AN INTEGRATED DYNAMIC MODEL OF ROAD TRAFFIC CONGESTION BASED ON SIMPLE CAR-FOLLOWING THEORY

Erik Verhoef
Department of Spatial Economics
Free University Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
Phone: +31-20-4446094
Fax: +31-20-4446004
E-mail: everhoef@econ.vu.nl
html://www.econ.vu.nl/vakgroep/re/members/everhoef/et.html

Key words: Road traffic congestion, congestion pricing, hypercongestion, car-following theory
JEL codes: R41, R48, D62

Abstract

This paper presents a dynamic model of road traffic congestion based on simple car-following theory, allowing for finite group velocity and discrete vehicles. The model offers a full-fledged dynamic version of the standard static model of road traffic congestion based on the so-called 'fundamental diagram', depicting the relation between density and speed. Using this dynamic formulation, it is proven formally and illustrated numerically that the suggested hypercongested equilibria of the standard static model are in fact dynamically unstable. The model presented is fully consistent with the standard static model, and reproduces non-hypercongested stationary state outcomes found in that model for arrival rates of new users below the maximum capacity of the road. When this maximum capacity is exceeded, the model behaves in the same manner as Vickrey's dynamic model of bottleneck congestion. Therefore, the model presented here offers an integration and a generalization of these two important types of models.

1. Introduction

The standard static model of road traffic congestion, based on the so-called 'fundamental diagram of road traffic congestion' depicting the inverse relation between density and speed on a road segment, has been subject to extensive debate in the literature – see, for instance, Else (1981, 1982) versus Nash (1982); De Meza and Gould (1987) versus Alan Evans (1992); Andrew Evans (1992, 1993) versus Hills (1993); and more recently, discussions by Chu and Small (1996), Yang and Huang (1998) and Verhoef (1998). In these studies, primarily stationary state solutions to the model are considered. The present paper aims to develop a full-fledged dynamic model of road traffic congestion, which in stationary states yields equilibria consistent with those found with the standard static model. With this proposed model however, based on simple car-following theory, it is possible to study the dynamic behaviour of the system also during transitional phases. This will prove to offer more than just a minor extension of the static model. Instead, the specification turns out to result in an integrated model of road traffic congestion, which is capable of producing both stationary state equilibria with flow congestion as described by the standard static model, and equilibria with bottleneck congestion such as found in Vickrey's (1969) model. Furthermore, the dynamic specification allows one to study the dynamic stability of the equilibria suggested by the standard static model. These analyses show that equilibria associated with hypercongestion are in fact dynamically unstable.

The paper has the following structure. In the next section, the standard static model of road traffic congestion will be briefly reviewed, and the simple transformations that yields the dynamic model are discussed. Section 3 presents the simulation model. In Section 4, a number of general propositions based on the new formulation are derived, and illustrative numerical illustrations are provided. Section 5 considers the implications for models where the demand side of the market is modelled more elaborately. Section 6 concludes.

2. Reformulating the standard static model on the basis of some simple dynamic car-following theory

The standard static model of flow congestion on a single-lane/single-entry/single-exit/constant-capacity road is deceptively simple. It is based on the so-called 'fundamental diagram' of road traffic congestion, depicting how speed S (in meters per second in the sequel) decreases with an increasing density D (vehicles per meter) on the road. For the density-speed relation (DS-curve) in the first quadrant in Figure 1[1], it is assumed that a certain maximum free-flow speed S^* can be sustained for positive densities (the DS-curve starts with a flat segment); and that there is some maximum density D_{max} for which speed falls to zero. Because traffic flow F (vehicles per second) is equal to the product of D and S for the single lane model, F obtains a maximum for some combination of speed and density. This combination is denoted $S^\#$ and $D^\#$.

[1] Figures 1 and 2 are taken directly from Verhoef (1998).

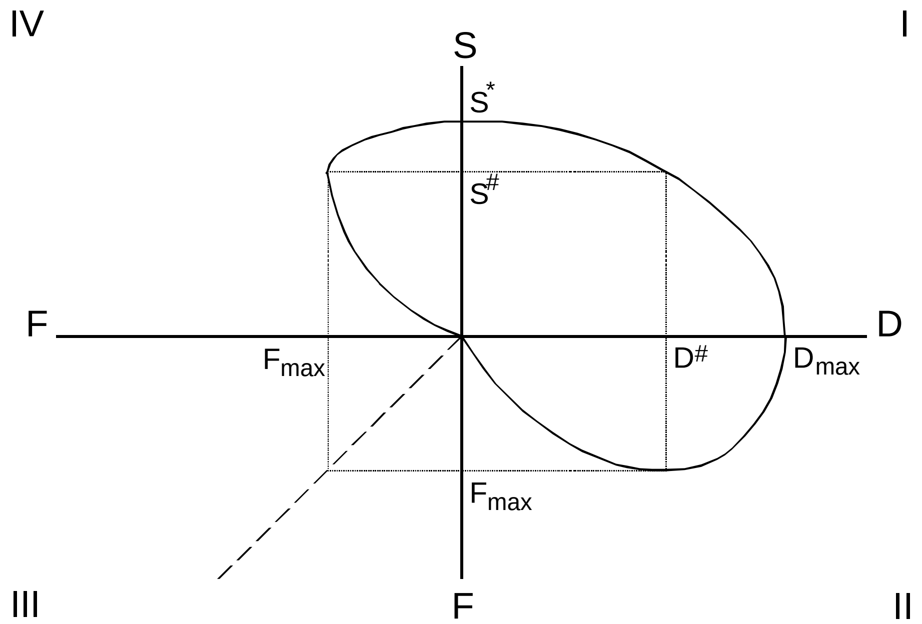

Figure 1. The density-speed curve (I), the speed-flow curve (IV) and the density-flow curve (II)

This gives rise to the familiar backward-bending speed-flow curve (SF-curve) in the fourth quadrant of Figure 1, and density-flow curve (DF-curve) in the second quadrant of Figure 1. Next, under the assumption that only time costs matter for generalized user costs, the speed-flow curve in Figure 1-IV can subsequently be combined with the inverse relation between speed and travel times to obtain the standard backward-bending average social cost function depicted in Figure 2. The lower section of this AC-curve, where speeds are relatively high and travel times relatively short, corresponds with the upper section of the SF-curve in Figure 1-IV. Likewise, the upper section of the AC-curve, representing situations that are usually referred to as 'hypercongestion', corresponds with the lower section of the SF-curve. As speeds go to zero in Figure 1-IV, generalized user costs go to infinity in Figure 2. Therefore, each level of flow, except the maximum level and zero flow, appears to be obtainable at two cost levels: a low one (denoted with superscript l), where the density is relatively low and the speed relatively high; and a high one (denoted with superscript h), where the opposite holds.

The backward-bending cost curve in Figure 2 has led to heated debate in the literature, because the confrontation with a demand curve such as E may lead to puzzling and sometimes counterintuitive results. These include the possibility of multiple equilibria (denoted x, y and z in the diagram), and the fact that in the decentralized optimum with tolling, the traffic flow may actually exceed the non-intervention flow (note that no marginal costs are depicted in Figure 2 in order to keep the diagram decipherable). Space is lacking here to repeat the arguments in detail (see, for instance, Else, 1981, 1982; Nash, 1982; De Meza and Gould, 1987; Alan Evans, 1992; Andrew Evans, 1992, 1993; Hills, 1993; Chu and Small, 1996; Yang and Huang, 1998; and Verhoef, 1998).

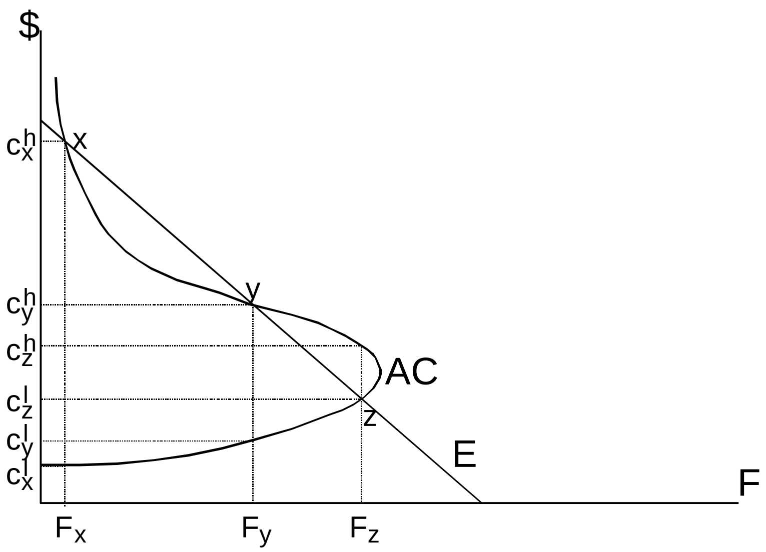

Figure 2. The backward-bending average cost curve (AC)

In Verhoef (1998) it was suggested that the backward-bending segment of the AC-curve is dynamically unstable. However, in order to make that statement rigorously, one actually needs a full-fledged dynamic model, that is consistent with the stationary state equilibria implied by Figures 1 and 2, but that is also capable of describing the situation on the road outside stationary states. In this respect, it can be observed that the fundamental diagram in Figure 1-I can only imply a certain speed if the density is unambiguously defined. This means that the density on a certain point along the road should have the same value upstream as downstream. Since this property should hold for every point along the road, the logical consequence is that density should be constant along the road. Moreover, if density would vary along the road, even a local measure for density is not unambiguously defined, because it then depends on the road stretch considered when determining density. Since the fundamental diagram therefore presupposes a constant density along the road, it implies equal speeds along the road, and hence reproduces constant densities for subsequent instants. Hence the inherently static nature of the standard model.

In order to develop a dynamic model consistent with the stationary states represented in Figures 1 and 2, two variables used above will be manipulated in a very basic manner. First, it is used that in stationary states, the density D (cars per meter) is equal to the inverse of the distance between subsequent cars, δ (meters between cars): $D=1/\delta$. Hence, we can directly manipulate the DS-curve in Figure 1-I into an equivalent distance-speed (δS-curve), which should produce exactly the same stationary state equilibria as the original DS-curve. The equivalent DS and δS-curves are sketched in Figures 3-I and 3-II, respectively. The maximum density D_{max} is equal to $1/\delta_{min}$, where δ_{min} gives the minimum distance between cars, for which speeds fall to zero. Since the distance δ should be measured from front-to-front (or back-to-back) of subsequent cars, δ_{min} will be slightly more than the length of a car. Next, $\delta^{\#}=1/D^{\#}$ gives the distance between cars for which the equilibrium flow is at its maximum. Finally, the limit of D=0 corresponds with $\delta\rightarrow\infty$. Since S^* can be sustained for positive densities, however,

S^* can apply for stationary states with non-infinite values of δ. It is assumed throughout the paper that the function S(δ) is continuous for all δ and smooth for all $\delta > \delta_{min}$.

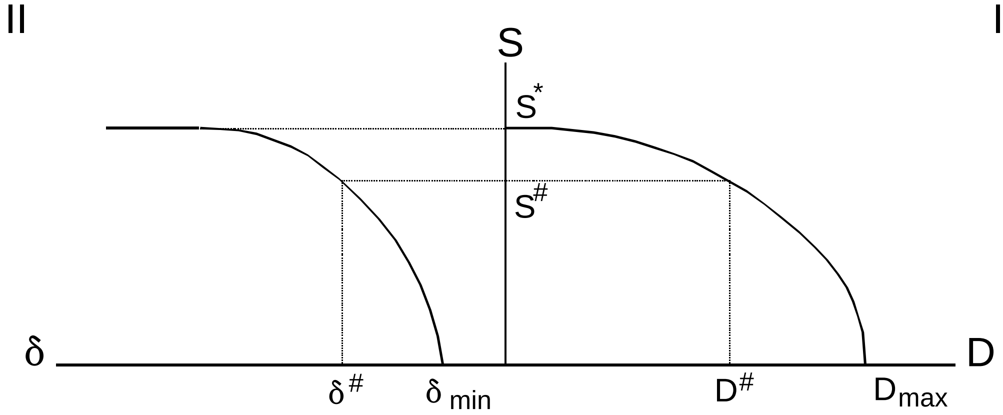

Figure 3. The density-speed curve (I) and the equivalent distance-speed function (II)

The second basic transformation to be made first involves the observation that for stationary states, the arrival rate of new vehicles at the entrance of the road ρ (cars per second) should be equal to the equilibrium flow F on the road. We can next use the identity that the time span τ between the arrival of two vehicles at the entrance of the road (seconds between cars) is the inverse of ρ: $\tau = 1/\rho$. It should be noted that this requires a fully deterministic arrival of new vehicles at the entrance of the road, but the same assumption of course must underlie the deterministic relations depicted in Figures 1 and 2. A formulation with stochastic arrivals is certainly not considered as irrelevant, but rather as a topic for future research after the behaviour of the model with deterministic arrivals has been explored in the present paper.

So far, we have only manipulated variables in the standard model, so we literally still have exactly the same model. Nevertheless, the stage is already nearly set to spell out the dynamic model that is consistent with the static one outlined above for stationary states. An important assumption that is made first is that when a driver chooses a speed according to δS-curve in Figure 3-II, δ denotes the distance to his predecessor.[2] In other words, knowing that the driver(s) upstream – behind him – behave in exactly the same manner as he does himself, a driver can safely ignore them in cases that the distance to his predecessor and to his follower differ. This assumption is consistent with the reasonable assumptions that (1) a driver can and will always speed up when the density behind him is relatively high, but the density in front of him allows such speeding up; and (2) a driver following a relatively slow platoon (which he cannot overtake because of the single-lane assumption) will not change his speed if and because faster cars, at a relatively large distance, are approaching him from behind.

Now starting from any initial situation on the road, and given an arrival rate of new users ρ possibly inconsistent with this initial state, the above manipulations and additional assumption imply a simple dynamic car-following model that should allow us to study also

[2] Strictly speaking, the δS-curve would give the *maximum* speed a driver can choose for a given distance to his predecessor. However, Proposition 1 in Section 4.1 asserts that this maximum speed coincides with the actual speed that a travel time minimizing user chooses.

non-stationary processes. The preposition 'simple' is not merely qualitative. It also reflects that the standard classical car-following model is defined using a second-order differential equation, where the acceleration or deceleration of a driver i depends on the difference in speed with his predecessor (see for instance Zhang and Jarret, 1997). Hence, denoting the location along the road as x, with 0≤x≤X (x=0 denotes the entrance of the road which is X meters long, so that the exit is at x=X), the classical model uses a formulation of the type:

$$\ddot{x}_i = f(\dot{x}_{i-1} - \dot{x}_i) \quad (1)$$

(ignoring terms related to reaction time for convenience). In the current model, however, a first-order differential equation is used, where the speed of a driver depends on the distance with his predecessor. Hence, the present model uses a formulation of the following type:

$$\dot{x}_i = f(x_{i-1} - x_i) \quad (2)$$

where f gives a distance-speed function such as shown in Figure 3-II. This simple first-order structure in the first place follows directly from the static model that we wish to make dynamic. Next, it can be noted that also with (2) as the fundamental car-following equation, the acceleration or deceleration of a driver depends on the difference in speed between him and his predecessor. For instance, if a driver is driving slower than his predecessor, the distance between them is increasing, and therefore this driver must be accelerating. Hence, (2) is certainly not inconsistent with the type of behaviour assumed in (1), be it that with (2), the acceleration for a given speed difference may be different for different absolute speeds of the two drivers considered. This is, however, not necessarily an unrealistic assumption.

The dynamic model therefore hinges on the assumptions that the δS-function in Figure 3-II also applies during transitional phases, and that it is only the distance to the predecessor, not to the follower, that matters in the selection of the speed. However, two more assumptions have to be made before the model is fully determined. The first one has to do with the entrance of the road. It is assumed that queuing is possible when upon arrival of a driver, his predecessor has not yet travelled δ_{min} meters. The queuing process used is simple and, as in Vickrey's (1969) model, in itself congestion-free: if the queue involves multiple users, they maintain their order of arrival during their queuing time, and a driver will start driving at a positive speed and pass the actual entrance of the road as soon as his predecessor has travelled (marginally farther than) δ_{min} meters.

The second assumption involves the behaviour near the exit of the road. Clearly, when his predecessor has completed his trip, we can no longer determine the speed of a driver according to (2). The possible assumption that a driver would then immediately speed up to S^* would in the first place be inconsistent with real-world situations where predecessors do not vanish into thin air when leaving the road, but would often take the same exit-link – which is outside the present one-link model. Secondly, because one of the conclusions below will be that the hypercongested equilibria depicted in Figure 2 are dynamically unstable and will therefore normally not result as stationary state equilibria, it is important that this conclusion should not possibly be dependent on drivers speeding up to S^* near the road's exit. On the other hand, the introduction of some bottleneck at the road's exit is also not an attractive option, since this too may become an unwarranted additional driving force in the model.

Therefore, the rather mild and general assumptions are made that if a driver has not decelerated (accelerated) up to the moment that his predecessor leaves the road, he will also not do so afterwards; and if he has had a constant speed for a positive number of last seconds that his predecessor was still on the road, he will maintain that speed also when his predecessor has left. For the numerical simulation model to be used below, a more specific assumption has to be made. There it is assumed that over the last meters where drivers no longer have a predecessor, they behave as if the road were longer and their predecessor were still there. This assumption technically means that the behaviour of the first drivers has to be calculated as if the road were much longer than X, although only their speeds over the given length X of the road are considered as relevant for their travel times. In practice it implies that the speed as a function of location x will be smooth curves for all drivers, without any kinks near the exit. It is thus implicitly assumed throughout that X is 'sufficiently large': X is much larger than the values of δ considered, so that the 'predecessor-less' part of the trip near the exit is always relatively small and does not become a major driving force in the model.

This concludes the general discussion of the model. In its present form, the model is of course primarily concerned with the cost side of the market. This follows from the aim of studying the dynamic stability of equilibria suggested by the standard static model, which will lead to a rejection of the shape of the standard AC-curve depicted in Figure 2. The demand side will be largely neglected in the next sections: only exogenously determined arrival rates, consistent with inelastic demand, will be considered. In Section 5, however, the implications for models with more elaborate demand structures will be discussed.

Finally, some distinguishing features of the present specification and set of assumptions, compared to other dynamic models of flow congestion, can be mentioned. First, the present model does not have the unrealistic properties of either 'zero group velocity' as in Henderson (1974, 1981) and Chu (1995), where drivers maintain the same speeds during their entire trip; or 'infinite group velocity' as in Agnew (1973), where all drivers simultaneously present on the road on a certain instant all have the same speed regardless of their position. Still, the finite positive group velocity considered here is no novelty; see the hydro-dynamic models of Lighthill and Whitham (1955) and Newell (1988). In contrast to these models, however, the present model uses a rather realistic specification by treating vehicles as discrete entities instead of considering a homogeneous flow, and by its micro-basis on behavioural assumptions. Like these models, however, the present model has the possible disadvantage that it does not have simple closed-form solutions for variables like traffic flow or speed at a certain point along the road at a certain moment, other than in stationary states.

3. The numerical model

This section briefly outlines the illustrative numerical model to be used below. For this model, the following distance-speed function $S(\delta)$, depicted in Figure 4-I, is used:

$$S(d) = 0 \quad if \quad d \leq= 5$$
$$S(d) = 33\tfrac{1}{3} - \frac{33\frac{1}{3}}{(100-5)^5} \cdot (100-d)^5 \quad if \quad 5 < d \leq 100 \qquad (3)$$
$$S(d) = 33\tfrac{1}{3} \quad if \quad d > 100$$

It is thus assumed that δ_{min}=5 m, and that the maximum speed S^* of $33\frac{1}{3}$ m/s (120 km/hr) is obtained if δ≥100. For intermediate values of δ, an arbitrary polynomial function was chosen, that however secures the δS-relation be continuous at δ=5 and δ=100, and smooth at δ=100.

For stationary states, the flow F must by definition be equal to S(δ)/δ. The implied speed-flow curve for stationary states is shown in Figure 4-II. The maximum flow of F_{max}=0.965 veh./s is consistent with a speed $S^{\#}$=17.551 m/s and a distance between cars of $\delta^{\#}$=18.195 meters (hence, the density at this maximum flow $D^{\#}$=0.055 vehicles per meter).

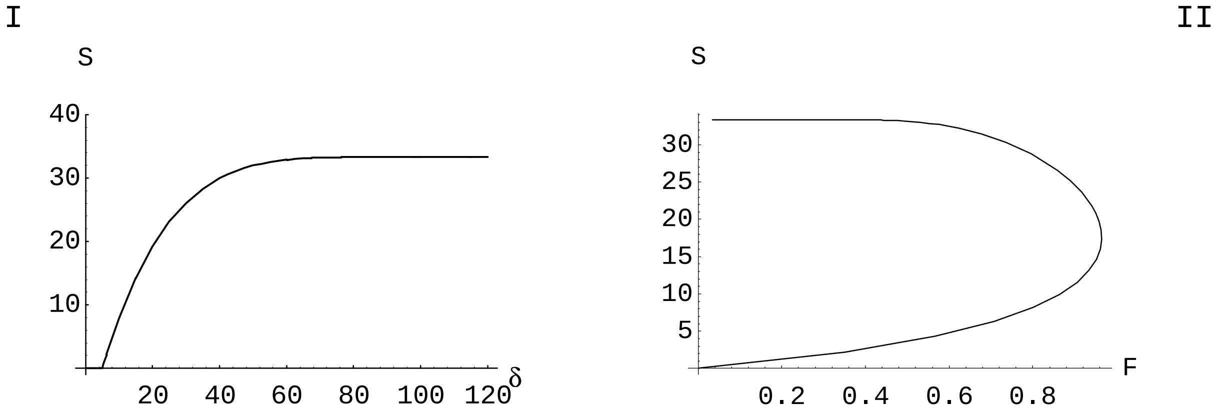

Figure 4. The distance-speed function (I) and the implied speed-flow function for stationary states (II) for the numerical simulation model

The simulation model uses around 30 subsequent car users (depending on the specific simulation), and derives their speeds and positions over time by numerical solution of the first-order differential equations as implied by (3) (note that a driver's speed is the time derivative of his position: $S = \dot{x}$; and that the distance between a driver i and his predecessor i–1 is the difference between their positions: $\delta_i = x_{i-1} - x_i$). All simulations start with an initial stationary state; that is: the very first driver modelled has a constant speed.

4. Dynamic properties of the model

4.1. Introduction

In this section, the dynamic properties of the model outlined above are investigated. The general procedure will be to first provide a general proposition, and next to give an illustration obtained with the simulation model.

Before proceeding, it is first necessary to briefly discuss the notation to be used, and to define precisely the different types of stationary and steady states that will be considered. Table 1 defines these – where the various ways in which subscripts and superscripts will be used are exemplified only for the variable speed (S), and where the logic in the applied nomenclature for various types of stationary states can be deducted from the examples given:

D	density
F	flow
i	index for driver i
S	speed
$S_i^{t=a}$ ($S_i^{x=a}$)	speed of driver i at clock time t=a (speed of driver i when he passes the position x=a)
S^0 (S^1)	equilibrium speed consistent with a stationary state with $F=F^0$ ($F=F^1$) according to Figure 4-II
S^{0h} (S^{1h})	equilibrium speed consistent with the hypercongested stationary state with $F=F^0$ ($F=F^1$) according to Figure 4-II
S^{0n} (S^{1n})	equilibrium speed consistent with the non-hypercongested stationary state with $F=F^0$ ($F=F^1$) according to Figure 4-II
$S_{I,i}^t$ ($S_{II,i}^t$)	the speed for driver i for a given trip I (trip II) at clock time t
t	clock time
x	position along the road, with $0 \leq x \leq X$
δ_i	distance between driver i and his predecessor i–1: $\delta_i = x_{i-1} - x_i$
θ	rate at which the queue before the entrance of the road grows, if there is one
ρ	arrival rate of new users at the entrance of the road – or at the back of the queue before the entrance of the road, if there is one
τ_i	time span driver i's and his predecessor's arrival time at the entrance of the road – or at the back of the queue before the entrance of the road, if there is one
$\tau_i^{x=a}$	time span driver i's and his predecessor's clock time that the point x=a is passed; $\tau_i^{x=0} = \tau_i$ if both driver i–1 and driver i can immediately start with a positive speed upon arrival at the entrance
SSR	Stationary state on the road; situation where the S^x, δ^x and F^x remain constant for all x, and F^x (necessarily) has the same value for all x
SS|SSR	Single speed (SS) SSR: SSR where S^x has the same value for all x
QF|SS|SSR	Queue free (QF) SS|SSR; SS|SSR with no queue before the entrance of the road
SGQ|VS|SSR	'Variable speed' (VS) SSR, where S^x does not have the same value for all x, and with a steadily growing queue (SGQ) before the entrance of the road
trip	a unique path of $S_i^t=0$ for the waiting time in the queue if there is one, and S_i^x defined over all $0 \leq x \leq X$

Table 1. Variables, notation, abbreviations and some definitions

We start with two rather intuitive propositions that will prove to be useful later on. Proposition 1 asserts that there are no feasible trips, which are defined as trips with $S_i^t \leq S(\delta_i^t)$ for all relevant t, superior to (i.e., with an earlier instant of completion than) the trip implied by $S_i^t = S(\delta_i^t)$ for all relevant t.

Proposition 1 A driver that chooses a speed $S_i^t = S(\delta_i^t)$ for all relevant t has the maximum possible x_i^t at every instant, and hence has the earliest possible instant of completion of the trip, when $x_i^t = X$. Travel time minimizing drivers can therefore be assumed to choose $S_i^t = S(\delta_i^t)$ for all relevant t.

Proof Defining the trip with $S_i^t = S(\delta_i^t)$ for all relevant t as trip I, by definition there exists no feasible alternative trip II with $S_{II,i}^t \geq S_{I,i}^t$ for all relevant t and $S_{II,i}^t > S_{I,i}^t$ for some relevant t. Hence, all alternative feasible trips must involve $S_{II,i}^t < S_{I,i}^t$ for at least some relevant t. This can only result in a higher x_i^t later on if trip II implies 'overtaking' of trip I for some relevant t.

However, this is inconsistent with the features that for trip I, $S_{I,i}^t=S(\delta_i^t)$ is the maximum possible speed for δ_i^t, and that x_{i-1}^t is independent of driver i's behaviour.■

From now on, we will therefore assume that all drivers always choose $S_i^t=S(\delta_i^t)$, unless explicitly stated otherwise (in particular in constructions in proofs for propositions).

Proposition 2 states that if driver i–1 has a strictly faster (slower) trip, with the speed never lower (higher) than in the alternative trip, for a given τ_i driver i will also have a strictly faster (slower) trip, with the speed never lower (higher) than in the alternative case.

Proposition 2 Consider two alternative trips I_{i-1} and II_{i-1} for driver i–1, a given τ_i, and the implied trips I_i and II_i for driver i. If driver i–1 has trip I_{i-1} with $S_{I,i-1}^t \geq S_{II,i-1}^t$ for all relevant t and $S_{I,i-1}^t > S_{II,i-1}^t$ for some relevant t in a time interval $[t_1,t_2]$ with $t_2>t_1$, then trip I_i also has $S_{I,i}^t \geq S_{II,i}^t$ for all relevant t and $S_{I,i}^t > S_{II,i}^t$ for some relevant t in a time interval $[t_1^*,t_2^*]$ with $t_2^*>t_1^*$. Driver i therefore completes his trip earlier if driver i–1 has trip I_{i-1} instead of trip II_{i-1}.

Proof First observe that $S_{I,i}^t=S_{II,i}^t$ and $\delta_{I,i}^t=\delta_{II,i}^t$ for $t<t_1$. Now suppose that driver i–1 has the trip I_{i-1}. From the moment t_1 onwards, $x_{I,i-1}^t > x_{II,i-1}^t$ for all $x_{i-1}^t \leq X$ by construction. Hence, if driver 1 would follow trip II_i while driver i–1 follows I_{i-1}, $S_i^t<S(\delta_i^t)$ from t_1^* onwards, where t_1^* gives $MAX\{t_1,t_i^{x=0}\}$. The reason is that $x_{II,i}^t$ would apply throughout, $x_{I,i-1}^t > x_{II,i-1}^t$ from $t_1 \leq t_1^*$ onwards, and $\delta_i^t=x_{i-1}^t-x_i^t$. Since $x_{I,i-1}^t > x_{II,i-1}^t$ from t_1 onwards and since $x_{I,i-1}^t - x_{II,i-1}^t$ is growing during $[t_1,t_2]$ because $S_{I,i-1}^t > S_{II,i-1}^t$ during $[t_1,t_2]$, one can easily construct feasible trips with $S_i^t > S_{II,i}^t$ during some time interval $[t_1^*,t_2^*]$, and $S_i^t \geq S_{II,i}^t$ for all relevant t. From t_1^* onwards, these trips should only satisfy:

$$\int_{t_1^*}^{t^*}\left(S_i^{\ t} - S_{II,i}^{\ t}\right)dt < \left(x_{I,i-1}^{\ t} - x_{II,i-1}^{\ t}\right) \text{ for all } t_1^* \leq t^* \leq t_2^* \tag{4}$$

By Proposition 1, trip I_i must have the lowest possible travel time among all such trips. For trip I_i, $t_2^* \geq t_2$ is required, otherwise the same reasoning could be applied for some instants later than t_2^*. $S_{I,i}^t<S_{II,i}^t$ will not occur after t_1^* because (1) during $[t_1^*,t_2^*]$, $\delta_{I,i}^t>\delta_{II,i}^t$ because $S_{I,i}^t>S_{II,i}^t$; (2) at t_2^*, $\delta_{I,i}^t$ must therefore have approached $\delta_{II,i}^t$ from above; (3) both trips I_i and II_i satisfy $S_i^t=S(\delta_i^t)$; and (4) $S_{I,i-1}^t=S_{II,i-1}^t$ from $t_2^* \geq t_2$ onwards by construction. Before t_1^*, we already saw that $S_{I,i}^t=S_{II,i}^t$.■

4.2. The stationarity of equilibria

Proposition 3 deals with the question of whether all points on a speed-flow curve as depicted in Figures 1-IV or 4-II, and hence an average cost curve as in Figure 2, can be stationary states. They can.

Proposition 3 All points on the speed-flow curve, both non-hypercongested ($S \geq S^{\#}$) and hypercongested ($S<S^{\#}$), can be QF|SS|SSR configurations.

Proof Consider any initial situation with $\{S^0(\delta^0), D^0=1/\delta^0, F^0=S^0/\delta^0\}$ according to the speed-flow curve. Observe that $F^0=S^0/\delta^0$ by construction of the speed-flow curve. Consider an arrival rate $\rho^0=1/\tau^0=F^0$. When driver i appears at the entrance of the road and there is no queue, driver i–1 has travelled $S^0\cdot\tau^0$ meters. That means that $\delta_i^{x=0}=S^0\cdot\tau^0=S^0/\rho^0=S^0/F^0=\delta^0>\delta_{min}$. Hence, driver i needs not queue, obtains δ^0 and S^0 straight away, and will maintain this speed throughout his trip since we started with an initial situation with constant speeds S^0.■

Proposition 3 is illustrated by considering the two points in Figure 4-II for which F=0.7, namely {δ,S}={44.33,31.03} and {δ,S}={8.8,6.17}, where the latter is the 'hypercongested' equilibrium, and setting ρ=F=0.7. Figures 5-I and 5-II show for both equilibria the 'clock time-speed functions' (tS-functions) for 20 subsequent drivers. These functions show a driver's speed as a function of clock time, starting with his arrival at the entrance of the road (or at the back of the queue, if there is one) up to the instant the road's exit is passed. The fact that for both cases, these 20 individual functions merge into one horizontal line illustrates the proposition: speeds, and hence – as a necessary condition as well as a consequence – also densities, and therefore also flows, all are and remain constant over time and place.

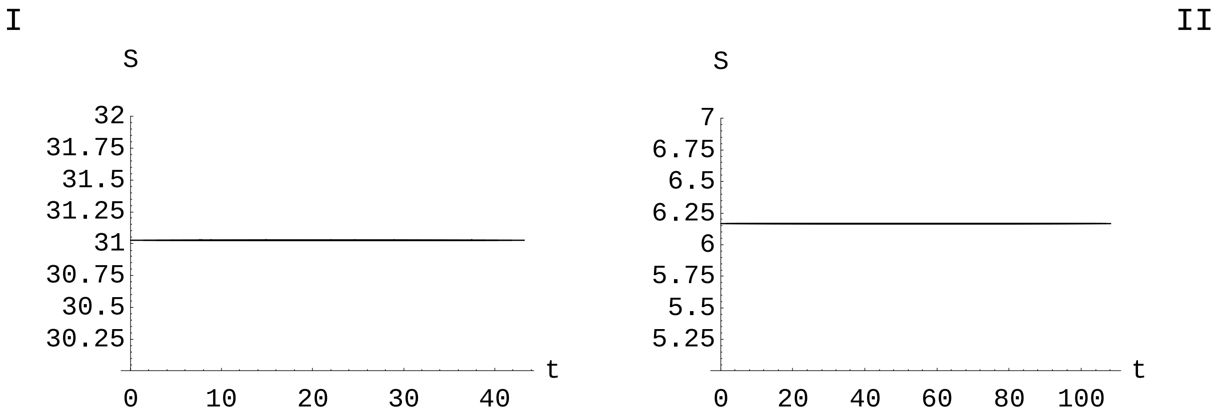

Figure 5. The clock time-speed functions for the non-hypercongested (I) and hypercongested (II) stationary states with r=F=0.7

4.3. The dynamic stability of non-hypercongested stationary states

Proposition 4 states that the non-hypercongested stationary states on a speed-flow curve as depicted in Figures 1-IV or 4-II, and hence an average cost curve as in Figure 2, are dynamically stable.

Proposition 4 All non-hypercongested points on the speed-flow curve, with $S \geq S^\#$, are dynamically stable QF|SS|SSR configurations: starting with an initial non-hypercongested QF|SS|SSR $\{S^{0n}(\delta^{0n})\geq S^\#, D^{0n}=1/\delta^{0n}, \rho^0=F^0=S^{0n}/\delta^{0n}\}$, a persistent change of ρ^0 to $\rho^1 \leq F_{max}$ will cause the system to converge to the non-hypercongested QF|SS|SSR $\{S^{1n}(\delta^{1n})\geq S^\#, D^{1n}=1/\delta^{1n}, \rho^1=F^1=S^{1n}/\delta^{1n}\}$.

Proof Consider an initial QF|SS|SSR $\{S^{0n}(\delta^{0n})>S^{\#},\ D^{0n}=1/\delta^{0n},\ \rho^0=F^0=S^{0n}/\delta^{0n}\}$. Label the last driver with $\tau_i=1/\rho^0$ as driver i=0, and the subsequent first driver with $\tau_i=1/\rho^1$ as driver i=1. Denote the clock time at which driver 0 starts as t=0. Driver 0 maintains S^0 by Proposition 3. Consider $\rho_1>\rho_0$. Observe that driver 1 starts at $t=1/\rho_1=\tau_1$, so that we have:

$$d_{\min} < d^{1n} = \frac{S^{1n}}{r^1} < d_1^{t=t_1} = \frac{S^{0n}}{r^1} < d^0 = \frac{S^{0n}}{r^0} \quad \Rightarrow \quad S^{1n} < S_1^{t=t_1} < S^{0n} \tag{5}$$

Therefore, driver 1 starts at a speed below driver 0's speed, and both δ_1 and (hence) S_1 start growing immediately after driver 1's start; until $\delta_1=\delta^{0n}$ and $S_1=S^{0n}$. Driver 1's speed is therefore strictly greater than S^{1n} throughout his trip, but strictly smaller than S^{0n} for a positive first number of meters. Reapplying the same reasoning to subsequent drivers with i>1, it is clear that S^{1n} remains a lower bound for each of these drivers' speeds, and that $d_i^{t=i\cdot t_1} > d_{\min}$ for all i so that no queuing will occur. Still, average speeds (over the entire trip) must be decreasing for subsequent drivers by Proposition 2. Since the non-hypercongested QF|SS|SSR $\{S^{1n}(\delta^{1n})\geq S^{\#},\ D^{1n}=1/\delta^{1n},\ \rho^1=F^1=S^{1n}/\delta^{1n}\}$ is the only queue free stationary state with $S^x\geq S^{1n}$ for all x consistent with ρ^1, the system must asymptotically approach this stationary state. The proof for $\rho_1<\rho_0$ is analogous to the present one, and is therefore omitted.■

Proposition 4 is illustrated in Figure 6 by considering the initial stationary state with $\rho^0=0.7$ depicted in Figure 5-I, and considering a change to $\rho^1=0.6$ and $\rho^1=0.8$ in Figures 6-I and 6-II, respectively. The diagrams show the tS-functions for 30 subsequent drivers, and depict how the new stationary states with $S^1=32.5$ and $S^1=28.7$ are approached asymptotically. The simulation results revealed, as expected, that the flow along each point of the road also approached the new stationary state values of 0.6 and 0.8, respectively.

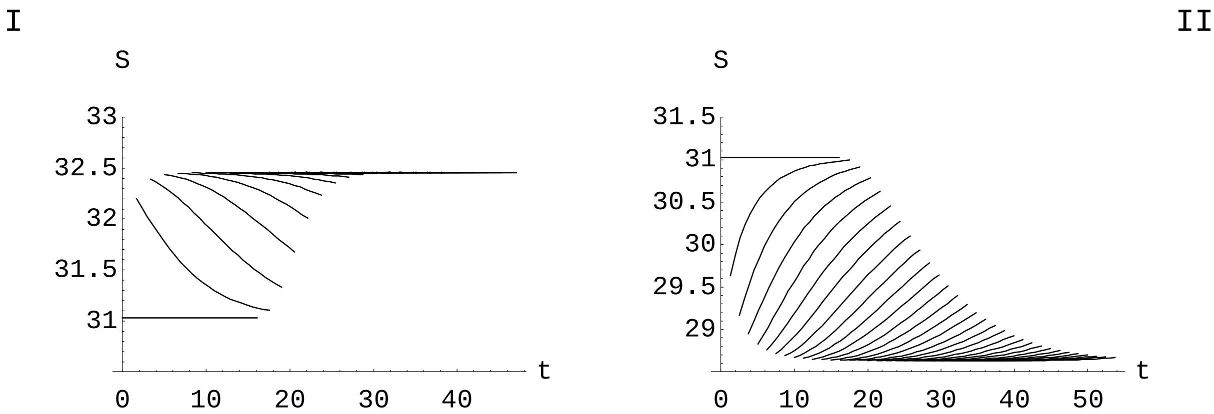

Figure 6. The clock time-speed functions for an initial non-hypercongested stationary state with r=F=0.7 after a change in r to r=0.6 (I) and r=0.8 (II)

4.4. The dynamic stability of hypercongested stationary states

Propositions 5 and 6 to be given below imply that the hypercongested stationary states on a speed-flow curve as depicted in Figures 1-IV or 4-II, and hence an average cost curve as in Figure 2, are dynamically unstable.

Proposition 5 Starting with an initial hypercongested QF|SS|SSR $\{S^{0h}(\delta^{0h})<S^{\#}, D^{0h}=1/\delta^{0h}, \rho^0=F^0=S^{0h}/\delta^{0h}\}$, a persistent reduction of ρ^0 to ρ^1 with $\rho^1<\rho^0\leq F_{max}$ will cause the system to converge to the non-hypercongested QF|SS|SSR $\{S^{1n}(\delta^{1n})>S^{\#}, D^{1n}=1/\delta^{1n}, \rho^1=F^1=S^{1n}/\delta^{1n}\}$, and hence *not* to the hypercongested QF|SS|SSR $\{S^{1h}(\delta^{1h})<S^{\#}, D^{1h}=1/\delta^{1h}, \rho^1=F^1=S^{1h}/\delta^{1h}\}$.

Proof Consider the initial QF|SS|SSR $\{S^{0h}(\delta^{0h})<S^{\#}, D^{0h}=1/\delta^{0h}, \rho^0=F^0=S^{0h}/\delta^{0h}\}$. Again label the last driver with $\tau_i=1/\rho^0$ as driver i=0, and the subsequent first driver with $\tau_i=1/\rho^1$ as driver i=1. Denote the clock time at which driver 0 starts as t=0. Driver 0 maintains S^{0h} by Proposition 3. Observe that driver 1 starts at $t=1/\rho_1=\tau_1$, so that we have:

$$d_{\min} < d^{1h} = \frac{S^{1h}}{r^1} < d^{0h} = \frac{S^{0h}}{r^0} < d_1^{t=t_1} = \frac{S^{0h}}{r^1} \quad \Rightarrow \quad S^{1h} < S^{0h} < S_1^{t=t_1} \tag{6}$$

Therefore, driver 1 starts at a speed exceeding S^{0h}, which is driver 0's speed, and both δ_1 and (hence) S_1 start declining immediately after driver 1's start; until $\delta_1=\delta^{0h}$ and $S_1=S^{0h}$. Driver 1's speed is therefore strictly greater than driver 0's speed for a positive first number of meters of his trip and will not fall below S^{0h} throughout the rest of the trip, and is hence strictly greater than S^{1h} throughout the trip. By Proposition 2, driver 2 will have a strictly higher speed than driver 1 has had over the first meters, and will never have a lower speed. Reapplying the same reasoning to subsequent drivers with i>3, it is clear that average speeds (over the entire trip) must be increasing for subsequent drivers by Proposition 2 and that queuing will not occur. Since the non-hypercongested QF|SS|SSR $\{S^{1n}(\delta^{1n})>S^{\#}, D^{1n}=1/\delta^{1n}, \rho^1=F^1=S^{1n}/\delta^{1n}\}$ is the only queue free stationary state with $S^x\geq S^{0h}$ for all x consistent with ρ^1, this system must asymptotically approach this stationary state.■

Proposition 5 is illustrated in Figures 7-I and 7-II. In order to be able to show the full transitional phase, we start with an initial stationary state with $\rho^0=0.96$ at a hypercongested speed $S^{0h}=16$, which is quite close to (but of course below) $S^{\#}=17.551$. Furthermore, ρ^1 is set at 0.6. Figure 7-I shows the tS-functions for 33 subsequent drivers, and depicts how the new stationary state with $S^{1n}=32.5$ is approached asymptotically. Figure 7-II shows the flow levels at the entrance (F_{en}) and the exit (F_{ex}) of the road, which are defined as $1/(t_i^{x=0}-t_{i-1}^{x=0})$ and $1/(t_i^{x=X}-t_{i-1}^{x=X})$, respectively. The flow at the entrance immediately falls to the level equal to the new arrival rate of $\rho=0.6$. It takes a while before the flow at the exit reaches this level; this happens when the speed at x=X has reached the new stationary state level of S=32.5. Figure 7 thus demonstrates how the initial hypercongestion vanishes after a reduction in the arrival rate, and how the system evolves to the non-hypercongested rather than the hypercongested stationary state consistent with the new arrival rate. Finally, Figure 7-II shows that during the transitional phase, neither the flow at the entrance nor the flow at the exit reaches F_{max}. In other words, the system does not somehow 'follow' the speed-flow curve (Figure 4-II) when moving from a hypercongested stationary state to a non-hypercongested one: transitional phases certainly need not be sequences of stationary states.

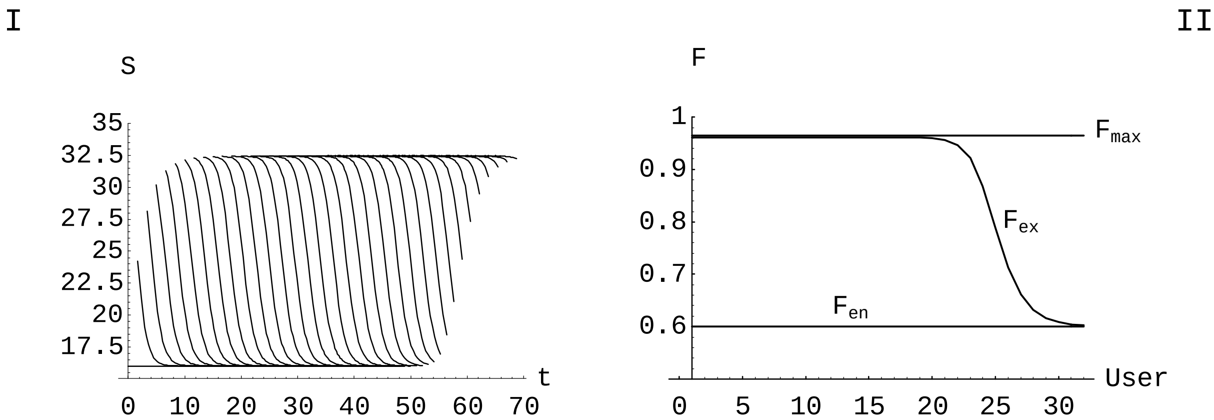

Figure 7. The clock time-speed function (I) and the flow at the entrance and exit of the road (II) after a reduction in the arrival rate to r=0.6 following an initial hypercongested stationary state with r=F=0.96

Proposition 6 Starting with an initial hypercongested QF|SS|SSR $\{S^{0h}(\delta^{0h})<S^{\#}, D^{0h}=1/\delta^{0h}, \rho^0=F^0=S^{0h}/\delta^{0h}\}$, a persistent increase of ρ^0 to ρ^1 with $\rho^0<\rho^1\leq F_{max}$ will *not* cause the system to converge to a QF|SS|SSR; neither the hypercongested QF|SS|SSR $\{S^{1h}(\delta^{1h})<S^{\#}, D^{1h}=1/\delta^{1h}, \rho^1=F^1=S^{1h}/\delta^{1h}\}$, nor the non-hypercongested QF|SS|SSR $\{S^{1n}(\delta^{1n})\geq S^{\#}, D^{1n}=1/\delta^{1n}, \rho^1=F^1=S^{1n}/\delta^{1n}\}$. Instead, provided the road is sufficiently long, the system will converge to a SGQ|VS|SSR with $\theta=\rho^1-\rho^0$, $F^x=\rho^0$ for all $0<x\leq X$, but where S^x and δ^x vary along the road.

Proof Consider the initial QF|SS|SSR $\{S^{0h}(\delta^{0h})<S^{\#}, D^{0h}=1/\delta^{0h}, \rho^0=F^0=S^{0h}/\delta^{0h}\}$. Again, label the last driver with $\tau_i=1/\rho^0$ as driver i=0, the subsequent first driver with $\tau_i=1/\rho^1$ as driver i=1 and the clock time at which driver 0 starts as t=0. Driver 0 maintains S^{0h} by Proposition 3. Observe that driver 1 starts at $t=1/\rho_1=\tau_1$, so that we have:

$$d_1^{t=t_1}=\frac{S^{0h}}{r^1}<d^{0h}=\frac{S^{0h}}{r^0}<d^{1h}=\frac{S^{1h}}{r^1}<d^{1n}=\frac{S^{1n}}{r^1}\quad\Rightarrow\quad S_1^{t=t_1}<S^{0h}<S^{1h}<S^{1n} \tag{7}$$

Therefore, driver 1 starts at a speed below driver 0's constant speed S^{0h}, and hence certainly below S^{1h} and S^{1n}. Hence, δ_1 starts growing immediately after driver 1's start; and so does S_1, provided $d_1^{t=t_1}\geq d_{\min}$; otherwise, driver 1 starts his trip queuing and obtains a positive speed only at $t=\delta_{min}/S^{0h}$. In both cases, driver 1 will therefore be accelerating from either $t=\tau_1$ (without queuing) or $t=\delta_{min}/S^{0h}$ (with queuing) onwards, until $\delta_1=\delta^{0h}$ and $S_1=S^{0h}$. Driver 1's speed is therefore strictly below driver 0's speed for a positive first number of meters of his trip and will not exceed S^{0h} throughout the rest of the trip, and is hence strictly below S^{1h} and S^{1n} throughout the trip. As a consequence, driver 2 must have a strictly lower speed than driver 1 has had over the first meters $(d_2^{t=2\cdot t_1}<d_1^{t=t_1})$. Driver 2 will therefore also be accelerating over the first meters, and will also never have a speed exceeding S^{0h}. Hence, driver 2 will have a lower average speed (over the entire trip) than driver 1, which is in line with Proposition 2. Reapplying the same reasoning to subsequent drivers with i>3, it is clear that average speeds (over the entire trip) must be decreasing for subsequent drivers by Proposition 2, that all drivers will accelerate over the first meters and that speeds will never exceed S^{0h}. Since there exist no QF|SS|SSR $\{S^1(\delta^1), D^1=1/\delta^1, \rho^1=F^1=S^1/\delta^1\}$ with $S^x \leq S^{0h}$ for all x consistent with ρ^1, the system cannot converge to a QF|SS|SSR.

To prove that the system will converge to a SGQ|VS|SSR with $\theta=\rho^1-\rho^0$ and $F^x=\rho^0$ for all $0<x\leq X$, first observe that there is a feasible trip II_1 for driver 1 that implies 'passing times' at all points x>0 along the road exactly τ^0 seconds later than driver 0's passing time at those points. Driver 1 should for this trip II_1 wait $1/\rho^0-1/\rho^1$ seconds at the entrance, and drive at a constant speed S^0 throughout. This trip II_1 is feasible by Proposition 3. By Proposition 1, driver 1's actual trip I_1 with $S_1^t=S(\delta_1^t)$ should have $x_{I,1}^t \geq x_{II,1}^t$. Now defining in general trip II_i for driver i ($i\geq 1$) as waiting $i\cdot(1/\rho^0-1/\rho^1)$ seconds at the entrance and driving at a constant speed S^0 afterwards, it is clear that since $x_{I,1}^t \geq x_{II,1}^t$ and since trip II_2 is feasible given trip II_1 by Proposition 3, it should also be feasible given trip I_1. Hence, $x_{I,2}^t \geq x_{II,2}^t$ by Proposition 1, and hence trip II_3 is feasible. Reapplying the same reasoning for subsequent drivers, it follows that all trips II_i are feasible and that ρ^0 therefore remains a lower bound on the flow at any point of the road: no driver i will arrive more than $i\cdot\tau^0=i/\rho^0$ seconds after driver 0 at any point of the road because trip II_i is always feasible and since $x_{I,i}^t \geq x_{II,i}^t$.

If the drivers have a constant speed over a last positive number of meters of their trip, this lower bound on the flow implies a lower bound on that constant speed equal to S^{0h}. Since speeds at the same time cannot exceed S^{0h} as proven above, it follows that if the trips end with a constant speed in the new stationary state, it must be S^{0h} implying ρ^0. Because speeds are below S^{0h} over a first positive number of meters of the road, there is a minimum positive acceleration drivers must have during these first meters. The reason is that the flows were shown not to fall below ρ^0 at any point of the road, while the time needed to travel δ meters (which is the inverse of the flow) at a speed $S(\delta)<S^{0h}$ would exceed $1/\rho^0$ seconds. Hence, if the road is sufficiently long, the accelerating drivers must obtain the constant speed S^{0h} at some point. Also, because flows cannot be below ρ^0 at any point of the road, there exist positive minimum speeds for all x>0. Hence, the decreasing speeds (over time) over the first meters of the road have positive lower bounds which increase with x, and the system must asymptotically be approaching a stationary state. A stationary state implies that the flow be equal along the road. Since this flow will be ρ^0 provided the road is sufficiently long, whereas the arrival rate is $\rho^1>\rho^0$, a queue must be growing at a rate $\theta=\rho^1-\rho^0$. Because queuing occurs, drivers must start with a zero speed.■

Figure 8 illustrates Proposition 6 using an initial hypercongested stationary state with $\rho^0=0.7$ and $S^0=6.17$ and considering a shock to $\rho^1=0.8$. Figure 8-I shows how the subsequent drivers' starting speeds drop quickly and reach 0 already for driver i=4, implying queuing time. What is clear from Figure 8-I, however, is that once in motion, drivers will have a rather rapid acceleration, which is consistent with the fact that flows cannot fall below $\rho=0.7$ also on that part of the road where the speed is below S^0. Note that the flat segments at S=0 in the clock time-speed functions from driver i=4 onwards represent the queuing of these drivers. Figure 8-II shows the flow at the entrance and exit of the road (for determining the flow at the entrance, a queuing driver's instant of passing the entrance is determined as the moment at which his predecessor has travelled δ_{min} meters, so that the driver himself can start speeding up). Only during the transitional phase can the flow at the entrance temporarily reach the level of 0.8. From driver i=4 onwards however, who starts with a zero speed, the flow at the entrance drops quickly to the (old and new) stationary state level of 0.7.

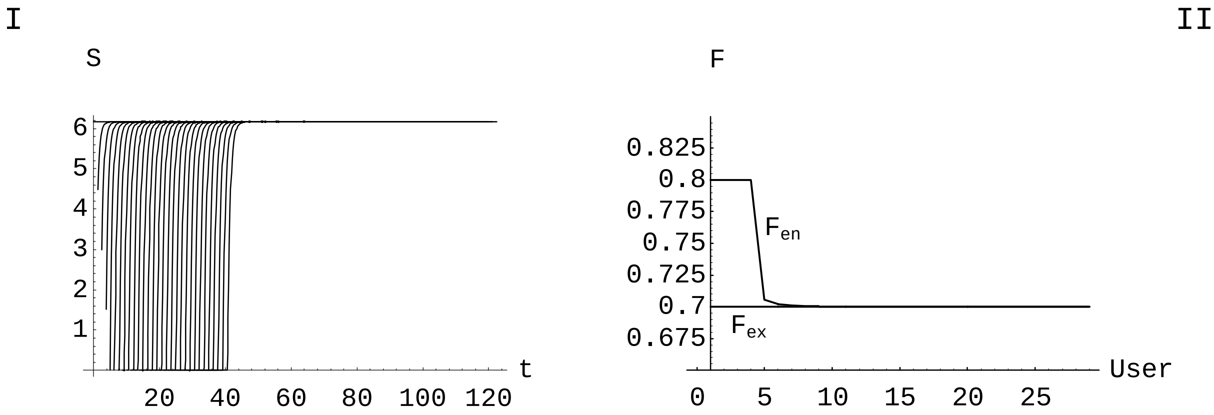

Figure 8. The clock time-speed function (I) and the flow at the entrance and exit of the road (II) after an increase in the arrival rate to r=0.8 following an initial hypercongested stationary state with r=F=0.7

Propositions 5 and 6 together imply that the hypercongested stationary states on a speed-flow curve as depicted in Figures 1-IV or 4-II, and hence an average cost curve as in Figure 2, are in fact dynamically unstable. Although these configurations can be stationary states, there are no equilibrium paths from other initial stationary states that would lead to these hypercongested stationary states. Coming from non-hypercongested stationary states, the system will remain in the non-hypercongested regime as long as ρ remains below F_{max}. Coming from hypercongested stationary states, the system will move to a non-hypercongested stationary state after a reduction in ρ, and will maintain the initial flow level in conjunction with a growing queue after an increase in ρ.

Therefore, the only possible route towards hypercongestion that is still open would involve hypercongestion caused by an arrival rate ρ exceeding F_{max}. This last possibility will be explored in the following sub-section.

4.5. Arrival rates exceeding the road's maximum capacity

Proposition 7 states that, when starting with an initial non-hypercongested state, an increase in the arrival rate to a value exceeding F_{max} will lead to a variable speed stationary state consistent with F_{max} on the road, while the queue before the entrance of the road will be growing at a rate $\theta=\rho-F_{max}$.

Proposition 7 Starting with an initial non-hypercongested QF|SS|SSR $\{S^{0n}(\delta^{0n})\geq S^{\#}, D^{0n}=1/\delta^{0n}, \rho^0=F^0=S^{0n}/\delta^{0n}\}$, a persistent increase in ρ^0 to $\rho^1>F_{max}$ will cause the system to converge to a maximum flow SGQ|VS|SSR with $F^x=F_{max}$ for all $0<x\leq X$, and $\theta=\rho^1-F_{max}$, provided the road is sufficiently long.

Proof Consider the initial QF|SS|SSR $\{S^{0n}(\delta^{0n})\geq S^{\#}, D^{0n}=1/\delta^{0n}, \rho^0=F^0=S^{0n}/\delta^{0n}\}$. Again, label the last driver with $\tau_i=1/\rho^0$ as driver i=0, the subsequent first driver with $\tau_i=1/\rho^1$ as driver i=1 and the clock time at which driver 0 starts as t=0. Driver 0 maintains S^{0n} by Proposition 3. Observe that driver 1 starts at $t=1/\rho_1=\tau_1$, so that we have:

$$d_1^{t=t_1} = \frac{S^{0n}}{r^1} < d^{0n} = \frac{S^{0n}}{r^0} \quad \Rightarrow \quad S_1^{t=t_1} < S^{0n} \tag{8}$$

In the same manner as in the proof to Proposition 6, it follows that driver 1's speed is therefore strictly below driver 0's speed for a first positive number of meters of his trip and will not exceed S^{0n} throughout the rest of the trip, that driver 1 will therefore be accelerating from either $t=\tau_1$ (without queuing) or $t=\delta_{min}/S^{0n}$ (with queuing) onwards, until $\delta_1=\delta^{0n}$ and $S_1=S^{0n}$, and that by reapplying the same reasoning to subsequent drivers with i>2, it can be proven that average speeds (over the entire trip) must be decreasing for subsequent drivers by Proposition 2, that the speed over a first positive number of meters of the road must be decreasing over time, that all drivers will be accelerating over the first meters, and that speeds will never exceed S^{0n}.

Now define trip I_i again as the trip driver i actually chooses, with $S_i^t=S(\delta_i^t)$; and define trip II_i for driver i ($i\geq 1$) as waiting $i\cdot(1/F_{max}-1/\rho^1)$ seconds at the entrance and driving at $S^{\#}$ afterwards. Observe that such a trip II is feasible for driver 1 since $S^{\#}\leq S^{0n}$ (the latter is his predecessor's speed), and since he will start driving not sooner than at $t=1/F_{max}$, at which for δ_1 we find:

$$d_1^{t=\frac{1}{F_{max}}} = \frac{S^{0n}}{F_{max}} \geq d^{\#} = \frac{S^{\#}}{F_{max}} \tag{9}$$

Next, since trip II_2 is feasible given trip II_1 by Proposition 3, it would certainly be feasible given trip I_1 since $x_{I,1}^t\geq x_{II,1}^t$ by Proposition 1. Hence, $x_{I,2}^t\geq x_{II,2}^t$ by Proposition 1, and trip II_3 is feasible by Proposition 3. Reapplying the same reasoning to subsequent drivers, it is evident that all trips II_i are feasible. It thus follows that $x_{I,i}^t\geq x_{II,i}^t$ for all i. Driver i's actual 'passing time' at any point x along the road, $t_{I,i}^x$, can therefore be no later than $t_{II,i}^x$, so that:

$$t_{I,i}^{\ x} \leq t_{II,i}^{\ x} = \frac{i}{F_{max}} + \frac{x}{S^{\#}} \tag{10}$$

Observing that the flow at a certain point x along the road at a certain time clock time t can be written as $F^x=1/(t_i^x-t_{i-1}^x)$, (10) implies that F_{max} is a lower bound on the stationary state flow at any point x along the road (F_{max} of course needs not be a lower bound on the *actual* flow during the transitional phase; observe for instance that we start with $F^0<F_{max}$ from the outset).

Because speeds are decreasing over time while the flow remains positive, the system must asymptotically converge to a stationary state. If the drivers do reach a constant speed during some part of their trip during this stationary state it must be $S^{\#}$, because F_{max} is a lower bound on the flow. Other speeds can only apply if drivers are accelerating when driving that speed different from $S^{\#}$, because otherwise they can not travel the associated distance of δ meters in $1/F_{max}$ seconds (or less), as is required by the lower bound on flow. However, if drivers do accelerate, speeds cannot exceed $S^{\#}$ in a stationary state with a flow at least equal to F_{max}. The reason is that acceleration would be needed to keep the flow at F_{max} (or more) for any $S>S^{\#}$. This however implies unlimited acceleration after the speed has exceeded $S^{\#}$, whereas an upper bound on speeds of S^{0n} was identified above. Hence, $S^{\#}$ is the maximum speed that can occur when F_{max} is the minimum flow. If the road is sufficiently long, this implies that the system must converge to a stationary state where drivers accelerate up to a speed of $S^{\#}$ and maintain this speed for the rest of their trip. The implied flow is then F_{max}, so that $\theta=\rho^1-F_{max}$. Because queuing occurs, drivers must start with a zero speed.■

Proposition 7 is illustrated in Figure 9, where the non-hypercongested stationary state with $\rho^0=0.7$ is followed by an increase in flow to $\rho^1=1.8$. Note that Figure 9-II shows that flows can temporarily – during transitional phases – exceed F_{max}. This can of course only occur if drivers are accelerating.

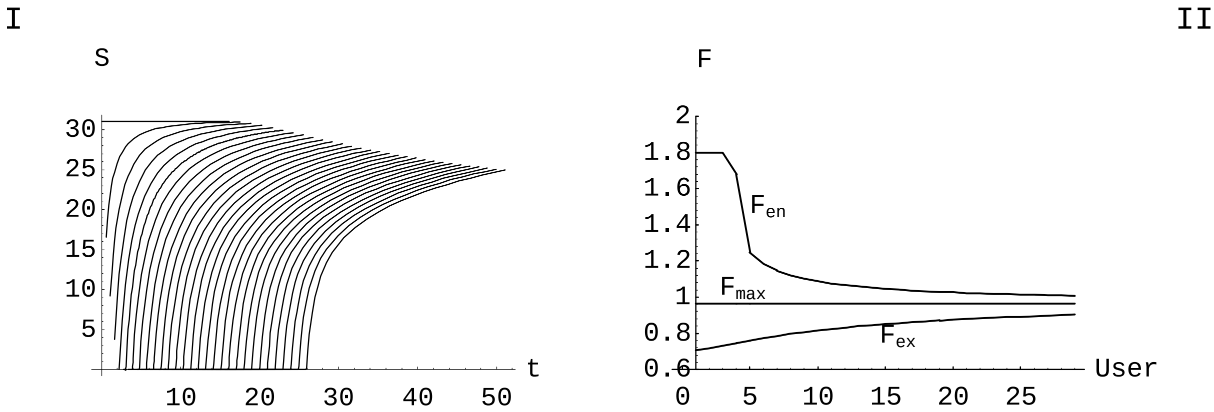

Figure 9. The clock time-speed function (I) and the flow at the entrance and exit of the road (II) after an increase in the arrival rate to r=1.8 (>F_{max}) following an initial non-hypercongested stationary state with r=F=0.7

Otherwise, the interpretation of Figure 9 resembles that of Figure 8. Clearly then, also an increase in the arrival rate to values exceeding F_{max} does not lead to a hypercongested equilibrium, at least not to hypercongested single speed equilibria with $S<S^{\#}$ as illustrated in Figures 1 and 3. Although speeds below $S^{\#}$ will certainly occur, the distinguishing factor is that the flow will not fall below F_{max}. The explanation for this paradox is that positive speeds below $S^{\#}$ only occur when drivers are accelerating.

4.6. Conclusion to Section 4

The above propositions imply that the hypercongested segments of the standard backward bending speed-flow and average cost curves do not represent dynamically stable market equilibria. The reason that these configurations are nevertheless suggested as possible equilibria by the standard model is that this model is inherently static in nature, and therefore presupposes an initial state equal to the equilibrium suggested, completely ignoring questions of dynamic stability. Likewise, the standard static model does not predict and is not able to describe the steady states, with the maximum flow on the road and steadily growing queues before the entrance, that would result from persistent arrival rates of new users exceeding the capacity of the road.

5. Implications for models of road traffic congestion with elastic continuous demand or with peak demand

The analyses in the previous section were performed under the assumption of a rather rigid demand structure, namely inelastic demand at a fixed, exogenous rate, which could change in value only once in order to study the dynamic stability of equilibria. In this section, the implications of the results presented above for models with more elaborate demand structures will be discussed briefly and only qualitatively; partly in order to save space, and partly because the conclusions confirm those already predicted in Verhoef (1998).

As in Verhoef (1998), two archetypes of demand will be considered here. The first is 'continuous demand', where the demand function is stable over time, and demand is unrelated over time: a driver does not consider rescheduling, but only takes a go-no-go decision based on the actual costs of the trip. This type of demand is in the same spirit as the type of demand implicitly assumed in static models of road traffic congestion, and actually in most standard static demand-supply diagrammatic analyses of market equilibria. The second type of demand is 'peak demand', which is the type of demand assumed in dynamic models of road traffic congestion based on bottleneck congestion (Vickrey, 1969; Arnott, De Palma and Lindsey, 1993, 1998), or flow congestion (Henderson, 1974, 1981; Chu 1995).

For models with a not completely inelastic 'continuous' demand function, the propositions in the previous sections imply a general shape of the average cost function for stationary states. First of all, the backward-bending segment of the AC curve in Figure 2 can be ignored in a market analysis, since the implied equilibria are not dynamically stable. Hence, the suggested problem of multiple equilibria with a demand curve like E in Figure 2 will also not be relevant, since only the equilibrium labelled z will be dynamically stable. For demand curves that do have an intersection with the AC curve on the lower segment, such as E in Figures 2 and 10, this particular intersection can thus be expected to be the single long run stationary state equilibrium (F_E in Figure 10). For demand curves that do not have such an intersection, such as E' in Figure 10, the demand exceeds the capacity of the road for any cost level associated with non-hypercongested speeds exceeding $S^{\#}$ in absence of queuing. Hence, for such cases, one would expect an arrival rate exceeding F_{max}, and hence a queue to be building up, for at least some period of time. In contrast to the inelastic exogenous demand assumed in the previous section, we should now take account of the fact that because of the growing queue and falling speeds during such a transitional phase, total travel times will be increasing. Therefore, because of demand elasticity, the arrival rate will be decreasing during this phase (see the arrows). Proposition 7 asserts that the stationary state flow on the road will not be below F_{max}, so that as soon as the arrival rate has fallen to a value of F_{max}, one can expect a stationary state equilibrium with F_{max} and variable speeds on the road, and a queue of constant length before the road's entrance. The implied waiting time costs in the queue (qc) is then just sufficient to keep away excessive demand (arrival rates exceeding F_{max}) in the non-intervention outcome.

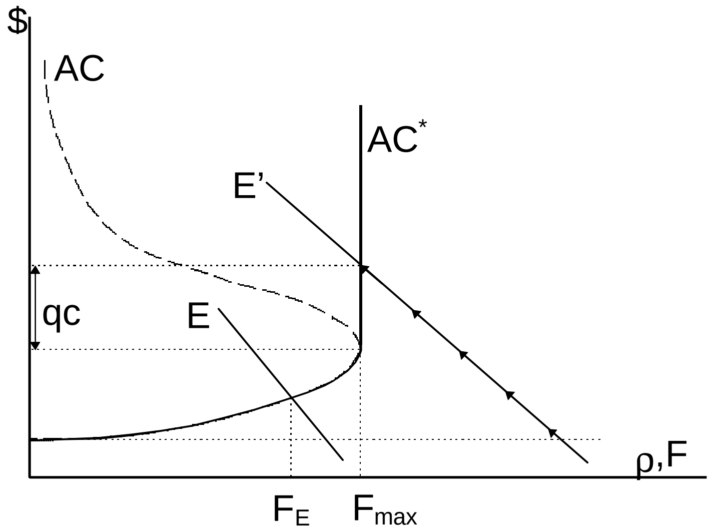

Figure 10. Dynamically stable stationary state equilibria (AC^*) and dynamically unstable (dashed part of AC) equilibria with continuous demand

Therefore, the curve denoted AC^* in Figure 10 gives the average cost function for dynamically stable stationary states with continuous demand, as suggested by the analyses in the previous section. The hypercongested equilibria on the upper segment of the original AC-curve are not included in AC^*, because of their dynamic instability. However, all non-hypercongested equilibria on the lower segment of the original AC are included. Furthermore, the possible stationary state equilibria with an arrival rate of $\rho=F_{max}$, a maximum flow F_{max} on the road, and a queue of constant length, are represented by the vertical segment of AC^*.

Given the shape of AC^*, an implied marginal social cost curve MC^*, located above AC^*, could now be derived using the identity $MC^*=AC^*+F\cdot\partial AC^*/\partial F$. The optimal stationary state flow levels could subsequently be identified as the unique intersection between this MC^* curve and the prevailing demand curve, and the optimal toll as $MC^*-AC^*=F\cdot\partial AC^*/\partial F$ in the optimum. For reasons of graphical clarity, this MC^*-curve is not shown in Figure 10 (see also Verhoef, 1998). The reader may easily verify, however, that with this curve the optimal flow will never exceed the flow in the unregulated equilibrium.

Next, in models of peak demand, it is assumed that a certain group of drivers, with an endogenous size when elastic demand is considered, wish to use the road during the same period. Normally, it is assumed that all potential users have the same desired arrival time, and that deviations from this desired arrival time imply 'scheduling costs'. The dynamic equilibrium then requires the sum of scheduling costs, travel delay costs and tolls (if levied) to be constant over the peak. The distribution of trips, scheduling costs and travel delay costs over time are thus determined endogenously.

For such demand structures, the present model can be expected to behave in a way similar to the bottleneck model as proposed by Vickrey (1969), and extended in various directions by Arnott, De Palma and Lindsey (1993, 1998). The model presented above shares the property with the bottleneck model that when the arrival rate of users exceeds the maximum capacity, a queue will be building up, while the road (or the bottleneck) operates at maximum capacity. In contrast to the pure bottleneck model, in the present model the flow on the road will not immediately take on the value F_{max} as soon as $\rho \geq F_{max}$. However, this transition need not take much time and may therefore be of only limited practical importance. For instance, in the numerical example presented in Figure 9, already after some six or seven drivers starting after the moment that ρ exceeded F_{max} for the first time, both F_{en} and F_{ex} approach F_{max} rather closely.

A further distinction between the models is that in the bottleneck model a queue starts growing immediately after the first driver has started his trip, and does not vanish completely until the last driver has passed the bottleneck. The present model would normally have some periods of flow congestion without queuing over the first and last phases of the unregulated peak. Likewise, it is likely that for the present formulation, not all congestion will be eliminated in the optimum, although all queuing will be. In contrast, like in the dynamic models of flow congestion with zero group velocity as presented by Henderson (1974, 1981) and Chu (1995), some flow congestion – that is, speeds below S^* – can be expected to remain existent in the optimum.

6. Conclusion

A new approach to model road traffic congestion was presented, based on simple car-following theory, allowing for finite group velocity and discrete vehicles. The model constitutes a full-fledged dynamic version of the standard static model of road traffic congestion based on the so-called 'fundamental diagram of road traffic congestion'. It was proven formally and illustrated numerically that the suggested hypercongested equilibria of the standard static model are in fact dynamically unstable. The model presented is fully consistent with, and reproduces non-hypercongested stationary state outcomes found in the standard static model, as long as the arrival rate of new users is below the maximum capacity of the road. When this maximum capacity is exceeded, the model behaves in the same manner as Vickrey's (1969) model of bottleneck congestion, by predicting a maximum flow on the road in conjunction with a growing queue before the road's entrance. Therefore, the model presented above offers an integration and a generalization of these two important types of models. An important next step would be to integrate the model with a more elaborate demand structure, and to study the behaviour of such a model during transitional phases more explicitly than in the qualitative last section of the paper.

The conclusion that the hypercongested equilibria suggested by the standard static model of road traffic congestion are in fact dynamically unstable should not be mistaken to imply that every observation of hypercongested speeds in reality actually falsifies the model presented here. In contrast, the present model would only suggest that whenever such speeds are observed, it is unlikely that the cause is to be found in flow congestion on that road itself, but that the true reason for such speeds may often be a downstream bottleneck. Therefore, optimal pricing rules should then not primarily be based on the road's characteristics, but rather on the bottleneck's capacity, which will in general imply different optimal road prices.

References

Agnew, C.E. (1973) "The dynamic control of congestion – prone systems through pricing". Report No. 6, Stanford University Center for Interdisciplinary Research.

Arnott, R., A. de Palma and R. Lindsey (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 and R. Lindsey (1998) "Recent developments in the bottleneck model". In: K.J. Button and E.T. Verhoef (1998) *Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility* Edward Elgar, Cheltenham (forthcoming).

Chu, X. (1995) "Endogenous trip scheduling: the Henderson approach reformulated and compared with the Vickrey approach" *Journal of Urban Economics* **37** 324-343.

Chu, X. and K.A. Small (1996) "Hypercongestion" Paper prepared for the meeting of the American Real Estate and Urban Economics Association, New Orleans, Jan. 1997.

De Meza, D. and J.R. Gould (1987) "Free access versus private property in a resource: income distributions compared" *Journal of Political Economy* **95** (6) 1317-1325.

Else, P.K. (1981) "A reformulation of the theory of optimal congestion taxes" *Journal of Transport Economics and Policy* **15** 217-232.

Else, P.K. (1982) "A reformulation of the theory of optimal congestion taxes: a rejoinder" *Journal of Transport Economics and Policy* **16** 299-304.

Evans, Alan W. (1992) "Road congestion: the diagrammatic analysis" *Journal of Political Economy* **100** (1) 211-217.

Evans, Andrew W. (1992) "Road congestion pricing: when is it a good policy?" *Journal of Transport Economics and Policy* **26** 213-243.

Evans, Andrew W. (1993) "Road congestion pricing: when is it a good policy?: a rejoinder" *Journal of Transport Economics and Policy* **27** 99-105.

Henderson J.V. (1974) "Road congestion: a reconsideration of pricing theory" *Journal of Urban Economics* **1** 346-365.

Henderson J.V. (1981) "The economics of staggered work hours" *Journal of Urban Economics* **9** 349-364.

Hills, P. (1993) "Road congestion pricing: when is it a good policy?: a comment" *Journal of Transport Economics and Policy* **27** 91-99.

Lighthill, M.H. and G.B. Whitham (1955) "On kinematic waves, II: a theory of traffic flow on long crowded roads" *Proceedings of the Royal Society (London)* **A229** 317-345.

Newell, G.F. (1988) "Traffic flow for the morning commute" *Transportation Science* **22** 47-58.

Verhoef, E.T. (1998) "Time, speeds flows and densities in static models of road traffic congestion and congestion pricing" Mimeo, Free University, Amsterdam.

Vickrey, W.S. (1969) "Congestion theory and transport investment" *American Economic Review* **59** (Papers and Proceedings) 251-260.

Yang, H. and H.J. Huang "Priciple of marginal-cost pricing: how does it work in a general road network?" *Transportation Research* **32A** (1) 45-54.

Zhang, X. and D.F. Jarrett (1997) "Stability analysis of the classical car-following model" *Transportation Research* **31B** (6) 441-462.