Congestion costs in bottleneck equilibrium with stochastic capacity and demand

Fosgerau, Mogens

Technical University of Denmark

2008

Congestion costs in bottleneck equilibrium with stochastic capacity and demand*

Mogens Fosgerau, mf@transport.dtu.dk
Technical University of Denmark

February 11, 2008

Abstract

I analyse congestion costs in the Vickrey bottleneck model of a congestible facility with a peak load in demand. The shape of the peak is endogenous, being the sum of individual scheduling decisions. Capacity and demand are random, which introduces uncertainty into the individual scheduling choices. These are essential features of actual peak loads. Based on work by Arnott, de Palma and Lindsey, I derive the expected marginal and total congestion costs and compare to the case with fixed capacity and demand. Using stylised values for scheduling costs relative to the value of time, I find that randomness of capacity and demand increases congestions cost by up to 50 percent relative to the deterministic case. The bound is general for any distribution of random capacity and demand.

KEYWORDS: Bottleneck model; Scheduling; Reliability; Congestion
JEL codes: D01; D81

*Thanks to Bruno de Borger for comments to a draft of this paper. Financial support from the Danish Social Science Research Council is acknowledged.

1 Introduction

Traffic congestion imposes significant costs on societies and understanding the phenomenon is important in order to devise appropriate policy responses. Congestion as a large component of travel cost is also an important determinant of urban sprawl (e.g. Anas and Rhee, 2007). Traffic congestion is an extremely complicated phenomenon. Not only do travel times increase as traffic volumes increase towards capacity, travel times also become increasingly random and unpredictable due to the chaotic behaviour of traffic at the micro level. Day-to-day variation in the number of travellers are hard to predict and causes further variation in travel times. Finally, scheduling decisions by travellers add to the complexity. For example in the morning peak, commuters may decide to depart earlier or later in order to avoid the high mean and standard deviation of travel time in the middle of the peak. Then the shape of the peak is endogenous, being an aggregate of individual scheduling decisions.

This paper analyses the simplest structural model for congestion, the bottleneck model, allowing for the endogeneity of scheduling choices. The bottleneck model was proposed by Vickrey (1969) and analysed in detail by Arnott et al. (1993). The model describes an isolated free access facility, where capacity is governed by a bottleneck with some maximum service rate or capacity. The model is presented for the case of a congested road. It could also be taken to describe other facilities such as many public recreational facilities, and computer systems. Roads are arguably the best and most practically important example (Arnott et al., 1999).

A queue builds up whenever the rate of arrivals at the bottleneck is greater than the service rate and diminishes again when the rate of arrivals is less than the service rate. Users choose departure time in order to minimise travel costs and scheduling costs, where the scheduling costs depend on whether the user is early or late relative to some preferred arrival time. In equilibrium, no user can reduce costs by changing departure time, which means the user cost is the same for all users. It turns out the marginal external congestion cost, the cost to everybody else of adding one user, is exactly equal to the user cost of one user.

Arnott et al. (1999) present an analysis of the bottleneck model where capacity and demand are allowed to be random, as is the case on actual congested roads. They concentrate on the value of providing information to travellers and do not consider marginal costs. The main difficulty here is that they do not describe the equilibrium explicitly, they are only able to derive some features of the equilibrium.

The contribution of the current paper is to provide explicit expressions for the expected costs in this model. This has not previously been achieved under random capacity, but as will be shown, it is possible to make the derivation without explicitly finding the departure rate in the random bottleneck model. These results are employed to find the contribution of randomness to total and marginal external congestion costs. A bound is provided for the additional costs due to random capacity and demand, this bound holds for any distribution of random capacity and demand.

Daniel (1995) applied the bottleneck model with random arrivals and fixed capacity to analyse congestion pricing at a large hub airport. The scheduling cost function was used by Small (1982), who considered the individual timing of trips, when the travel time depends on the departure time. This paper did not consider equilibrium and did not consider random travel times. Noland and Small (1995) introduced random travel times into the scheduling model. They were able to derive the value of reliability for some special cases, taking the travel time distribution as exogenous and without considering equilibrium. Fosgerau and Karlstrom (2007) solved the general case to show that the user cost is linear in the mean and the standard deviation of random travel time, assuming the simplest formulation of scheduling costs and any exogenous, possibly time-dependent, distribution of travel times. Fosgerau and Karlstrom still take the individual perspective and do not take account of the endogeneity of the distribution of random travel time. In applications, this endogeneity must be handled through an additional model for traffic flow. This issue does not arise with the present analysis that takes both randomness and equilibrium into account.

The paper is organised as follows. Section 2 the deterministic and the random bottleneck model. Section 3 derives expressions for congestion costs for the random bottleneck model and compares to the deterministic case. Section 4 concludes with some perspectives for further research.

2 Review of the bottleneck model

2.1 The deterministic bottleneck model

The model presented here is from Arnott et al. (1993) and is an extension of the Vickrey (1969) deterministic bottleneck queuing model. It describes an isolated free access facility, where capacity is governed by a bottleneck with a maximum service rate or capacity of $1/\phi$ vehicles per time unit, where ϕ is the serve time. Service within capacity is instantaneous.[1] A queue builds up whenever the rate of arrivals at the bottleneck is greater than the service rate. There are N individuals, treated as a continuum. They arrive at the bottleneck at the rate $\rho(t)$, starting at some time t_0 and ending at time t_L. The cumulative number of arrivals at time t is denoted

[1] A fixed travel time on a section of road can be added at no loss of generality.

by $R(t) = \int_{t_0}^{t} \rho(t)dt$ and the length of the queue at time t is

$$Q(t) = \left(R(t) - \frac{t - t_0}{\phi}\right)^+,$$

when R is concave, $\phi\rho(t_0) > 1$ and where x^+ is the positive part of x.[2] The queuing time is then

$$q(t, \phi) = Q(t)\phi = (t_0 + \phi R(t) - t)^+$$

and individuals departing at time t are served at time $t + q(t, \phi)$.

Individuals all have time 0 as their preferred arrival time. When they arrive earlier or later than this time they incur a schedule delay cost, defined as a function of the deviation between the serve time and the preferred arrival time. Take the schedule delay cost to be piecewise linear according to

$$D(t) = \beta t^- + \gamma t^+.$$

The user cost is defined as

$$C(t) = D(t+q(t, \phi))+\alpha q(t, \psi) = D\left(t + (t_0 + \phi R(t) - t)^+\right)+\alpha(t_0+\phi R(t)-t)^+,$$

where α is the marginal disutility of travel time per se (DeSerpa, 1971).

[2] More precisely, $x^+ = x$ if $x > 0$ and $x^+ = 0$ otherwise. x^- is defined by $x = x^+ - x^-$.

Assume an equilibrium where t_0, t_L and $R(t)$ are such that $C(t)$ is constant for $t \in [t_0 : t_L]$. Note that there is always queue inside this interval, since otherwise travellers could change their departure time towards the preferred arrival time and reduce costs. In order to find the equilibrium we may therefore differentiate the user cost with respect to time t and set the derivative to zero. Let t^* defined implicitly by $\phi R(t^*) + t_0 = 0$ be the time where travellers change from being early to being late. Then

$$C'(t) = \phi\rho(t)\left(\gamma - (\gamma + \beta)1_{\{t<t^*\}}\right) + \alpha(\phi\rho(t) - 1) = 0.$$

Solve for $\rho(t)$ to find

$$\rho(t) = \frac{\alpha}{\phi(\alpha + \gamma - (\gamma + \beta)1_{\{t<t^*\}})},$$

such that the departure rate ρ is constant on both sides of t^*. We may now find that

$$\begin{aligned}
R(t_0) &= 0 \\
R(t^*) &= (t^* - t_0)\frac{\alpha}{\phi(\alpha - \beta)} \\
R(t_L) &= N = R(t^*) + (t_L - t^*)\frac{\alpha}{\phi(\alpha + \gamma)}
\end{aligned}$$

We also know that the queue is exactly gone at time t_L, since otherwise the last traveller could delay departure to save travel time without changing the arrival time. That is,

$$\phi N = t_L - t_0.$$

Combining this information we find that

$$\begin{aligned} t^* &= \frac{\beta}{\alpha} t_0 \\ t_0 &= -\phi N \frac{\gamma}{\beta + \gamma} \\ t_L &= \phi N \frac{\beta}{\beta + \gamma} \end{aligned}$$

The user costs of all travellers are equal, so we find for example the user cost of the first traveller to be

$$C(t_0) = -\beta t_0 = \phi N \frac{\beta\gamma}{\beta + \gamma}. \tag{1}$$

The total user cost for all users is then $TC = NC(t_0)$ and the marginal cost is

$$\frac{dTC}{dN} = 2\phi N \frac{\beta\gamma}{\beta + \gamma}, \tag{2}$$

where half of this is internal cost to the additional traveller and the other half is the marginal external congestion cost, denoted by MEC_d and the subscript d denotes that this is the deterministic case.

2.2 The stochastic bottleneck model

I use the results of Arnott et al. (1999) as a starting point for the present analysis. A brief summary of their findings follows. I use the piecewise linear schedule delay cost function, although Arnott et al. (1999) use a more general formulation. The difference relative to the deterministic bottleneck model above is that now the serve time ϕ is taken as random with cumulative distribution J and density j. Assume that ϕ is bounded above by ϕ^+. Normalise at no loss of generality to let $N = R(t_L) = 1$. We assume an equilibrium that equalises the expected costs $EC(t)$ for all travellers. The expected cost for a traveller departing at time t is

$$EC(t) = D(t)J\left(\frac{t - t_0}{R(t)}\right) + \int_{\phi^Q(t)}^{\phi^+} [D(t + q(t, \phi)) + \alpha q(t, \phi)]j(\phi)d\phi,$$

where $\phi^Q(t)$ is the minimal serve time ϕ at which queue occurs at time t. Arnott et al. (1999) then present the following findings.

- The cumulative departure rate $R(t)$ is concave in t.
- The first traveller departs earlier than the preferred arrival time and the last traveller departs later: $t_0 < 0 < t_L$.
- The minimal serve time at which queue occurs at time t is $\phi^Q(t) = \frac{t-t_0}{R(t)}$. In particular, $q^Q(t_L) = t_L - t_0$. Queue occurs at time t whenever

$\phi > \phi^Q(t)$.

- The expected cost at the first departure time is $EC(t_0) = -\beta t_0$.
- This is equal to the expected cost at the last departure, which is

$$EC(t_L) = \gamma t_L J(t_L - t_0) + \gamma \int_{t_L - t_0}^{\phi^+} (t_0 + \phi) j(\phi) d\phi + \alpha \int_{t_L - t0}^{\phi^+} (\phi - t_L + t_0) j(\phi) d\phi.$$

- The time between the first and last departures is given by

$$t_L - t_0 = J^{-1}\left(\frac{\alpha}{\alpha + \gamma}\right).$$

3 Congestion costs

In this section I extend the results of Arnott et al. (1999) by deriving expressions for the expected total, internal and marginal external cost. Expected costs are equal at all departure times and hence it suffices to find the marginal cost at t_0. In doing this we may utilise that the times where travel starts and ends, t_0 and t_L are given by $EC(t_0) = EC(t_L)$ and $t_L - t_0 = J^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)$. It is particularly useful that the cost for the first traveller is not random, since she never experiences a queue but only the delays cost associated with being early.

Elaborate the model slightly by letting the random number of travellers be $N\nu_N$, where $E(N\nu_N) = N$. Similarly, the random serve time is $\psi\nu_\psi$, where $E(\psi\nu_\psi) = \psi$. Now let $\phi = N\nu_N\psi\nu_\psi$, $\sigma = N\psi$ and $\nu = \nu_N\nu_\psi$ such that $J(\phi) = H(\nu)$ and $h(\nu)d\nu = j(\phi)d\phi$. We may then rewrite the equations governing the start and end times as follows.

$$\begin{aligned}
EC(t_0) &= C(t_0) = -\beta t_0 && (3)\\
EC(t_L) &= t_L\frac{\alpha\gamma}{\alpha+\gamma} + \gamma\int_{H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)}^{\nu^+}(t_0+\sigma\nu)h(\nu)d\nu \\
&+ \alpha\sigma\int_{H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)}^{\nu^+}\left(\nu - H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)\right)h(\nu)d\nu && (4)\\
t_L - t_0 &= \sigma H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right) && (5)
\end{aligned}$$

We are then ready to find the marginal value of changes to σ.

Proposition 1 *The marginal cost of changes in σ is constant and given by*

$$\frac{dEC(t_0)}{d\sigma} = \beta\frac{\alpha+\gamma}{\gamma+\beta}\int_{\frac{\alpha}{\alpha+\gamma}}^{1} H^{-1}(x)dx.$$

Proof Differentiate (5) with respect to σ to find that

$$\frac{dt_L}{d\sigma} - \frac{dt_0}{d\sigma} = H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right). \tag{6}$$

Differentiate also the expected costs at times t_0 and t_L.

$$\begin{aligned}\frac{dEC(t_0)}{d\sigma} &= -\beta\frac{dt_0}{d\sigma}\\ \frac{dEC(t_L)}{d\sigma} &= \frac{dt_L}{d\sigma}\frac{\alpha\gamma}{\alpha+\gamma}+\gamma\int_{H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)}^{\nu^+}\left(\frac{dt_0}{d\sigma}+\nu\right)h(\nu)d\nu\\ &+ \alpha\int_{H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)}^{\nu^+}\left(\nu-H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)\right)h(\nu)d\nu\\ &= \left(H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)+\frac{dt_0}{d\sigma}\right)\frac{\alpha\gamma}{\alpha+\gamma}+\frac{dt_0}{d\sigma}\frac{\gamma^2}{\alpha+\gamma}\\ &+ (\alpha+\gamma)\Gamma\left(\frac{\alpha}{\alpha+\gamma}\right)-\alpha H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)\frac{\gamma}{\alpha+\gamma}\\ &= \frac{dt_0}{d\sigma}\frac{\alpha\gamma}{\alpha+\gamma}+\frac{dt_0}{d\sigma}\frac{\gamma^2}{\alpha+\gamma}+(\alpha+\gamma)\Gamma\left(\frac{\alpha}{\alpha+\gamma}\right)\end{aligned}$$

where

$$\Gamma(x)=\int_{H^{-1}(x)}^{\nu^+}\nu h(\nu)d\nu=\int_x^1 H^{-1}(x)\,dx$$

Now solve to find that $\frac{dt_0}{d\sigma}=-\frac{\alpha+\gamma}{\gamma+\beta}\Gamma\left(\frac{\alpha}{\alpha+\gamma}\right)$ such that we may use (3) to find that $\frac{dEC(t_0)}{d\sigma}=\beta\frac{\alpha+\gamma}{\gamma+\beta}\int_{\frac{\alpha}{\alpha+\gamma}}^1 H^{-1}(x)dx.$ ∎

This proposition is all that we require to establish explicit expressions for user costs in the model. The following proposition states the results.

Proposition 2 *The total expected cost for all travellers in equilibrium* *is*

$$E(TC) = N^2\beta\frac{\alpha+\gamma}{\gamma+\beta}\psi\int_{\frac{\alpha}{\alpha+\gamma}}^{1} H^{-1}(x)dx.$$

Hence the marginal expected cost per user is

$$2N\beta\frac{\alpha+\gamma}{\gamma+\beta}\psi\int_{\frac{\alpha}{\alpha+\gamma}}^{1} H^{-1}(x)dx,$$

of which half is the internal expected cost and half is the marginal expected external cost, MEC_r.

Proof The marginal total expected cost of increasing the number of travellers is

$$\begin{aligned}\frac{dE(TC)}{dN} &= \frac{dE(N\nu_N C(t_0))}{dN} = \frac{dNC(t_0)E(\nu_N)}{dN}\\ &= \frac{dNC(t_0)}{dN} = C(t_0) + N\frac{dC(t_0)}{dN}.\end{aligned}$$

Proposition 1 shows that

$$\frac{dC(t_0)}{dN} = \beta\frac{\alpha+\gamma}{\gamma+\beta}\psi\int_{\frac{\alpha}{\alpha+\gamma}}^{1} H^{-1}(x)dx.$$

This, in turn, leads to

$$\frac{dE(TC)}{dN} = -\beta t_0 + N\beta\frac{\alpha+\gamma}{\beta+\gamma}\psi\int_{\frac{\alpha}{\alpha+\gamma}}^{1} H^{-1}(x)dx,$$

where $-\beta t_0$ is the internal expected cost and $MEC_r = N\beta\frac{\alpha+\gamma}{\gamma+\beta}\psi\int_{\frac{\alpha}{\alpha+\gamma}}^{1} H^{-1}(x)dx$ is the marginal expected external cost. The expected cost for one traveller is

$$EC = C(t_0) = \int_0^N \frac{dC(t_0)}{dN} dN = N\beta\frac{\alpha+\gamma}{\gamma+\beta}\psi\int_{\frac{\alpha}{\alpha+\gamma}}^{1} H^{-1}(x)dx.$$

Now the total expected cost is $NE(C)$. $\blacksquare$

It is informative to compare the difference in the marginal external cost of a change in the mean number of travellers between the random and deterministic bottleneck cases. We take the fixed serve time ϕ in the deterministic case to be $\psi\int_0^1 H^{-1}(x)dx$.

$$\begin{aligned}\Delta MEC &= N\beta\frac{\alpha+\gamma}{\beta+\gamma}\psi\int_{\frac{\alpha}{\alpha+\gamma}}^{1} H^{-1}(x)dx - N\frac{\beta\gamma}{\beta+\gamma}\psi\int_0^1 H^{-1}(x)dx \\ &= N\frac{\beta}{\beta+\gamma}\psi\left[\alpha\int_{\frac{\alpha}{\alpha+\gamma}}^{1} H^{-1}(x)dx - \gamma\int_0^{\frac{\alpha}{\alpha+\gamma}} H^{-1}(x)dx\right] \qquad (7)\end{aligned}$$

Note that the difference is zero whenever one of α, β, γ or N is zero. The difference tends to zero as the distribution of ν collapses on its mean. Using that H^{-1} is increasing, it is easy to show that the difference is positive.

$$\begin{aligned}\Delta MEC &> N\frac{\beta}{\beta+\gamma}\psi\left[\alpha\left(1-\frac{\alpha}{\alpha+\gamma}\right)H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right) - \gamma\left(\frac{\alpha}{\alpha+\gamma}-0\right)H^{-1}\left(\frac{\alpha}{\alpha+\gamma}\right)\right] \\ &= 0\end{aligned}$$

It is also possible to find an upper bound for the difference, using that

$H^{-1}(x) > 0$. From (7),

$$\begin{aligned}\Delta MEC &< N\frac{\alpha\beta}{\beta+\gamma}\psi\int_0^1 H^{-1}(x)dx \\ &= N\frac{\alpha\beta}{\beta+\gamma}\psi E(\nu) \\ &= \frac{\alpha}{\gamma}MEC_d.\end{aligned}$$

This bound cannot be improved in general, since the distribution H may have arbitrarily much of the mass below the $\frac{\alpha}{\alpha+\gamma}$-quantile located at zero. We collect these insights into a proposition.

Proposition 3 *The marginal external congestion cost in the case with uncertain capacity and demand increases by a factor between* 1 *and* $1+\frac{\alpha}{\gamma}$ *relative to the deterministic case. The difference is zero whenever* $\alpha = 0, \gamma = 0$ *or* $\beta = 0$*. The difference increases in the value of time* α *and in the cost of earliness* β*.*

The same results hold for total costs and internal costs.

Stylised values of $\alpha = 2$, $\beta = 1$ and $\gamma = 4$ are often used. With these values we find that the difference is at most a factor 1.5.

4 Perspectives

Given the cost of capacity provision, the results in this paper may be applied directly to find the optimal capacity. It is also easy to equip the average demand with an elasticity like in Arnott et al. (1999) in order to determine the optimal time-invariant toll.

Arnott et al. (1993) determine the optimal time-varying toll for the deterministic case such that the average cost excluding the toll is minimal. A main outstanding question is to determine an optimal time-varying toll for the case of random capacity and demand. Here it seems to be hard to make progress without being able to find the endogenous arrival rate ρ. Moreover, concavity of the cumulative arrival rate is crucial for the model. It is not obvious that concavity holds in the presence of an optimal time-varying toll.

References

Anas, A. and Rhee, H. J. (2007) When are urban growth boundaries not second-best policies to congestion tolls? *Journal of Urban Economics* **61**(2), 263–286.

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

Daniel, J. I. (1995) Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues *Econometrica* **63**(2), 327–370.

DeSerpa, A. C. (1971) A Theory of the Economics of Time *The Economic Journal* **81**, 828–845.

Fosgerau, M. and Karlstrom, A. (2007) The value of reliability *Working Paper* .

Noland, R. B. and Small, K. A. (1995) Travel-Time Uncertainty, Departure Time Choice, and the Cost of Morning Commutes *Transportation Research Record* **1493**, 150–158.

Small, K. (1982) The scheduling of Consumer Activities: Work Trips *American Economic Review* **72**(3), 467–479.

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