CONGESTION COSTS

Camilla Riff Brems
Niels Buus Kristensen
COWI A/S
Birgitte Sloth
Department of Economics, University of Copenhagen

Congestion costs have gained increased attention in the last couple of years. Among other things it is mentioned among the external costs that should be evaluated according to the European Commission's White Paper on European Transport Policy. This states, that

> *"… the price for using the infrastructure should include both infrastructure costs and external costs, which include costs of accidents, emissions, noise and congestion. …"*

This paper focuses on the effects of congestion on the road network. The paper presents a newly developed method to evaluate external costs of congestion including both increased travel times and increased variation in travel times. The method is applied to three different road segments in Copenhagen representing an urban road, an arterial road, and a motorway. It is demonstrated how the methods and its results can be used to analyse the levels and variation of congestion costs.

The work presented here is part of a research project on Congestion, which is carried out by University of Copenhagen, Technical University of Denmark, the National Road Authorities, the Greater Copenhagen Development Council, the Municipality of Copenhagen and COWI. The project is financed by the Danish Transport Research Council and the Danish Ministry of Transport.

1 CONGESTION

When considering the congestion on a road segment the focus is usually the effects on travel time described by the speed-flow curve indicating how travel speed decreases as flow increases. When the capacity of the road is reached, both the speed and the flow decrease at the same time.

However, the effects of congestion are not limited to this reduction of speed. Among the other effects are
- increased inconvenience for drivers
- increased variation in travel time
- deterrence of trips (change of route or of departure time or abandoned trips)
- environmental effects
- effects on accidents

The method suggested in this paper includes in the measure of external costs of congestion the increase in travel time and in variation of travel time, whereas the other effects are omitted.

The *increased inconvenience* for drivers includes that driving in congested roads require more attention to other drivers as well as problems with overtaking or changing lanes. The effect was included in the project through measuring of traffic densities and a small postcard survey, where drivers should state if they felt the traffic inconvenient. However, there were no clear indications of a relationship.

A measure that includes the *deterrence of trips* would require knowledge of the size of deterrence effects. Since these effects include change of route, change of time and abandoned trips, the measurement would require a larger survey area as well as detection of abandoned trips. None of these were included in this present project. A full description of deterrence effects would require an approach with a general equilibrium model that integrates a traffic model for the entire transport system with an economic model to take proper account of substitution between different hours, routes, modes as well as substitution between transportation and other consumption goods or production inputs. Since deterrence of trips under some conditions, i.e. when traffic is close to the capacity of the road, may be a very important effect of increasing traffic, it is important also to try to develop measures taking this into account, but in this paper we will not do so.

The *environmental effects* of congestion in terms of air pollution can be both positive and negative. If the effects of congestion are merely a limited reduction of speed, the effects are positive on motorways and some arterial roads, since the cars will be driving at a more efficient speed. On urban roads the effects would be negative, because the speed is already below the most efficient speed. If congestion increases frequencies of accelerating and decelerating, e.g. by creating stop-go traffic, this will have negative environmental effects. Similarly, the effects of congestion on *accidents* may be positive or negative depending upon exactly how the congestion changes the traffic speed and density. Rough estimates of both environmental and accidental costs in the project indicated that both are small compared to the costs of effects to travel speed, variation, and deterrence of trips.

The congestion effects as described in the traditional speed flow curve has been dealt with in e.g. Hall et al. (1986), Newbery (1990), Verhoef (1999) and May et al. (2000). With the exception of Newbery (1990), these references focus on the traffic effects and not on the measuring of the costs of congestion.

2 METHODOLOGY

No road user expects to have the same travel time in peak hours as in the middle of the night. A road user tends to include some additional time in his or her estimation of travel time given the present level of traffic. This additional time increases if the road user has a preferred arrival time, e.g. at work, and will make sure not to be late.

In section 2.1 the distinctions between travel time and delay, early and late arrival and other relevant terms are described. Based on this, measures of travel time, optimal planned travel time as well as total and marginal cost of congestion are given in section 2.2.

2.1 Terminology

As mentioned, the road users' nuisance of congestion is not only increased travel times but also increased variation in travel times. The size and variation of travel time is of most importance to those road users, who have a preferred arrival time. They need to depart earlier in order to be sure to arrive on time. When a road user has to decide how much earlier he or she will depart from home the evaluation of travel time, of being early and of being late goes into the considerations.

In planning the departure time, the road user is - maybe unconsciously - making an evaluation of the distribution of travel time and the costs associated with being either early or late. Usually, it is worse being late than early. In the following this is described by the value of time that the road users assign to travel time and the penalties they assign to early and late arrival. Road users without a preferred arrival time are assumed to have no penalties for early and late arrival. Road users with a preferred arrival time, e.g. going to an important meeting, may have high penalties on late arrival relative to the penalty of early arrival or to the value of travel time. Usually, the penalty assigned to late arrival is 1-3 times[1] the value of travel time. Here a penalty of late arrival, which is twice the value of travel time, is used. This means that the road users would rather be leaving home 15 minutes before necessary than being 5 minutes late. Likewise, the penalty of being early is less than or equal to the value of travel time depending on the possibility to make use of the hidden waiting time associated with early arrival.

In **Fejl! Ukendt argument for parameter.** two observed distributions of travel time at different levels of traffic intensities on the arterial road are shown. As might be expected both the increase in travel time compared to free flow and the variation in travel time are small when traffic intensities are low. Some of the road users observe travel times that are identical to the free flow travel time, and all road users at this level of traffic observed increases not exceeding 15 seconds per km. When traffic intensities increase, both the travel time and the variation of travel time increase.

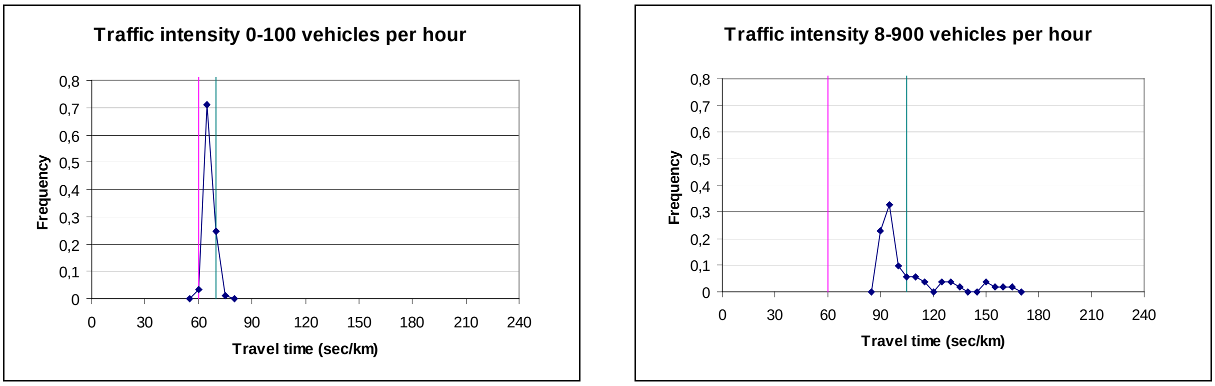

Figure ***Fejl! Ukendt argument for parameter.*** *Observed variation in travel time at different levels of traffic on the arterial road. The first vertical line indicates free flow travel time whereas the second indicates the optimal planned travel time when the penalty of late arrival is valued twice the value of travel time.*

The travel time that minimises the total cost of the trip, i.e. cost of travel time and of early or late arrival is called the *optimal planned travel time*. If the optimal planned travel time is 30 minutes, the road user must depart 30 minutes before the preferred arrival time in order to minimise the costs associated with the trip. With the value of early arrival equal to the value of travel time and the value of late arrival three times as high, the vertical lines most to the right in **Fejl! Ukendt argument for parameter.** shows the optimal planned travel time for these distributions of travel time. For intensities between 800 and 900 vehicles per hour the observed travel time varies from 90 to 170 sec per km. This results in an optimal planned travel time of 105 sec per km. Even though the road users are planning their departure, some may be up to 15 sec early for each km whereas some may be up to 65 sec late for each km. The average travel time and average hidden waiting time is calculated later. Here, the difference between the free flow travel time of 60 sec/km and the optimal planned travel time of 105 sec/km is called expected delay, whereas travel time exceeding 105 sec/km is called unexpected delay.

---

[1] Usually the delay included in value of time studies is delay in the conventional approach, i.e. all delay compared to free flow time. Observed values for Copenhagen are between 1.7 and 3.4. This results in penalties between 1 and 3. These values cannot be transferred to the approach used in this paper, which only values late arrivals at the value. However, a penalty of late arrival of twice the value of travel time is used as an estimate.

Fejl! Ukendt argument for parameter. below shows the relations between the conventional distinction between free flow time and delay and the optimal planned travel time, with expected and unexpected delay.

Travel time		
Free flow	Delay	
Optimal planned travel time		Unexpected delay
Free flow	Expected delay	Unexpected delay

Figure ***Fejl! Ukendt argument for parameter.*** *Splitting travel time in free flow, expected and unexpected delay.*

The purpose of this distinction is that in practice road users do not expect to observe free flow travel times when travelling in peak hours. Depending on the experience of the traffic patterns in the area they tend to build in some delay in their planned travel time. This delay is not causing any late arrivals and is therefore not as big a problem as the unexpected delays, which cause being late.

2.2 Expressions for congestion time and cost

With the above terminology it is now possible to derive total and marginal costs of congestion that includes effects of both size and variation of travel time. The expressions are based on distributions of travel time describing size and variation as well as on valuations for travel time, early arrival and late arrival.

The distribution of travel time is described by a probability density $f(\tau, v)$, which describes the probability of the road user observing travel time τ and traffic flow v. For some purposes the marginal density describing the travel time at a specific traffic flow, v, is used. This is described by the probability density $f_v(\tau)$. The two densities in **Fejl! Ukendt argument for parameter.** are examples of such marginal probabilities.

The value of travel time is denoted β, and is identical to the traditional values of time. The penalty for being early is denoted α, and the penalty for being late is denoted γ.

Optimal planned travel time
Based on these assumptions the cost of a trip with planed travel time t, i.e. departure t before preferred arrival, is expressed by,

$$C(t) = \int_{\underline{t}}^{\infty} \beta \tau f_v(\tau) d\tau + \int_{\underline{t}}^{t} \alpha (t-\tau) f_v(\tau) d\tau + \int_{t}^{\infty} \gamma (\tau - t) f_v(\tau) d\tau \qquad (1)$$

The first term is the costs associated with the travel time, where $\underline{t}$ is the free flow travel time, which is the minimum travel time observed. The second and third terms are the costs of early and late arrival, respectively. In the formula τ is the realised travel time from the distribution of travel times, $f_v(\tau)$ at flow level v. When minimising $C(t)$ with respect to t, the optimal planned travel time, t^*, is obtained. By manipulating formula (1) it can be proven that t^* can be determined by

$$F_v(t^*) = \frac{\gamma}{\alpha + \gamma} \qquad (2)$$

Here F_v is the cumulative distribution of travel times associated with the density $f_v(\tau)$.

The assumption of no deterrence corresponds to the assumptions that the amount of trips is constant and that $f_v(\tau)$ is constant within the analysed period of time. These assumptions may be lessened in an approach, which includes deterrence.

With the penalty of late arrival as twice the penalty of being early (rather up to 10 minutes early than 5 minutes late), formula (2) gives that the optimal planned travel time is identical to the 67%-fraction of the distribution of travel time. This corresponds to the vertical lines most to the right in the two distributions in **Fejl! Ukendt argument for parameter.**.

Travel time and cost of congestion

Above it was described how the travel time for a road user depends on whether he or she has a preferred arrival time or not. The difference in travel time is described in the two formulas (3) and (4) below.

For road users with *no preferred arrival*, i.e. $\alpha = \gamma = 0$ the expected travel time at flow v is given by

$$Travel\ Time = \int_{\underline{t}}^{\infty} \tau f(\tau, v) d\tau \qquad (3)$$

This corresponds to the mean of the distribution, and the minimum possible travel time is the free flow travel time, $\underline{t}$.

The effects of variation in travel time becomes evident, when the same travel time is derived for those road users, who have a preferred arrival time, i.e. $\alpha = \beta$ and $\gamma > \beta$. The relationship between γ and α is used to determine t^*, which describes the averseness of being late. The higher the penalty γ relative to α, the longer the optimal planned travel time t^*.

$$Travel\ Time = t^* + \int_{t^*}^{\infty} (\tau - t^*) f(\tau, v) d\tau \qquad (4)$$

With a preferred arrival time no road user can observe a travel time less than t^*. However, in the evaluation of the trip the travel time exceeding t^* should be valued by γ in addition to the value of travel time.

Total cost

The total travel cost associated with the travel times above depends on the value of travel time, β, and on the penalties that road users assign to early arrival, α, and to late arrival, γ. For simplicity it is assumed that hidden waiting time due to early arrival is valued as travel time, i.e. $\alpha = 0$.

With these assumptions the total Travel Cost (TC) at flow level v is

$$TC = \beta \cdot v \left[t^* + \int_{t^*}^{\infty} \left(1 + \frac{\gamma}{\beta}\right) (\tau - t^*) f(\tau, v) d\tau \right] \qquad (5)$$

However, only part of these costs is related to congestion. The total Congestion Cost (CC) at flow level v is derived as TC subtracted the cost at free flow.

$$CC = \beta \cdot v\left[\left(t^* - \underline{t}\right) + \int_*^\infty \left(1 + \frac{\gamma}{\beta}\right)(\tau - t^*)\, f(\tau, v)\, d\tau\right] \tag{6}$$

CC can be used as part of the evaluation of infrastructure projects, which are expected to reduce congestion, or to provide information on the order of magnitude of congestion problems.

Marginal cost

However, in this paper the focus is on external costs of congestion. The marginal external cost is described by the extra cost which an additional road user's extra km poses on the other road users. The effect on the single road user is not included.

$$\begin{aligned} MCC &= \frac{\partial CC}{\partial v} \\ &= \beta\left[\left(t^* - \underline{t}\right) + \int_*^\infty \left(1 + \frac{\gamma}{\beta}\right)(\tau - t^*)\, f(\tau, v)\, d\tau\right] \\ &+ \beta \cdot v\left[\frac{dt^*}{dv} - \int_*^\infty \left(1 + \frac{\gamma}{\beta}\right)\frac{dt^*}{dv} f(\tau, v)\, d\tau + \int_*^\infty \left(1 + \frac{\gamma}{\beta}\right)(\tau - t^*)\frac{\partial f(\tau, v)}{\partial v}\, d\tau\right] \end{aligned} \tag{7}$$

Formula (7) does not look very applicable. However, with different approximations applied in section 3 it proves possible to get reasonable results.

3 APPLICATION

The methodology described above has been applied to three different roads in Copenhagen. The roads are selected such that they represent different types of roads, and such that the selected roads have some problems of congestion. The three roads are

Motorway:	Motorring 3
Arterial road:	Ellebjergvej/P Knudsens Gade
Urban road:	Vesterbrogade

External costs of congestion in the form of marginal congestion costs are derived for each of the three roads. In addition, sensitivity analyses stating the dependency of the external costs on some of the posed assumptions are carried out.

The derived distributions show that at low traffic volumes the road user has a low travel time with little variation, whereas at higher volumes both the travel time and the variation in travel time increases. A total of 12 different distributions were generated for each of the three different road types (one with four subsections) and two directions.

Formulas (3) and (4) state the average travel times for road users with no preferred arrival and with preferred arrival. **Fejl! Ukendt argument for parameter.** shows these travel times for the inbound direction on Ellebjergvej/P Knudsens Gade.

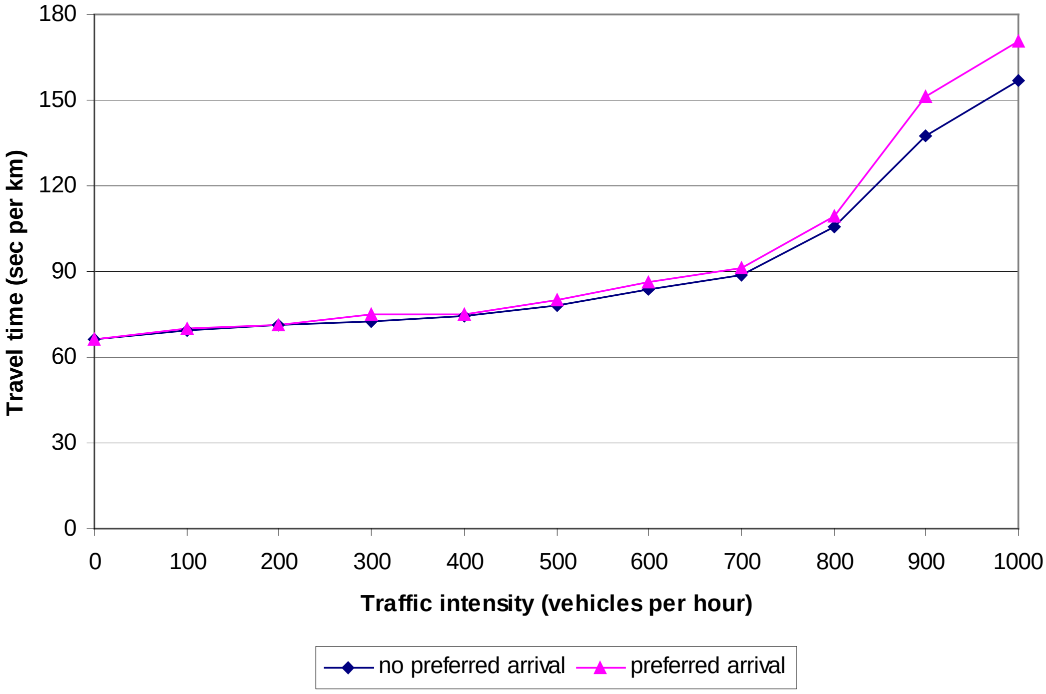

*Figure **Fejl! Ukendt argument for parameter.** Difference in average travel time for preferred and no preferred arrival for Ellebjergvej/P Knudsens Gade.*

The travel time for road users with no preferred arrival time is smaller than the travel time for road users with preferred arrival time. The difference is the hidden waiting time due to early arrival. For low traffic intensities the two sets of travel times are almost identical. This is the case, when there is not much variation in travel time. However, when the traffic intensity gets close to capacity, the variation increases and so does the difference in the two types of travel time.

Including preferred arrival time will influence the level of the marginal costs of congestion. The same applies for the value of late arrival compared to early arrival and for the value of time applied. The importance of all three effects is outlined below.

*Table **Fejl! Ukendt argument for parameter.** Effects of congestion on Ellebjergvej/P Knudsens Gade.*

	Free flow traffic (0-100 veh/h) $\gamma = 0$	Congested traffic (8-900 veh/h) $\gamma = 0$	Congested traffic (8-900 veh/h) $\gamma = 2\beta$
Average travel time	66 sec/km	106 sec/km	110 sec/km
Hidden waiting time	0 sec/km	0 sec/km	4 sec/km
Marginal congestion time	2 sec/km	128 sec/km	135 sec/km
Marginal congestion cost	0.05 kr./km	3.10 kr./km	6.25 kr./km

The Hidden waiting time is included in the average travel time.

In **Fejl! Ukendt argument for parameter.** the first two columns illustrate the effects of congestion if only the *increase in travel time* is included in the congestion costs.

When traffic increases from a free flow traffic situation with 0-100 vehicles per hour to a congested traffic situation with 8-900 vehicles per hour the average travel time increases from 66 sec/km to 106 sec/km. This corresponds to a reduction in speed to less than 2/3 of the free flow speed. The consequence regarding marginal time is an increase from 2 sec/km to 128 sec/km. This can be transferred into congestion costs by a value of travel time. In this case a value of 87.70 kr. per hour is used. This is derived from the average mix of traffic on purposes in Copenhagen and the purpose specific values of time recommended by the Danish National Road Authorities. The marginal costs of congestion then becomes less than 0.05 kr./km in the free flow traffic situation. This increases to 3.10 kr./km in the congested traffic situation.

By comparing the first and the last column in **Fejl! Ukendt argument for parameter.** the effects of both *increased travel time and increased variation in travel time* are illustrated.

The effect on congestion costs is here an increase from less than 0.05 kr./km in the free flow traffic situation to 6.25 kr./km in the congested traffic situation. **Fejl! Ukendt argument for parameter.** also shows the decomposition of these congestion costs into a number of effects.

The inclusion of a preferred arrival time and a penalty of being late, which is twice the value of travel time, results in an increase of average travel time to 110 sec/km. The 106 sec/km is still due to the direct increase in amount of traffic whereas the remaining 4 sec/km is hidden waiting time due to the averseness of being late. Both the average travel time and the hidden waiting time increases with level of traffic. Here, the marginal congestion time becomes 135 sec/km in the congested traffic situation. With the same value of travel time as above and a penalty of late arrival, which is twice the value of travel time, this results in a marginal cost of congestion of 6.25 kr./km.

In deriving these congestions costs assumptions regarding valuation of late arrival and value of time have been made. Sensitivity analyses have been carried out to analyse the importance of these assumptions.

First, different values of late arrival have been tested. From formula (2) it shows, that the value of late arrival compared to the value of early arrival or travel time determines the fraction of the cumulative distribution. Changing the value of late arrival from two to three times the value of travel time will change the fraction from 67% to 75%. For the marginal congestion costs this means no change for the free flow traffic situation and an increase from 6.25 kr./km to 7.20 kr./km in the congested situation. This is an increase of 15%. The valuation of late arrival proves to have some influence on the congestion costs, especially when the traffic intensities get close to capacity.

Second, the applied value of time has effects on marginal cost. From formula (7) it is evident that the value of time is just multiplied to the marginal congestion time. A 20% increase in value of time from 87.70 kr. per hour to 105.25 kr. per hour will therefore result in a 20 % increase in marginal congestion costs, i.e. from 6.25 kr./km to 7.50 kr./km.

Finally, **Fejl! Ukendt argument for parameter.** lists the levels of marginal costs of congestion for each of the three road types at different times of day, which have different levels of traffic intensities. All the costs are based on a penalty of being late, which is twice the value of travel time.

Table ***Fejl! Ukendt argument for parameter.*** *Congestion costs at different times of day (kr. per km).*

	Vesterbrogade	Ellebjergvej/P Knudsens Gade	Motorring 3
Peak hour - peak direction	3.70	11.15	2.25
Peak hour - opposite direction	3.70	1.50-6.00	2.25
Mid day	3.70	1.50	0-0.75
Night	0	0	0

The results in **Fejl! Ukendt argument for parameter.** show significant differences in marginal costs of congestion both for different road types and for different times of day. Ellebjergvej/P Knudsens Gade has the highest marginal costs of congestion in the peak hours, with Vesterbrogade on the second place. Motorring 3 with the highest capacity has the lowest marginal costs for an additional km driven.

It is seen that Vesterbrogade has the same level of marginal costs throughout the entire day. This is in line with the observations of traffic on the road. Between the peak periods there is a lot of shopping traffic on this road. Both Ellebjergvej/P Knudsens Gade and Motorring 3 have significant differences in the levels of marginal costs between peak periods and mid day. Motorring 3 has the same marginal costs in both directions in the peak period, since it is a ring motorway with an almost 50/50 split in directions. For Ellebjergvej/P Knudsens Gade with a clear split in directions there is difference in the levels for peak direction and opposite peak direction.

In all, the obtained marginal costs are within what could be expected, when regarding the traffic patterns on the analysed roads.

4 PERSPECTIVES

The methodology and applied results reported in this paper provides a way to measure the external costs of congestion including the costs of both increased travel time and increased variation of travel time as well as it indicates levels and sensitivity of the external costs.

In peak hours the cost of congestion is the most important external cost when compared to costs of emissions, noise and accidents. It is therefore important to include these costs in the infrastructure pricing. **Fejl! Ukendt argument for parameter.** showed that there is great variation in the costs of congestion during the day as there is for the amount of traffic.

Behind the above applications of methodology there are extensive and expensive data collections, which cannot be applied to all roads if a road pricing system should be applied generally. In the continuation of the project, it is therefore a goal to try to generalise the distributions of travel time generated in this part of the project.

Other affiliated projects are going to clarify some of the uncertainties related to the results presented. The Danish part of the EU project Progress, called AKTA, has questionnaires trying to reveal the share of road user with preferred arrival as well as the valuation of late arrival compared to early arrival. Also in AKTA, some field tests with road pricing have been carried out. By incident, one of the tested price levels is in the same level as the marginal costs derived in this project.

5 BIBLIOGRAPHY

Arnott, R., de Palma, A. and Lidsey, R. (1990)
"Economics of a Bottleneck"
Journal of Urban Economics, **27** 111-130.

Hall, F.L., Allen, B.L. and Gunter, M.A. (1986)
"Empirical Analysis of Freeway Flow-Density Relationships"
Transportation Research, **20** 197-210.

May, A.D., Sheperd, S.P. and Bates, J.J. (2000)
"Supply Curves for Urban Road Networks"
Journal of Transport Economics, **34** 261-290.

McFadden, D. (1974)
"The Measurement of Urban Travel Demand"
Journal of Public Economics, **3** 303-328.

Newbery, D.M. (1990)
"Pricing and Congestion: Economic Principles relevant to pricing Roads"
Oxford Review of Economic Policy, **6** 22-38.

Small, K.A. (1982)
"The Scheduling of Consumer Activities: Work Trips"
American Economic Review, **72** 467-479.

Verhoef, E.T. (1999)
"Time, speed, flows and densities in static models of road traffic congestion and congestion pricing"
Regional Science and Urban Economics, **29** 341-369.