows the indirect utility in social optimum, equal to the scheduling utility minus the toll. So the toll is visible as the difference between the two thin lines. The indirect utility in social optimum is decreasing but less steeply than under laissez-faire, such that difference ranges from negative to positive.

It is clearly visible how travellers located near the bottleneck lose in social optimum while those far away gain. The indirect utility difference between the individual nearest and furthest from the bottleneck is reduced from 2.9 under laissez-faire to 2.1 in social optimum.

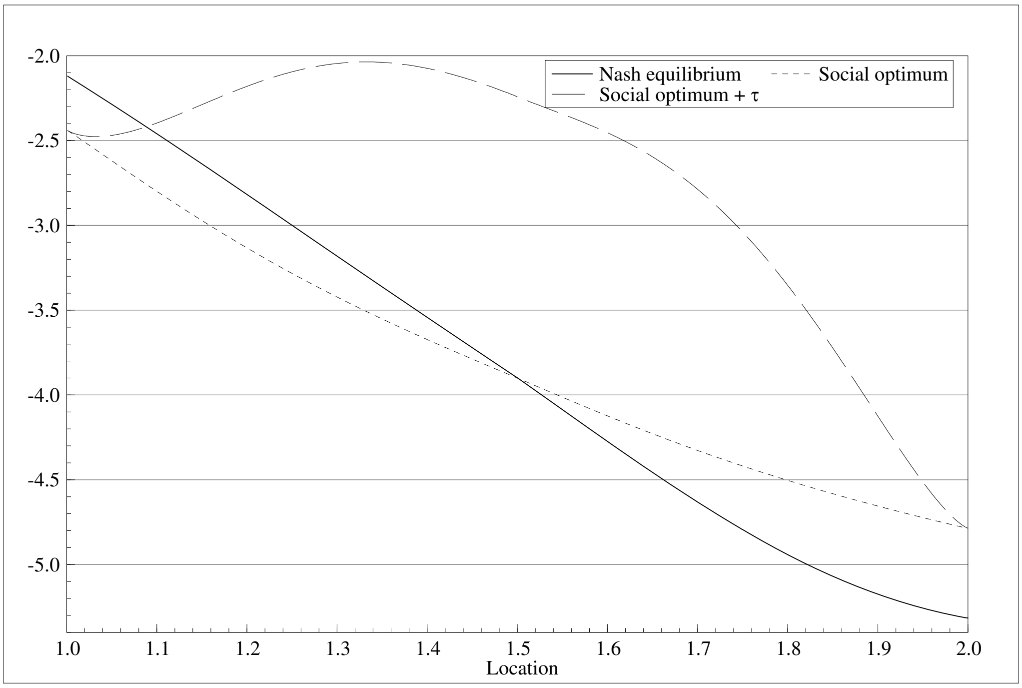


Figure 5: Utilities achieved at different locations

# 6 Conclusion

This paper has introduced spatial heterogeneity into the bottleneck model such that it can be used to represent a city with a central bottleneck. A number of new insights are generated from the model. Perhaps the most important insight is that travellers located near the bottleneck will tend to lose from optimal tolling, while those located far away will tend to gain, when the use of toll revenues is not accounted for. The paper also shows that a reason for the congested demand peaks to be uni-modal can be found in the properties of equilibrium in combination with our general specification of scheduling preferences.

A relevant question, raised by a reviewer, is whether equal redistribution of the toll revenue would make the social optimum a Pareto improvement relative to laissez-faire. This would be a significant finding, indicating that more elaborate mechanisms for revenue recycling might not be needed to gain acceptability of pricing.

The crucial property that generates sorting both under laissez-faire and in the social optimum is that the schedule of arrivals at the bottleneck changes with derivative having the same sign as $-(u_{11}+u_{12})$. Condition 3 then ensures that the derivative is strictly positive. Strict concavity of scheduling utility requires that

$u_{11} + 2u_{12} + u_{22} < 0$, but it is still possible to formulate scheduling utility with $u_{11} + u_{12} > 0$ for some values of $d, t$. Future work could investigate the properties of equilibria under such scheduling utility. The illustrative case of $\alpha - \beta - \gamma$ scheduling utility discussed in the Introduction does not fit into this framework as it is piecewise linear and hence not strictly concave.

The spatial distribution of travellers is a source of heterogeneity in the model. It would be of interest to introduce other sources of heterogeneity into the model. One issue would be the robustness of the sorting property. Another kind of extension would be to introduce risk into the model, for example in the form of random capacity (Arnott et al., 1999) or random queue sorting (de Palma and Fosgerau, 2011).

Perhaps the most interesting extension would be to make the location of individuals endogenous as in the Mirrlees (1972a) standard urban model. This would tie together congestion dynamics and urban economic models. For example Arnott (1998) combines a model of urban spatial structure with the $\alpha - \beta - \gamma$ bottleneck model; optimal tolling does not change transport costs for travellers so when the revenues are not returned, optimal tolling will have no effect on urban structure. As Arnott (1998) notes, this is in contrast to urban economic models with static congestion. However, the Arnott (1998) result is a consequence of space essentially being assumed away in the specification of preferences as was discussed in the Introduction to this paper.

# References

Arnott, R. A. (1998) Congestion Tolling and Urban Spatial Structure *Journal of regional science* **38**(3), 495–504.

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.

Arnott, R. A., de Palma, A. and Lindsey, R. (1994) The Welfare Effects of Congestion Tolls with Heterogeneous Commuters *Journal of Transport Economics and Policy* **28**(2), 139–161.

Arnott, R. A., de Palma, A. and Lindsey, R. (1999) Information and time-of-usage decisions in the bottleneck model with stochastic capacity and demand *European Economic Review* **43**(3), 525–548.

Arnott, R., de Palma, A. and Lindsey, R. (1991) A temporal and spatial equilibrium analysis of commuter parking *Journal of Public Economics* **45**(3), 301–335.

Arnott, R. and DePalma, E. (2011) The corridor problem: Preliminary results on the no-toll equilibrium *Transportation Research Part B: Methodological* **45**(5), 743–768.

Daganzo, C. F. (2007) Urban gridlock: Macroscopic modeling and mitigation approaches *Transportation Research Part B: Methodological* **41**(1), 49–62.

de Palma, A. and Fosgerau, M. (2009) Random queues and risk averse users *Working Paper* .

de Palma, A. and Fosgerau, M. (2011) Random queues and risk averse users *Mimeo* . DTU Transport.

Fosgerau, M. and de Palma, A. (2011) Parking fees as a substitute for roadpricing in a dynamic model of traffic congestion *Working Paper* .

Fosgerau, M. and Engelson, L. (2011) The value of travel time variance *Transportation Research Part B: Methodological* **45**(1), 1–8.

Fosgerau, M. and Small, K. A. (2011) Endogenous scheduling preferences and congestion *Working Paper* .

Geroliminis, N. and Daganzo, C. F. (2008) Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings *Transportation Research Part B: Methodological* **42**(9), 759–770.

Geroliminis, N. and Levinson, D. M. (2009) Cordon pricing consistent with the physics of overcrowding *Proceedings of the 18th International Symposium on Transportation and Traffic theory* .

Hendrickson, C. and Kocur, G. (1981) Schedule Delay and Departure Time Decisions in a Deterministic Model *Transportation Science* **15**(1), 62–77.

il Mun, S. (1999) Peak-Load Pricing of a Bottleneck with Traffic Jam *Journal of Urban Economics* **46**(3), 323–349.

Kuwahara, M. (1990) Equilibrium Queueing Patterns at a Two-Tandem Bottleneck during the Morning Peak *Transportation Science* **24**(3), 217–229.

Lindsey, R. (2004) Existence, Uniqueness, and Trip Cost Function Properties of User Equilibrium in the Bottleneck Model with Multiple User Classes *Transportation Science* **38**(3), 293–314.

Mirrlees, J. A. (1972a) The Optimum Town *Swedish Journal of Economics* **74**(1), 114–135.

Mirrlees, J. A. (1972b) The optimum town *Swedish Journal of Economics* **74**(1), 114–135.

Newell, G. F. (1987) The Morning Commute for Nonidentical Travelers *Transportation Science* **21**(2), 74–88.

Smith, M. J. (1984) The Existence of a Time-Dependent Equilibrium Distribution of Arrivals at a Single Bottleneck *Transportation Science* **18**(4), 385–394.

Tseng, Y. Y. and Verhoef, E. T. (2008) Value of time by time of day: A stated-preference study *Transportation Research Part B: Methodological* **42**(7-8), 607–618.

van den Berg, V. and Verhoef, E. T. (2011) Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay *Journal of Public Economics* **95**(7-8), 983–992.

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.

# A Appendix

**Proof of Theorem 1.** The optimal arrival time $a_* (c)$ exists uniquely by the assumptions on $u$. Note that the first order condition for utility maximization is that

$$u_1 (a_* (c) - c, a_* (c)) + u_2 (a_* (c) - c, a_* (c)) = 0$$

and that differentiating the first order condition with respect to $c$ shows that

$$a'_* (c) = \frac{u_{11} + u_{12}}{u_{11} + 2u_{12} + u_{22}}. \tag{2}$$

Then condition 3 implies that $a'_* (c) > 0$.If also condition 4 holds, then $a'_* (c) < 1$. ■

**Lemma 1** *If $a_0 \leq a_* (c_0)$, $a_0 + N/\psi = a_1 \geq a_* (c_1)$, and if there is no queue at time $a_1$, then, taking the behavior of all other travellers as given, any traveller will choose arrival time at the bottleneck in the interval $[a_0, a_1]$.*

**Proof.** Consider an arbitrary traveller located at $c$. Then $a_0 \leq a_* (c_0) \leq a_* (c) \leq a_* (c_1) \leq a_1$. Therefore the traveller prefers to arrive at the bottleneck at time $a_0$ to any time before, since there is no queue at time $a_0$. Similarly, he prefers arriving at the bottleneck at time $a_1$ to any time after, since the queue is gone at time $a_1$. Therefore he will choose to arrive at the bottleneck during $[a_0, a_1]$. ■

## A.1 Equilibrium

**Lemma 2** *The arrival rate given by (1) and*

$$\int_{c_0}^{c_1} a'(c)\, dc = N/\psi \tag{3}$$

*satisfies* $a(c_0) \leq a_*(c_0)$ *and* $a(c_1) \geq a_*(c_1)$ *when (5) holds.*

**Proof.** Note first that $a(c) > a_*(c)$ implies that $a'(c) > f(c)/\psi$. So it is not possible to have $a(c) > a_*(c)$ for all $c$, since then $a_1 - a_0 > N/\psi$ contradicting (3).

Assume now that $a_0 > a_*(c_0)$. Let $c'$ be the first $c$ with $a(c) = a_*(c)$. Then $a(c') > a_0 + F(c)/\psi > a_*(c_0) + F(c)/\psi > a_*(c')$, which is a contradiction. Hence $a(c_0) \leq a_*(c_0)$ follows.

Assume now that $a_1 < a_*(c_1)$. If $a_0 + F(c)/\psi < a_*(c)$ for all $c$ then $a'(c) < f(c)/\psi$ for all $c$, which is a contradiction with (3). So there is a last $c''$ with $a_0 + F(c'')/\psi = a_*(c'')$. Now $N/\psi - F(c'')/\psi = a_1 - a_*(c'') < a_*(c_1) - a_*(c'') < N/\psi - F(c'')/\psi$, by (5). This is a contradiction and hence $a(c_1) \geq a_*(c_1)$ follows. ■

**Proof of Theorem 2.** Note first that $a_*(c)$ is strictly increasing in $c$ such that $a_*(c_0) < a_*(c_1)$. Consider laissez-faire equilibrium and let $[a_0, a_1]$ be the smallest interval containing all arrivals at the bottleneck. Then $a_0 \leq a_*(c)$ for all $c$, since otherwise it would be possible for some to postpone arrival at the queue without meeting congestion and increase utility. Hence $a_0 \leq a_*(c_0)$. The argument for $a_1 \geq a_*(c_1)$ is similar. If $a_1 - a_0 > N/\psi$, then there will exist an interval where the bottleneck capacity is not fully utilized and where it will be possible for some to relocate to increase utility, since $a_*(c_1) - a_*(c_0) < N/\psi$ by Condition 5. This would contradict equilibrium. If $a_1 - a_0 < N/\psi$, then there is a residual queue at time $a_1$ and the last individual to arrive could postpone departure from the destination without delaying arrival. This would lead to a strict increase in utility which would contradict equilibrium. Hence $a_1 = a_0 + N/\psi$. ■

**Proof of Theorem 3.** Assume first that equilibrium exists. By Theorem 2, there is always queue during $[a_0, a_1]$, such that $R(a) \geq \psi(a - a_0)$. The first order condition for utility maximization for an individual located at distance $c$ is

$$0 = u_1\left(a - c, \frac{R(a)}{\psi} + a_0\right) + u_2\left(a - c, \frac{R(a)}{\psi} + a_0\right)\frac{\rho(a)}{\psi} \tag{4}$$

and the corresponding second order condition is (suppressing some notation)

$$0 > u_{11} + 2u_{12}\frac{\rho}{\psi} + u_{22}\left(\frac{\rho}{\psi}\right)^2 + u_2\frac{\rho'}{\psi}. \tag{5}$$

Denote the solution by $a(c)$. Achieved utility $u\left(a(c) - c, \frac{R(a(c))}{\psi} + a_0\right)$ for an individual at $c$ satisfies

$$\begin{aligned}\frac{\partial}{\partial c}u\left(a(c) - c, \frac{R(a(c))}{\psi} + a_0\right) &= u_1 \cdot (a' - 1) + u_2\frac{\rho}{\psi}a' \\ &= -u_1 < 0\end{aligned}$$

by the first order condition (4). Then utility is decreasing in the distance from the bottleneck.

Differentiate the first order condition (4) with respect to $c$ to find that

$$\begin{aligned}0 &= u_{11} \cdot (a' - 1) + u_{12} \cdot \frac{\rho}{\psi}(2a' - 1) + u_{22} \cdot \frac{\rho^2}{\psi^2} \cdot a' + u_2 \cdot \frac{\rho}{\psi}a' \\ &= \frac{\partial^2 u}{\partial a^2} \cdot a'(c) - (u_{11} + u_{12}).\end{aligned}$$

By Condition 3, $a'(c) > 0$. Then $a(\cdot)$ has an inverse $c(\cdot)$ with $a(c(a)) = a$ and $c'(a) = 1/a'(c(a)) > 0$. In this case, $R(a) = F(c(a))$, such that

$$\rho(a) = \frac{f(c(a))}{a'(c(a))}. \tag{6}$$

The first order condition (4) then shows that

$$a'(c) = \frac{f(c)}{\rho(a(c))} = -\frac{u_2\left(a(c) - c, \frac{F(c)}{\psi} + a_0\right)}{u_1\left(a(c) - c, \frac{F(c)}{\psi} + a_0\right)}\frac{f(c)}{\psi},$$

such that $a(\cdot) > 0$ is determined from $a(c_0) = a_0$ by $a(c) = a_0 + \int_{c_0}^{c} a'(\zeta)\,d\zeta$.

Equilibrium requires that $a(c_1) = a_0 + N/\psi$. This defines $a_0$ uniquely as the following argument shows. Note first that

$$\begin{aligned}\frac{\partial a'(c)}{\partial a_0} &= \left[\frac{u_2}{u_1^2}\left(u_{11}\frac{\partial a(c)}{\partial a_0}+u_{12}\right)-\frac{u_{12}\frac{\partial a(c)}{\partial a_0}+u_{22}}{u_1}\right]\frac{f(c)}{\psi}\\ &= \frac{u_2}{u_1^2}\left[\left(u_{11}-\frac{u_1}{u_2}u_{12}\right)\left(\frac{\partial(a(c)-a_0)}{\partial a_0}+1\right)+\left(u_{12}-\frac{u_1}{u_2}u_{22}\right)\right]\frac{f(c)}{\psi}.\end{aligned}$$

This is strictly positive by conditions 1 and 2 if $\frac{\partial(a(c)-a_0)}{\partial a_0} \geq 0$. Note next that

$$\frac{\partial(a(c)-a_0)}{\partial a_0} = \int_{c_0}^{c}\frac{\partial a'(c)}{\partial a_0}dc.$$

Then $\frac{\partial(a(c)-a_0)}{\partial a_0} > 0$ if $\frac{\partial(a(\zeta)-a_0)}{\partial a_0} \geq 0$ for all $\zeta < c$. Also, $\frac{\partial(a(c_0)-a_0)}{\partial a_0} = 0$ and $\frac{\partial a'(c_0)}{\partial a_0} > 0$ such that $\frac{\partial(a(c)-a_0)}{\partial a_0} \geq 0$ for $c$ in a small neighborhood around $c_0$. Therefore $\frac{\partial(a(c)-a_0)}{\partial a_0} > 0$ for all $c > c_0$. In particular, $\frac{\partial(a(c_1)-a_0)}{\partial a_0} > 0$. Since equilibrium requires that $a(c_1)-a_0 = N/\psi$, there can only be one equilibrium.

It remains to show that equilibrium exists, i.e. that there exists $a_0$ such that $a(c_1)-a_0 = N/\psi$. It is sufficient to show that there are values of $a_0$ such that $\int_{c_0}^{c_1} a'(c)\,dc$ can attain values both larger and smaller than $N/\psi$. Consider therefore first $a_0 + N/\psi < a_*(c_0)$. Then by conditions 1 and 2, if for some $c$

$$a(c) \leq \frac{F(c)}{\psi}+a_0 \tag{7}$$

then also

$$a'(c) = -\frac{u_2\left(a(c)-c, \frac{F(c)}{\psi}+a_0\right)}{u_1\left(a(c)-c, \frac{F(c)}{\psi}+a_0\right)} < 1.$$

But (7) holds near $c_0$ and therefore both inequalities hold for all $c$. Then

$$\int_{c_0}^{c_1} a'(c)\,dc < \int_{c_0}^{c_1}\frac{f(c)}{\psi}dc = \frac{N}{\psi}.$$

The opposite inequality can be obtained for $a_0 > a_*(c_1)$. This establishes existence of equilibrium.

We have shown that (1) together with (3) has a unique solution. This does not depend on the existence of equilibrium. Using the above arguments, it is then easy to see that no individual can improve his arrival time at the bottleneck within $[a_0, a_1]$. It remains to be shown that $a(c_0) \leq a_*(c_0)$ and $a(c_1) \geq a_*(c_1)$. But this is shown in Appendix Lemma 2. Therefore, by Lemma 1, the proposed solution does indeed define an equilibrium, which then exists uniquely. ■

**Proof of Theorem 4.** Using 1, the derivative of the equilibrium queue length is found to be

$$q'(c) = \frac{f(c)}{\psi} \frac{u_1\left(a(c) - c, a_0 + \frac{F(c)}{\psi}\right) + u_2\left(a(c) - c, a_0 + \frac{F(c)}{\psi}\right)}{u_1\left(a(c) - c, a_0 + \frac{F(c)}{\psi}\right)}.$$

Notation for the point where $u$ is evaluated is suppressed in the remainder of the proof. The derivative $q'$ has the same sign as $u_1 + u_2$. We shall show that $u_1 + u_2 < 0$ implies that $u_1 + u_2$ is decreasing. This is sufficient to guarantee that $q$ is quasiconcave. So compute

$$\begin{aligned}
\frac{\partial(u_1 + u_2)}{\partial c} &= (u_{11} + u_{12})(a' - 1) + (u_{12} + u_{22})\frac{f}{\psi} \\
&= -(u_{11} + u_{12})\left(\frac{u_2}{u_1}\frac{f}{\psi} + 1\right) + (u_{12} + u_{22})\frac{f}{\psi} \\
&< -(u_{11} + u_{12})\frac{u_1 + u_2}{u_1} + (u_{12} + u_{22})
\end{aligned}$$

where the second equality follows from (1) and the inequality follows from the assumptions of the Theorem. Then by Conditions 3 and 4, $u_1 + u_2 < 0 \implies \frac{\partial(u_1+u_2)}{\partial c} < 0$. ■

## A.2 Optimal tolling

**Proof of Theorem 5.** In optimum, arrivals at the bottleneck take place during an interval of length $N/\psi$. If there is queue at some point, then arrivals can be delayed such that welfare is improved and the queue is reduced. Hence there is no queue in social optimum and $R_\tau(a) = \psi(a - a_{\tau 0})$, $\rho_\tau(a) = \psi$. Denote by $a_\tau(\cdot)$ the optimal arrival time at the bottleneck under optimal tolling for an individual located at $c$. The first order condition for utility maximization is

$$\tau'(a_\tau(c)) = u_1(a_\tau(c) - c, a_\tau(c)) + u_2(a_\tau(c) - c, a_\tau(c))$$

and the second order condition is

$$\tau''(a_\tau(c)) > u_{11}(a_\tau(c) - c, a_\tau(c)) + 2u_{12}(a_\tau(c) - c, a_\tau(c)) + u_{22}(a_\tau(c) - c, a_\tau(c)).$$

Differentiate the first order condition with respect to $c$ to find that

$$u_{11} \cdot (a'_\tau - 1) + u_{12} \cdot (2a'_\tau - 1) + u_{22} a'_\tau - \tau'' \cdot a'_\tau = 0, \tag{8}$$

which implies that $a'_\tau > 0$ by the second order condition and condition 3. Then $R_\tau (a_\tau (c)) = F(c) = \psi (a_\tau (c) - a_{\tau 0})$, such that $a_\tau (c) = \frac{F(c)}{\psi} + a_{\tau 0}$, and the inverse of $a_\tau$ is $c_\tau (a) = F^{-1} (\psi (a - a_{\tau 0}))$. So now from the first order condition,

$$\tau'(a) = u_1 \left(a - F^{-1} (\psi (a - a_{\tau 0})), a\right) + u_2 \left(a - F^{-1} (\psi (a - a_{\tau 0})), a\right), \tag{9}$$

such that

$$\begin{aligned} \tau(a) - \tau(a_{\tau 0}) &= \int_{a_{\tau 0}}^{a} \tau'(s)\, ds \\ &= \int_{a_{\tau 0}}^{a} \left(u_1 \left(s - F^{-1} (\psi (s - a_{\tau 0})), s\right) + u_2 \left(s - F^{-1} (\psi (s - a_{\tau 0})), s\right)\right) ds. \end{aligned}$$

Any such toll ensures that individuals located at $c$ will prefer to arrive at the bottleneck at time $a_\tau (c)$ to any other time during $[a_{\tau 0}, a_{\tau 1}]$.

The average scheduling utility under tolling is

$$Eu = \int_{c_0}^{c_1} u \left(\frac{F(c)}{\psi} + a_{\tau 0} - c, \frac{F(c)}{\psi} + a_{\tau 0}\right) f(c)\, dc.$$

Differentiate with respect to $a_0$ to find that (writing $u_1$ for $u_1 \left(\frac{F(c)}{\psi} + a_0 - c, \frac{F(c)}{\psi} + a_0\right)$ etc.)

$$\begin{aligned} \frac{\partial Eu}{\partial a_0} &= \int_{c_0}^{c_1} (u_1 + u_2) f(c)\, dc \\ \frac{\partial^2 Eu}{\partial (a_0)^2} &= \int_{c_0}^{c_1} (u_{11} + 2u_{12} + u_{22}) f(c)\, dc < 0. \end{aligned}$$

So $Eu$ is concave as a function of $a_0$; $\frac{\partial Eu}{\partial a_0}$ is positive for some $a_0$ sufficiently smaller than $0$ and negative for some $a_0$ sufficiently larger than $0$. This follows from the assumption that $u_1 + u_2$ varies over all of $\mathbb{R}$. Then $Eu$ attains a global maximum. Make the change of variable $a = \frac{F(c)}{\psi} + a_{\tau 0}$ in the equation $\frac{\partial Eu}{\partial a_0} = 0$ to find that the optimal location of the interval of arrival at the bottleneck is determined by

$$0 = \int_{a_{\tau 0}}^{a_{\tau 0} + N/\psi} (u_1 (a - c, a) + u_2 (a - c, a))\, da.$$

Then by (9), $\tau(a_{\tau 0}) = \tau(a_{\tau 1})$.

There remains the possibility that individuals may prefer a time outside this interval. By Lemma 1, this will not happen if the toll is such that $\tau(a_{\tau 0}) = \tau(a_{\tau 1}) = 0$ and $a_{\tau 0} \leq a_*(c_0), a_*(c_1) \leq a_{\tau 1}$.

To verify that $a_{\tau 0} \leq a_*(c_0), a_*(c_1) \leq a_{\tau 1}$, note that $a(c) \lesseqgtr a_*(c) \iff \tau'(a) \gtreqless 0$. The toll is not constant, so there must be a point $c'$ where $a(c') = a_*(c')$. There can only be one such point by (6). The desired conclusion follows.

■

**Proof of Theorem 6.** For a person located at $c \in C$, examine the difference in utility between the cases with and without optimal tolling.

$$\Delta u(c) = u(a_\tau(c) - c, a_\tau(c)) - \tau(a_\tau(c)) - u\left(a(c) - c, \frac{F(c)}{\psi} + a_0\right).$$

Differentiate this expression with respect to $c$ and insert from first order conditions to find that

$$\frac{\partial \Delta u(c)}{\partial c} = u_1\left(a(c) - c, \frac{F(c)}{\psi} + a_0\right) - u_1(a_\tau(c) - c, a_\tau(c)).$$

Therefore the utility difference can be expressed as

$$\Delta u(c) = \Delta u(c_0) + \int_{c_0}^{c} \left[u_1\left(a(\zeta) - \zeta, \frac{F(\zeta)}{\psi} + a_0\right) - u_1(a_\tau(\zeta) - \zeta, a_\tau(\zeta))\right] d\zeta.$$

Note now that $a_0 + F(c_0)/\psi < a_*(c_0)$ and $a_0 + F(c_1)/\psi > a_*(c_1)$ and that the curves $a_0 + F(c)/\psi$ and $a_*(c)$ intersect in the interior of $C$. Similarly, $a_\tau(c_0) < a_*(c_0)$ and $a_\tau(c_1) > a_*(c_1)$ and the curves $a_\tau(c) = a_{\tau 0} + F(c)/\psi$ and $a_*(c)$ intersect in the interior of $C$. Thus, $\Delta u(c_0) \cdot \Delta u(c_1) < 0$, i.e. one of the endpoint differences must be positive and the other must be negative. Which is positive depends on the sign of $a_{\tau 0} - a_0$.

It must be the case that $a_{\tau 0} < a_0$ : Otherwise $a_\tau(c) \geq a_0 + F(c)/\psi \geq a(c)$ for all $c$, which implies that $\Delta u(c_0) \geq 0$, $\frac{\partial \Delta u(c)}{\partial c} < 0$ for interior $c$ by condition 3 and $\Delta u(c_1) \leq 0$, which is a contradiction.

Since there is no queue for commuters located at $c_0$ and $c_1$, we have $\Delta u(c_0) < 0 < \Delta u(c_1)$. There exists a $c$ with $a_\tau(c) > a(c)$: Otherwise the curves $a(\cdot)$ and $a_\tau(\cdot)$ do not intersect, which implies that

$$u_1(a_\tau(c) - c, a_\tau(c)) > u_1\left(\frac{F(c)}{\psi} + a_0 - c, \frac{F(c)}{\psi} + a_0\right) > u_1\left(a(c) - c, \frac{F(c)}{\psi} + a_0\right)$$

for all $c$. Then $\frac{\partial \Delta u(c)}{\partial c} < 0$, which is a contradiction.

The same argument shows that $\frac{\partial \Delta u(c)}{\partial c}|_{c=c_0} < 0$ and $\frac{\partial \Delta u(c)}{\partial c}|_{c=c_1} < 0$. $\blacksquare$