Traffic incidents and the bottleneck model

Stefanie Peer[1], Paul Koster[1], Erik T. Verhoef[1,2] and Jan Rouwendal[1]

1 Department of Spatial Economics, VU University Amsterdam, The Netherlands

2 Tinbergen Institute, The Netherlands

Abstract

This paper analyzes the costs of traffic incidents using a bottleneck model. A traffic incident results in queuing because of a sudden drop in road capacity. We derive the equilibrium departure rate. Equilibrium scheduling and queuing costs which are functions of the probability that an incident occurs can then be analyzed.

Corresponding author: Stefanie Peer. Tel.: +31 20 59 88326. Fax.: +31 20 59 86004.
Email addresses: speer@feweb.vu.nl (S. Peer), pkoster@feweb.vu.nl (P.R. Koster), everhoef@feweb.vu.nl (E.T.Verhoef).
Address: Faculty of Economics and Business Administration. De Boelelaan 1105, 1081HV Amsterdam, The Netherlands.

1. Introduction

Congestion is one of the major problems in urban areas. It results in time losses and it is a key element in the accessibility of places. Different types of congestion have been identified. One distinction is between recurrent and non-recurrent congestion. Recurrent congestion is the regular day-to-day congestion caused by the fact that the number of road users is higher than what is the maximum for freely flowing traffic. Non-recurrent congestion is non-regular congestion, for example caused by traffic incidents. This paper analyzes recurrent and non-recurrent congestion jointly, using a single origin, single destination bottleneck model where the expected queuing time depends on the expected number of drivers queuing at the bottleneck and the capacity of the bottleneck.

Road capacities are not fixed but variable. Incidents such as bad weather conditions, accidents or temporary roadworks might decrease road capacity. As incident moments are generally unknown, individuals face stochastic capacity over the peak, and as a consequence stochastic travel times. Often, road users are not informed whether and when an incident occurs on a specific day. But due to experience, they can learn the incident probability, and can adjust their departure time in order to minimize their expected travel costs. In our model, these consist of cost related to travel time and costs related to arriving earlier or later than their preferred arrival time (scheduling costs).

This paper makes two contributions to the literature. First, we analyze how capacity reductions caused by traffic incidents affect the departure time decisions of individuals during peak periods. Our paper differs from most other papers that study stochastic bottlenecks by the fact that capacity is not constant over the day. Instead, capacity drops can occur at any point in time during the peak period.

Second, we are able to find a unique equilibrium departure pattern under the assumption that traffic incidents occur with a certain probability per unit in time. Compared to the standard bottleneck model, we find that traffic incidents lead to an earlier equilibrium peak starting time and a steeper cumulative departure rate at the beginning of the peak. It also results in a flatter peak. As expected, an increase in the incident probability increases the costs of passing the bottleneck for all drivers.

2. Prior literature

The standard bottleneck model was developed by Vickrey (1969). A more structural formulation is given by Arnott et al. (1993) who explicitly take into account congestion technology and departure time decisions of travelers. In the bottleneck model, congestion occurs due to the fact that traffic demand exceeds the capacity of a bottleneck. As a consequence, a vertical queue develops.[3] The length of the queue at the moment of joining in, together with the capacity of the bottleneck, determines the queuing time of a driver passing the bottleneck. In the un-priced equilibrium, each traveler faces equal costs. Travelers arriving close to their preferred arrival time face the highest scheduling costs. In the social optimum, no queuing takes place as the departure rate is equal to the capacity of the bottleneck. Decentralization of the social optimum can be achieved using a time varying toll.

The basic model has been extended in many directions, among them the implementation of stochastic capacity.

[3] For an overview of other queuing models such as flow-congestion models and car-following models we refer to Verhoef (2003).

Arnott et al. (1996, 1999) and Fosgerau (2009) analyze a bottleneck where both demand and supply are variable. The probabilities of these fluctuations are expressed in the form of general distribution functions. Under this assumption, the authors compare various setups that differ with respect to which information drivers receive on the day-to-day realizations of demand and supply. Lindsey (1996) assumes that capacity is stochastic. If a time varying toll is implemented, he finds that the optimal departure is non-decreasing over the time of the peak. The according toll is concave in time. Under the assumption of a two-point distribution of capacity, the peak starts earlier, or at the same moment compared to the no-toll case. This is an indication that private costs are higher in the toll equilibrium than in the no-toll equilibrium.

Daniel (1995) develops a stochastic bottleneck model for congestion at hub airports. He uses numerical methods to analyze the equilibrium. Also Li et al. (2008) use numerical methods to solve a bottleneck where daily capacity is distributed uniformly between an upper and a lower bound. However, existence of an equilibrium depends on the assumptions of the model parameters. Also Fosgerau and Jensen (2008) show that in a bottleneck with stochastic demand as well as supply an equilibrium not always exists. Their analysis applies for a general distribution of road capacity. Koster and Rietveld (2010) use a simple model with elastic demand to study the effects of incidents and reductions in incident durations, assuming that the probability and the time loss because of an incident are a function of the number of drivers passing the bottleneck. However, they do not derive an equilibrium departure pattern. Furthermore, they assume that free flow travel time prevails if no incident has happened. Therefore only non-recurrent time-losses are included.

In contrast to the papers cited above, Fosgerau (2010) investigates road capacity changes that occur at random moments during the peak rather than on a day-to-day basis, again using the bottleneck model. He considers incidents of fixed duration that may end before the peak is over. Properties of the Nash equilibrium and social optimum are derived under these assumptions. Due to mathematical complexity, Fosgerau uses a “pretty good toll” instead of the socially optimum toll to decentralize the social optimum. The “pretty good toll” assumes that the departure rate is equal to the maximum capacity of the bottleneck over the entire peak.

Capacity changes within peak periods have also been studied using models other than the bottleneck model. For instance, Schrage (2006) uses a model of flow congestion to analyze the effects of accidents. She models the probability of accidents as Poisson process and assumes capacity to decrease sharply after an accident and to gradually return to its previous level thereafter. To derive the equilibrium traffic flow, she uses the Bellman-equation. She finds that the optimal toll captures future congestion costs. As these differ over the peak – they are higher in the early periods of the peak – the marginal social costs of congestion are not constant across the peak either.

3. Model Setup

3.1 Introduction

The setup of the model used in this paper is based on the standard, deterministic bottleneck model developed by Vickrey (1969). We assume a fixed number of individuals N, who travel along the same route and have the same preferred arrival time (PAT), which we set equal to 0. This setup may represent typical morning peak traffic, with all individuals having the same work starting time and the same destination to go to.

The cost function of each individual has three components. First the costs for travel time, second the costs of schedule delay early (SDE) and third, the costs for schedule delay late (SDL). In

Equation 1 we show the costs for an individual departing at time t with value of travel time α, value of SDE β, and value of SDL γ.

$$C(t) = \alpha \cdot (travel\ time) + \beta \cdot (SDE) + \gamma \cdot (SDL) \quad (1)$$

Following conventional assumptions, and empirical estimates by Small (1989), we assume $\beta < \alpha < \gamma$. The former inequality is required for an equilibrium without a mass departure in the deterministic model. Furthermore we assume that the value of travel time and of schedule delay early (β) and late (γ) are time independent. For simplicity, it is assumed that free flow travel time through the bottleneck is 0. Therefore, all costs related to travel time are due to queuing, i.e. the length of the queue at time t divided by the capacity of the bottleneck, denoted by s_0 when there is no incident. The queue is equal to the difference between cumulative departures ($R(t)$) and cumulative arrivals (A) up to point t. In the standard bottleneck model, cumulative arrivals are equal to $(t - tstart) * s_0$. In this paper, we assume that the capacity of the bottleneck is not necessarily constant across the peak.
Instead, an incident can occur at any time between the starting (t_{start}) and the end time of the peak (t_{end}), causing the capacity to drop from its initial level (s_0) to a lower level (s_1). We assume that this decrease in capacity then prevails for the rest of the day and that at most one incident can occur. The probability of a capacity drop (p), is constant over time during the peak, and therefore $p \cdot (t_{end} - t_{start})$ denotes the probability that an incident occurs during the entire peak period.

Day-to-day variations in travel times occur, since both the occurrence of an incident and the incident moments are stochastic. We are interested in the case where individuals have no information on whether and when an incident has occurred or will occur. Given these assumptions, we try to find an equilibrium departure pattern that is based on the individuals' expectations of travel times and schedule delays.

3.2 Probabilities and expected travel times

Given the incident probability p, an individual with departure time t can end up in two different regimes. In scenario 0, the driver is not affected by any incident. Thus, neither before his departure nor during his queuing time in front of the bottleneck an incident occurs. The probability of this scenario is denoted by p_0. In regime 1, the opposite is true. An incident occurs either before the road user has left from home or during his trip. The probability of this scenario is denoted by p_1 . It is equal to $1 - p_0(t)$.

In the following section, we will derive the associated probabilities for both regimes as well as day-to-day and expected travel times.

Regime 0: No incident

The probability of being in regime 0 is the complement of the probability that there is an incident between t_{start} and the arrival time at the end of the bottleneck:

$$p_0(t) = 1 - p \cdot \frac{R(t)}{s_0} \tag{1}$$

So if p=0, the probability of being in regime 0 equals 1. In that case the model reduces to the standard bottleneck model. If there is no incident before and during the trip, the travel time is given by the following expression:

$$T_0(t) = \frac{R(t)-(t-t_{start}\ )\cdot s_0}{s_0} \tag{2}$$

Clearly, in this regime the day-to-day travel time is equal to the expected travel time.

Regime 1: Incident before passing the bottleneck

In this model, the cumulative probability that an incident happens before the driver has arrived at the end of the bottleneck depends on the time that passes between t_{start} and the time the driver leaves the bottleneck. The latter is given by $t + \frac{R(t)-(t-t_{start}\ )s_0}{s_0}$ if there is no incident before the bottleneck is passed. The probability that an incident occurs before a driver departing at t has passed the bottleneck is therefore proportional (with a factor p) to $t - t_{start} + \frac{R(t)-(t-t_{start}\ )s_0}{s_0}$. This expression simplifies to $\frac{R(t)}{s_0}$. Therefore, the probability of regime 1 (an incident before completing the trip) is given by:

$$p_1(t) = p \cdot \frac{R(t)}{s_0} \tag{3}$$

This reflects that an incident must have happened before $t_{start} + \frac{R(t)}{s_0}$ to affect the traveler departing at *t*. If an incident happens, travel time corresponds to the queue length at time *t* divided by the post-incident capacity level s_1 and is given by equation 4. Cumulative arrivals consist of the cars served by the bottleneck before the time when the incident occurs (t_i) at capacity s_0 and the cars served after the incident moment at capacity s_1. The equation can be shown to hold for both incidents that occur before *t* as well as incidents that occur between *t* and $+\frac{R(t)}{s_0}$.

$$T_1(t, t_i) = \frac{R[t] - (t_i - t_{start}) \cdot s_0 - (t - t_i) \cdot s_1}{s_1} \quad (4)$$

To derive the expected travel time we integrate over all possible incident moments. The expected travel time is the integral of the day-to-day travel time function over all possible incident moments between t_{start} and $t_{start} + \frac{R(t)}{s_0}$. Conditional on departing at t and being in regime 1, incident moments are uniformly distributed during this time interval, implying a density of $\frac{s_0}{R(t)}$. The expected travel time for this regime is given by Equation 5.

$$ET_1(t) = \int_{t_{start}}^{t_{start} + \frac{R(t)}{s_0}} T_1(t, t_i) \cdot \frac{s_0}{R(t)} \, dt_i \quad (5)$$

3.3 Expected travel costs

The expected travel costs consist of expected travel time costs, scheduling costs and the time dependent toll. In the model we define three departure ranges being characterized by different costs functions. The first range indicates the range for which the individuals are early for sure, even if there is an incident. Even if an incident happens, they do not incur any schedule delay late. We denote t_e as the departure time of the last traveler in this departure range. In the second departure range individuals depart before the moment t_c and they might face either schedule delay early or late (depending whether and when an incident has occurred). Finally, in the third departure range (starting at time t_c), individuals will face schedule delay late in any case, even if no incident takes place. In each of the departure ranges, individuals can be in either in regime 0 or 1.

Departure range 1: $t_{start} \leq t \leq t_e$

The individual can be in one of the two regimes describes in section 3.2. The costs for individuals leaving before t_e are given by Equation 7.

$$EC_1(t) = p_0(t) \cdot \left(\alpha \cdot T_0(t) + \beta \cdot \left(PAT - t - T_0(t)\right)\right) + \ p_1 \cdot \left(\alpha \cdot ET_1(t) + \beta \cdot \int_{t_{tstart}}^{t_{start}+\frac{R(t)}{s_0}} \left(PAT - t - T_1(t, t_i)\right) \cdot \frac{s_0}{R(t)} dt_i\right) \quad (7)$$

For this departure range travelers only face scheduling costs for arriving early. The expected costs for the first regime are therefore given by integrating over the possible incident moments.

Departure range 2: $t_e \leq t \leq t_c$

Within this departure range, road users always face schedule delay early for regime 0. If an incident has occurred, both schedule delay early and late are possible, depending on the incident moment. Therefore we need to solve for the incident moment at which an individual departing at t does not face any schedule delay. This is the case when the driver arrives exactly at the preferred arrival time (PAT). This moment is denoted by $\breve{t}_i(t)$. The travel time this person encounters has been given in equation 4 (travel time in regime 1). Given departure time t and PAT, $\breve{t}_i(t)$ is therefore given in equation (8b), which results from solving the expression in (8a) with respect to $\breve{t}_i(t)$.

$$t + \frac{R[t] - (\breve{t}_i(t) - t_{start}) \cdot s_0 - (t - \breve{t}_i(t)) \cdot s_1}{s_1} = PAT \quad (8a)$$

$$\breve{t}_i(t) = \frac{R(t)-PAT \cdot s_1 + s_0 \cdot t_{tstart}}{s_0 - s_1} \tag{8b}$$

If the incident occurs later than $\breve{t}_i$ there will be schedule delay early, and if the incident occurs earlier than $\breve{t}_i$, there will be schedule delay late.

Suppose there is an individual who departs early and will arrive early if there is no incident. This individual will arrive on time if the incident occurs at $\breve{t}_i$. This implies that there is a possible benefit in terms of scheduling costs for this traveler because of the incident. However, because we assumed $\beta < \alpha$ the incident will always result in higher overall travel costs.

The expected costs for this departure range are then given by:

$$EC_2(t) =$$

$$p_0(t) \cdot (\alpha \cdot T_0(t) + \beta \cdot (PAT - t - T_0(t))) + p_1(t) \cdot$$

$$\left(\begin{array}{c} \alpha \cdot ET_1(t) + \beta \int_{\breve{t}_i}^{t_{start} + \frac{R(t)}{s_0}} \frac{PAT - t - T_1(t,t_i)}{\left(t_{start} + \frac{R(t)}{s_0} - \breve{t}_i\right)} dt_i + \\ \gamma \cdot \int_{t_{start}}^{\breve{t}_i} \frac{T_1(t,t_i) - (PAT - t)}{\breve{t}_i - t_{start}} dt_i \end{array} \right) + \tag{9}$$

<u>Departure range 3:</u> $t_c \leq t \leq t_{end}$

If an individual departs after t_c, he will in any case face schedule delay late for all departure ranges, even if no capacity reduction occurs. The cost function for this departure range is given by equation 10.

$$EC_3(t) = p_0(t) \cdot (\alpha \cdot T_0(t) + \gamma \cdot (T_0(t) - (PAT - t))) \; + \; p_1(t) \cdot \left(\alpha \cdot ET_1(t) + \gamma \cdot \int_{tstart}^{t_{start} + \frac{R(t)}{s_0}} (T_1(t, t_i) - (PAT - t)). \frac{s_0}{R(t)} dt_i \right) \tag{10}$$

3.5 No-toll Equilibrium Conditions

In this section we describe how the model can be solved for equilibrium given specific values of the value of time, value of schedule delays, incident probability and pre- as well as post-incident capacities. In equilibrium the expected costs must be equal for all drivers, so there is no incentive to change departure time. We use the cost functions derived in equations (7), (9) and (10). For all three departure ranges, we are able to find closed form expressions. We can therefore find the equilibrium using an analytical approach with the following conditions:

We try to find that peak starting time t_{start} that minimizes costs. Given a specific peak starting time, the following calculations are done:

1. *Expected Costs for each Departure Range*

Expected costs must be equal for all drivers. Since the costs of the first driver are deterministic and equal to $\beta * (PAT - t_{start})$, the problem reduces to maximize t_{start} subject to equality of costs for all drivers.

2. Cumulative Departures

The expected costs of each departure range are set equal to the costs of the first driver. Subsequently, we can solve for cumulative departures in each range. The kinks (thus, the changes from one into the next range) occur where cumulative departures are equal.

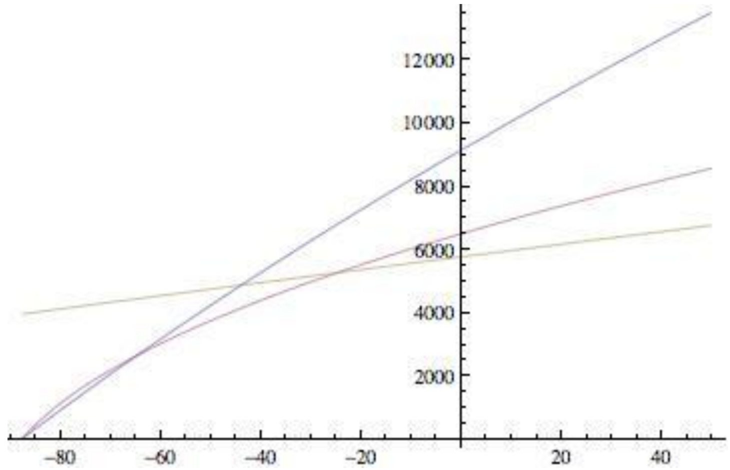
3. Expected Travel time of the last driver

In the case that no incident occurs over the entire peak, $(t_{end} - t_{start}) * s_0$ is be equal to the number of drivers *N*. This equivalent to stating that also if no incident happens during the entire peak, equilibrium departure rates are sufficiently high to make the bottleneck operate at its maximum capacity throughout the peak. Therefore, in the case of no incident during the entire peak, the last driver does not face queuing time. Since the probability of an incident is assumed to be larger than 0, the expected queuing time is positive as well. This also corresponds to the findings by Li et al., 2008.

Thus, all drivers must have departed at t_{end}. This condition, together with t_{end} being equal to $t_{start} + \frac{N}{s_0}$, enables us to solve for t_{start} and t_{end}.

An example of how cumulative departures look like is shown below. The cumulative departures are given by the lower envelope of the three curves. Compared to the deterministic bottleneck model, which only exhibits one kink, the slope is smoother in the stochastic model.

Figure 1: Example Cumulative Departures



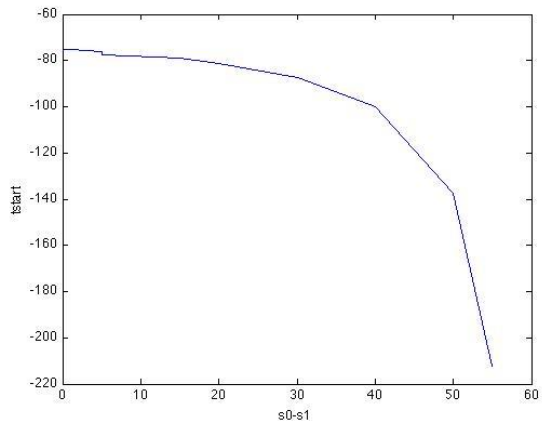
4.1 No-toll Results

This section presents some results on equilibrium departure patterns. We use the equilibrium conditions described above to find the equilibrium departure time and departure pattern. Since the expected costs are required to be equal for all drivers, they need to be equal to the costs of the first drivers. These are given by $\beta * (PAT - t_{start})$. Thus an earlier t_{start} is equivalent to higher costs for all drivers.

1) *Comparing different relative capacity decreases (s_0-s_1)*

The graphs below show a situation where the pre-incident capacity s_0 is equal to 60, and s_1 assumes values between 1 and 59. The difference between s_0 and s_1 is plotted on the x axis. As expected, we find that a higher capacity drop leads to an earlier begin of the peak and hence higher costs for all drivers. Also, it can be shown that the second departure range is longer in this case indicating that the phase when drivers do not know whether they will arrive early or late is extended. This intuitively makes sense since day-to-day travel times exhibit a higher variability if the capacity drop in case of an incident is larger.

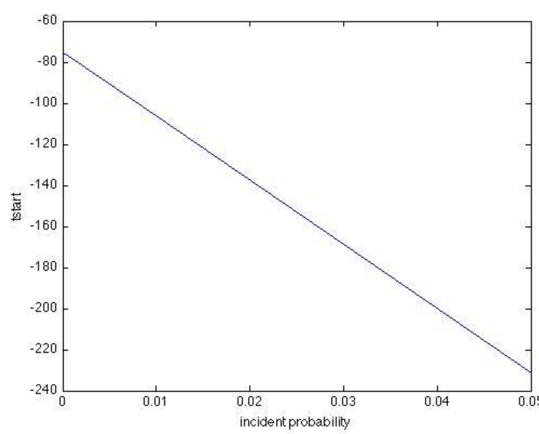
Figure 2: The effect of capacity reduction on (expected) costs/driver



2) *Comparing different incident probabilities*

A higher incident probability leads to similar effect as an increase in the capacity drop in case of an incident. Higher incident probabilities lead to an earlier start of the peak, a longer second departure range and higher costs for all drivers. As figure 3 shows, there is a linear relationship between t_{start} (and hence also the costs) and the incident probability.

Figure 3: The effect of probabilities on (expected) costs/driver



5.1 Conclusions and Extensions

The framework developed in this paper is useful for analyzing the effects of incidents and offers a good basis for future work. It clearly shows that higher incident probabilities and higher capacity drops in case of incidents lead to higher costs for drivers.

In future work, an optimal pricing regime shall be derived for the stochastic bottleneck. As stated also by Fosgerau (2010), this results in quite a complex mathematical problem.

Extensions other than tolling are also feasible. For example, the incident probability can be modelled such that it depends on the number of cumulative departures or the queue length, meaning that p is dependent on the number of drivers and therefore is not constant over the peak. Another extension could be to include the incident duration as a variable in the model. Also various information regimes (where drivers are informed on the incident occurrence and the incident moment) can be analyzed.

References

Arnott, R., Palma, A. de, and Lindsey, R., 1993. A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand, American Economic Review 83, 161–179.

Arnott, R., de Palma, L., Lindsey. R., 1996. Information and Usage of Free-Access Congestible Facilities with Stochastic Capacity and Demand. International Economic Review 37, 181-203.

Arnott, R., de Palma, L., 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., 1995. Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues, Econometrica 63, 327-370.

Fosgerau, M., Jensen, H., 2008. A Note on the Existence of Nash Equilibrium in the Stochastic Bottleneck model. Munich Personal RePEc Archive.

Fosgerau, M., 2009. Congestion costs in a bottleneck model with stochastic capacity and demand. Munich Personal RePEc Archive.

Fosgerau, M., 2010. Congestion with Incidents. Working Paper. Technical University of Denmark.

Koster, P.R. and Rietveld, P., 2010. Optimizing incident management on the road, Journal of Transport Economics and Policy, *forthcoming*.

Li, H., Bovy, P. , Bliemer, M. 2008. Departure Time Distribution in the Stochastic Bottleneck Model, International Journal of ITS Research 6, 79-86.

Lindsey, R. (1996), “Optimal departure scheduling for the morning rush hour when capacity is uncertain”, in D.A. Hensher and J. King (eds.), Proceedings of the 7th World Conference on Transport Research Vol. 2, Pergamon Press, 1996, 195-211.

Schrage, A., 2006. Traffic Congestion and Accidents. Regensburger Diskussionsbeiträge zur Wirtschaftswissenschaft, Working Paper No. 419.

Verhoef, E.T., 2003. Inside the queue: hypercongestion and road pricing in a continuous time–continuous place model of traffic congestion, Journal of Urban Economics 54, 531-565.

Vickrey, W. S. 1969. Congestion Theory and Transport Investment, The American Economic Review 59, 251-260.