THE STABILITY OF DOWNTOWN PARKING AND TRAFFIC CONGESTION*

Richard Arnott†
University of California, Riverside

Eren Inci‡
Sabanci University

01 May 2010

Abstract

Consider a transport facility in steady state that is operating at maximum throughput. How does it respond to a once-and-for-all increase in demand? The trip price must increase to ration the increased demand, but how? These questions have been the subject of a debate in transport economic theory dating back to Walters' classic paper (1961). The current wisdom is that the facility continues to operate at full capacity, with travel at reduced velocity and/or increased queuing serving to increase the trip price. This paper analyzes the transient dynamics and stability of steady states for a spatially uniform road

*The authors would like to thank Albert Erkip, Thomas Holmes, Robin Lindsey, Kenneth Small, Mete Soner, Erik Verhoef, William Strange (*the editor*), the two anonymous referees, the participants at the 3^{rd} International Conference on Funding Transport Infrastructure and 10^{th} Journée Transport, the Macroeconomics, Real Estate, and Public Policy Workshop, the 55^{th} Annual North American Meetings of the Regional Science Association International, especially the discussant Jeffrey Lin, the Far East and South Asia Meeting of the Econometric Society (2009), and the 24^{th} Annual Congress of the European Economic Association, and seminar participants at Bilkent University, the GRIPS (Tokyo), Istanbul Technical University, and the University of Victoria for valuable comments. The authors would like to thank Elijah DePalma for his expert numerical analysis, reported in footnotes 22 and 23. Arnott would like to thank the University of California Transportation Center for funding support (004436-002).

†*Address*: Department of Economics, University of California, Riverside, 4106 Sproul Hall, Riverside, CA 92521-0427 USA. *E-mail address*: richard.arnott@ucr.edu

‡*Corresponding Author*. *Tel.*: 90-216-483-9340; *fax*: 90-216-483-9250. *Address*: Sabanci University - FASS, Orhanli / Tuzla 34956 Istanbul TURKEY. *E-mail address*: ereninci@sabanciuniv.edu.

network with on-street parking, and finds in this context that the increase in demand may cause operation at reduced throughput.

Keywords: traffic congestion, on-street parking, steady states, stability, cruising for parking

JEL Classification: R41, L91

1 Introduction

Consider a transport facility (with possibly multiple congestible elements) in steady state that is operating at maximum throughput. How does it respond to a once-and-for-all increase in demand? The trip price must increase to ration the increased demand, but how? These questions have been the subject of a debate in transport economic theory dating back to Walters' classic paper (1961). The current wisdom (Small and Verhoef, 2007, and Verhoef, 2005) is that the facility continues to operate at maximum throughput, with travel at reduced velocity and/or increased queuing serving to increase the trip price. This paper analyses the transient dynamics and stability of steady states for a spatially uniform road network with on-street parking, and finds in this context that the increase in demand may cause operation at reduced throughput. An analogy is the occurrence of brownouts and blackouts on overloaded electricity distribution networks. The issue is central to our understanding of heavily congested traffic. It also has important implications for the magnitude of the efficiency loss due to underpriced congestion and for congestion management policy.

Until recently transport economists answered the questions posed above by analyzing steady-state equilibrium on a single link with only crude treatments of transient dynamics. In a series of papers (Small and Chu, 2003, and Verhoef, 1999, 2001, 2003, and 2005), Small and Verhoef have raised the level of analysis, treating alternative concepts of equilibrium, extending the analysis to more realistic networks, and providing more sophisticated treatments of traffic dynamics. Verhoef (2001) studies a finite road of uniform width subject to a simple form of flow congestion (as described by a simple car-following model). With high demand, there is a unique steady state in which the road operates at maximum throughput and a vertical queue is present at the entry point, whose endogenous length serves to ration demand. Verhoef (2005) extends the analysis to a two-element network, a congestible road with a flow bottleneck at the exit point. High demand is rationed by increased travel time on the road rather than a vertical queue at the entry point, but the network continues to operate at maximum throughput. On the basis of these results and the analyses of their other papers, in their magisterial textbook (Small and Verhoef 2007) Small and Verhoef argue that, with high demand, operation at maximum throughout is characteristic of transport facilities.

This paper contributes to the debate by providing a (we believe) persuasive treatment of the transient dynamics and stability of steady states of a particular two-element transport facility – a spatially uniform downtown road network with on-street parking, as modeled in Arnott and Inci (2006) – and subject to a particular specification of demand. If the on-street parking capacity constraint binds, cruising for parking arises, which is essentially a random access queue that interferes with traffic flow. We first determine the steady states of the model. We then consider the model's transient dynamics from all feasible initial conditions when the demand function is stationary over time, which allows us to determine the stability of the various steady states. Finally, we explore the model's transient dynamics from one steady state to another in response to a once-and-for-all increase or decrease in demand.

We find among other things: (*i*) Gridlock is always a stable, steady-state equilibrium, and is the only stable, steady-state equilibrium when demand is very high. (*ii*) Except when demand is very high, there is another stable, steady-state equilibrium. The properties of this stable equilibrium depend on the demand intensity. With low demand intensity, parking is unsaturated (not fully occupied) and travel is congested (the normal traffic situation). With intermediate demand intensity, parking is saturated and travel is congested. With high demand intensity, parking is saturated, and travel is hypercongested (a traffic jam situation). With very high demand intensity, this stable equilibrium disappears and only the gridlock equilibrium remains. (*iii*) Depending on parameter values, there may be an interval of demand intensity over which the non-gridlock equilibrium has the comparative static property that an increase in demand intensity results in a fall in throughput. (*iv*) Even when steady-state demand intensity is not very high, a demand pulse may lead to a "catastrophic" transition to the gridlock equilibrium. (*v*) Except when demand is very high, there is a third equilibrium that is saddle-path stable.

Results (i), (iii), (iv), (v) and the last part of (ii) are inconsistent with Small and Verhoef's argument. That our model provides a counterexample to their argument raises doubts about the generality of their conclusion that, with high demand, operation at maximum throughput is characteristic of transport facilities. Since ours is a very particular model, we do not claim that its properties extend to other transport systems. We do conjecture however that, under conditions of high demand, increased demand leading to reduced throughput is a widespread phenomenon.

The paper is organized as follows. Section 2 provides a thorough review of the debate. Section 3 presents the structural model that is analyzed in the rest of the paper, and discusses how it differs from previous models and what is special about it that allows comprehensive analysis of its transient dynamics. Section 4 derives the steady-state equilibria of the model and explores their properties. Section 5 carries out the stability analysis followed by a discussion of the results, and Section 6 concludes. An appendix contains technical details.

2 A Review of the Debate

To understand why fully satisfactory answers to the questions posed at the beginning of the paper have eluded the experts for almost half a century requires a reasonably thorough review of the literature.

Imagine a homogeneous road between two locations with a constant flow, f, of cars entering it, traveling along it, and exiting it. And assume, as we do throughout the paper, in keeping with the classical treatment of flow congestion,[1] that both in and out of steady state there is a technological relationship between velocity, v, and density, V, with velocity being inversely related to density. For the sake of concreteness, we assume Greenshield's Relation (1935), which specifies a negative linear relationship between velocity and density:

$$v = v_f\left(1 - \frac{V}{V_j}\right) \quad \text{or} \quad V = V_j\left(1 - \frac{v}{v_f}\right) \quad , \tag{i}$$

where v_f is free-flow velocity and V_j is jam density.

[1] By "the classical treatment", we mean what is variously called the hydrodynamic model, kinematic wave theory, and the Lighthill-Whitham-Richards (LHR) model (see Daganzo, 1997).

The Fundamental Identity of Traffic Flow is that flow equals density times velocity

$$f = Vv \quad . \tag{ii}$$

Combining (i) and (ii) gives flow as a function of velocity:

$$f = \frac{V_j\,(v_f - v)\,v}{v_f} \quad , \tag{iii}$$

which is an inverted and translated parabola and is displayed in Figure 1.

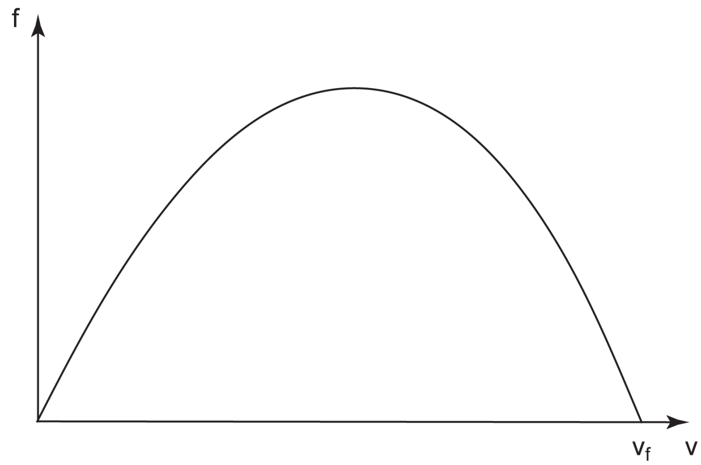


Figure 1: Flow as a function of velocity

Maximum flow is referred to as capacity (flow). There are two velocities associated with each level of flow below capacity flow. Following Vickrey, economists refer to travel at the higher velocity as *congested* traffic flow and travel at the lower velocity as *hypercongested* flow. Congested traffic flow is informally interpreted as smoothly flowing traffic and hypercongested traffic flow as a traffic jam situation.

Assume to simplify that the money costs of travel are zero and that the value of travel time is independent of traffic conditions and is the same for all cars. Then the user cost of a trip, c, which equals its price, is simply the value of travel time, ρ, times travel time, t, which is the inverse of velocity, times the length of the street, which we normalize to one, without loss of generality: $c = \rho t = \rho/v$ or $v = \rho/c$. Substituting this into (iii) gives the relationship between trip cost and flow:

$$f = \frac{V_j \left(v_f c - \rho\right) \rho}{v_f c^2} \quad . \tag{iv}$$

Figure 2 plots trip cost/price on the $y-$axis against flow on the $x-$axis. The upward-sloping portion of the curve corresponds to congested travel; the backward-bending portion corresponds to hypercongested travel. In the literature, this curve is referred to as the user cost curve or the supply curve of travel. The trip demand curve relates the (flow) demand for travel to trip price. Assume that no toll is applied, so that trip price equals user cost, and trip demand can be expressed as a function of user cost. Now draw in a linear trip demand curve that intersects the user cost curve three times, once on the upward-sloping portion and twice on the backward-bending portion of the user cost curve. The first intersection point is a congested equilibrium, the latter two are hypercongested equilibria. Label the three equilibria e_1, e_2, and e_3.

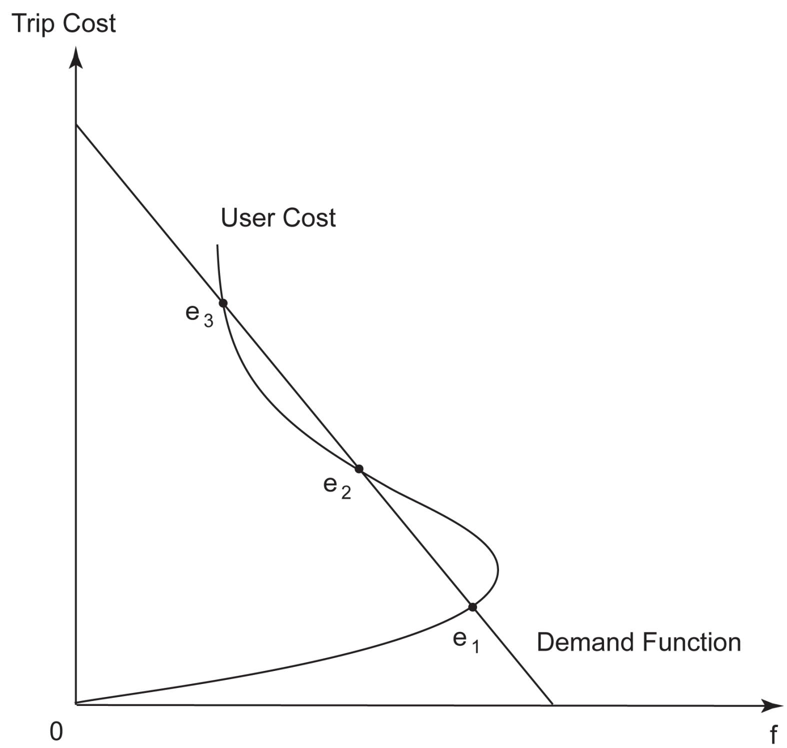


Figure 2: Stability of equilibria

The issue that has been much debated concerns the stability of equilibria on the backward-bending portion of the user cost curve – in terms of Figure 2, e_2 and e_3. Suppose, for the sake of argument, that an equilibrium is defined to be stable if, when a single car is added to or subtracted from the entry flow (a demand perturbation), the steady-state traffic flow returns to that equilibrium's level. Even if the traffic inflow rate, apart from the added car, is held constant, solving for the transient dynamics of traffic flow using the classical model is very difficult. But perhaps one should also take into account that the added car will affect traffic flow, hence user cost/trip price, and hence the traffic inflow rate, in the future, which makes the analysis even more difficult. To circumvent this complexity, Else (1981) and Nash (1982), viewing equilibrium as the intersection of demand and supply curves, apply conventional economic adjustment dynamics without reference to the physics of traffic flow. Assuming a density/price perturbation and adjustment via flows/quantities (akin to Walrasian price dynamics), Else argues that e_3 is locally stable. Assuming instead a flow/quantity perturbation and adjustment via densities/prices (Marshallian dynamics), Nash argues that e_3 is locally unstable.[2]

There is now broad agreement that this stability issue cannot be resolved without dealing explicitly with the dynamics of traffic flow. Unfortunately, providing a complete solution even for traffic flow on a uniform point-input, point-output road with an exogenous inflow function is formidably difficult.[3] The literature has responded in four qualitatively different ways to the intractability of obtaining complete solutions to this class of problems:

1. Derive qualitative solution properties, while fully respecting the physics of traffic flow.[4] This approach is the ideal but is mathematically demanding.

[2] Applying Else's analysis to e_2 leads to the conclusion that it is locally unstable, applying Nash's that it is locally stable. Applying either analysis to e_1 leads to the conclusion that it is locally stable.

[3] One inserts an equation relating velocity to density – we assume Greenshield's Relation – into the equation of continuity (the continuous version of the conservation of mass), which yields a first-order partial differential equation. Applying the appropriate boundary conditions, one can in principle solve for density as a function of time and location along the road. Unfortunately, the partial differential equation does not have a closed-form solution for any sensible equation relating velocity and density, and derivation of even the qualitative properties of equilibrium is difficult.

[4] Lindsey (1980) considers an infinite road of uniform width subject to classical flow congestion, with no cars entering or leaving the road, and proves that, if there is hypercongestion at no point along the road at some initial time, then there will be hypercongestion at no point along the road in the future. Verhoef (1999) considers a finite road of uniform width with a single entry point and a single exit point, and argues (Prop. 2b) that if there is hypercongestion at no point along the road at some initial time, then there will no hypercongestion along the road in the future. Verhoef (2001) develops the argument further using a simplified variant of car-following theory in which drivers

2. Employ an assumption that simplifies the congestion technology, while continuing to treat location and time as continuous. One example is the "zero propagation" assumption that a car's travel time on the road depends only on either the entry rate to the road at the time the car enters the road (Henderson, 1981) or the exit rate from the road at the time the car exits the road (Chu, 1995). Another example is the "infinite propagation" assumption that the velocity of all cars on the road at a point in time depends on either the entry rate to the road or the exit rate from it (Agnew, 1977). None of these assumptions is consistent with classical flow theory. The question then arises as to whether the qualitative results of a model employing such assumptions are spurious.

3. Replace the partial differential equation with a discrete approximation – discretizing time and location – and then solve the resulting difference equation numerically. One such discrete approximation is Daganzo's cell transmission model (Daganzo, 1992). Again, there is the concern that such approximations may give rise to spurious solution properties.

4. Adopt an even simpler traffic geometry in which the road system is isotropic, so that the entry and exit rates, as well as travel velocity, density, and flow, are the same everywhere on the network. This eliminates the spatial dimension of congestion so that the partial differential equation reduces to an ordinary differential equation. The second model of Small and Chu (2003) adopts this simplification, as do we in this paper. Unlike the previous two approaches, this approach does not entail any dubious approximation, but one may reasonably question the generality of results derived from models of an isotropic network.

Whatever approach is adopted, the issue arises as to the appropriate concept of stability to apply. This paper considers only steady-state equilibrium, in which the inflow rate and traffic flow remain constant over time. The most familiar concept of stability is local stability. Start in a steady-state equilibrium. Perturb it, which implies an infinitesimal change. If the system always returns to that steady-state equilibrium, it is said to be locally stable with respect to the assumed adjustment dynamics. In their textbook discussion of the stability of steady-state equilibrium, Small and Verhoef employ a different concept of stability – dynamic stability. They define a steady state to be dynamically stable if it can arise as the end state following

control their velocities directly.

some transitional phase initiated by a once-and-for-all change to a *constant inflow rate*. We employ a similar definition of stability, but with what will turn out to be an important difference. We specify the adjustment dynamics so that the inflow rate is responsive to trip price, and hence define a steady state to be dynamically stable if it can arise as the end state following some transitional stage initiated by a once-and-for-all change to a *stationary demand function*.

In their textbook, Small and Verhoef argue that, when steady-state demand for the road is so high that its use cannot be rationed through congested travel, equilibrium exists, is unique and dynamically stable (according to their definition), and entails a steady-state queue or quasi-queue whose length adjusts to clear the market, with the road operating at full capacity. In line with this argument, they replace the backward-bending portion of the user cost curve with a vertical segment at capacity flow. They base their textbook argument on the analysis of a variety of different models presented in several papers, which we now review.

Small and Chu (2003) considers two network geometries, one a uniform highway with a downstream bottleneck of fixed flow capacity, the other an isotropic network of downtown streets. For each network geometry, they examine first traffic flow with an exogenous demand spike and then an endogenous scheduling equilibrium. For both network geometries, the demand spike analysis shows that hypercongestion can occur as a transient phenomenon. In the endogenous scheduling model, a fixed number of identical commuters with a common origin, a common destination, and a common desired arrival time each decides when to depart. Travel at the peak of the rush hour is more congested (higher travel time cost) but entails arrival at a more convenient time (lower schedule delay cost). In the endogenous scheduling equilibrium (first introduced by Vickrey, 1969, in the bottleneck model), the time pattern of departures is such that trip cost, the sum of travel time cost and schedule delay cost, is equalized over the rush hour. The reduced-form supply curve relates equilibrium trip cost to the number of commuters. The main result for both network geometries is that, while hypercongested travel may occur for a portion of the rush hour, the reduced-form supply curve is upward sloping.

Verhoef (1999) defines a steady-state equilibrium to be dynamically stable if there exists a constant inflow rate such that it can be reached as the end point starting from some other steady-state equilibrium (with a different constant inflow rate). The paper argues that, on a road of uniform width, hypercongested equilibria are dynamically unstable, and that high demand is rationed through the formation of a steady-state queue at the entry point. Verhoef (2001) formalizes the argument presented in Verhoef (1999) assuming that a car's velocity is determined by a simple car-following rule. Verhoef (2003) considers the endogenous scheduling equilibrium[5] for a two-link serial network in which the upstream link has a greater capacity than the downstream link. Like Small and Chu (2003), a fixed number of identical commuters with a common desired arrival time is assumed. The paper shows that travel on the upstream link may be hypercongested over an interval of the rush hour, while travel on the downstream link asymptotically approaches capacity flow from below. Essentially hypercongestion on the upstream link takes the place of the queue in the bottleneck model, and may therefore be referred to as a quasi-queue. Verhoef (2005) examines steady-state equilibria in the same two-link serial network as Verhoef (2003), and concludes that, while hypercongestion can occur on the upstream link, flow on the downstream link is always at capacity.

As well as presenting our model and analysis, in this paper we shall attempt to identify why our results concerning the stability of steady-state equilibria are so at variance with the conclusions of Small and Verhoef's textbook argument. Keeping track of the diverse models in Small and Chu (2003) and Verhoef (1999, 2001, 2003, and 2005) is difficult. Fortunately, Verhoef's papers follow a logical progression, while Small and Chu (2003) does not address the stability of steady-state equilibria. Since Small and Verhoef's textbook argument is fully consistent with the analysis in Verhoef (2005), we shall compare our model and analysis with those of Verhoef (2005).

While there are other differences, we shall argue that the divergent conclusions derive from different adjustment dynamics. In particular, in his stability analysis Verhoef (2005) assumes a constant inflow rate (which would be appropriate with perfectly inelastic demand), whereas we assume a stationary demand function, with the inflow rate depending on the trip price. To illustrate the importance of this distinction, consider gridlock, which we find to be a stable equilibrium but which Verhoef's paper does not mention. In determining whether gridlock is a stable equilibrium according to Verhoef's stability criterion, one would proceed as follows. Starting from any

[5] Verhoef refers to endogenous scheduling equilibria as dynamic equilibria and to steady-state equilibria.

steady-state equilibrium other than gridlock, hold the inflow rate constant at zero. Obviously the traffic system will move to a situation of no traffic. Thus, according to Verhoef's criterion, gridlock is not a dynamically stable equilibrium. According to our stability criterion, in contrast, gridlock would be a stable equilibrium if, given the stationary demand function, there exists a feasible initial traffic state such the traffic system becomes (and stays) gridlocked.

Imagine that the initial traffic state is a traffic jam that is almost gridlocked (generated perhaps by a traffic accident), with a trip price such that the entry flow exceeds the exit flow. The traffic jam will get worse, resulting in both a decrease in the exit flow and an increase in the trip price, and the increase in trip price will in turn lead to a decrease in the entry flow. Whether the entry flow will continue to exceed the exit flow depends on the congestion technology and the form of the demand function, but if it does gridlock is eventually reached, at which point both the entry and exit flows equal zero. This line of reasoning establishes the plausibility of a stable gridlock equilibrium but does not prove it. We now turn to our model and analysis, which will prove the assertions stated in the introduction, including the existence and stability of the gridlock equilibrium.

3 Model Description

The model is aimed at describing downtown traffic and its interaction with on-street parking. Two parking régimes are considered. In the *saturated* parking régime, all on-street parking spaces are occupied, cars are cruising for parking, and as soon as a parking space is vacated it is taken by a car cruising for parking. In the *unsaturated* parking régime, there are vacant on-street parking spaces, and cars spend no time cruising for parking.[6] A detailed description of a slightly different version of the model, which focuses only on the steady-state equilibrium under saturated parking conditions and does not consider its stability, can be found in Arnott and Inci (2006).

[6] Arnott and Rowse (1999) provides a more sophisticated treatment of unsaturated parking in which cruising for parking occurs. In contrast to the model of this paper, their city is located on an annulus. On the basis of the parking occupancy rate, a driver decides how far from his destination to start cruising for parking, takes the first available vacant space, and walks from there to his destination. Adapting this more sophisticated treatment of unsaturated parking here should not substantially alter our results.

The focus here is on the transient dynamics of the variant of the model considered in this paper, especially the stability of equilibria, taking into account transitions between saturated and unsaturated parking conditions.

The downtown area has an isotropic (spatially homogeneous) network of streets.[7] For concreteness, one can imagine a Manhattan network of one-way streets. We assume that all travel is by car and that all parking is on street.[8] Each driver enters the downtown area, drives to his destination, parks there immediately if a vacant parking space is available and otherwise circles the block until a parking space becomes available, visits his destination for an exogenous length of time, and then exits the downtown area.[9] Drivers differ in driving distance and visit length. Driving distance is Poisson distributed in the population with mean m, and visit length is Poisson distributed with mean l.

Downtown parking spaces are continuously provided over the space. There may be three kinds of cars on streets: cars in transit, cars cruising for parking, and cars parked. Apart from the street architecture (the street layout and the proportion of curbside allocated to parking), travel velocity depends on the density of cars in transit and cars cruising for parking, with a car cruising for parking contributing at least as much to congestion as a car in transit.

Let T be the pool of cars in transit per unit area, C be the pool of cars cruising for parking per unit area, and P be the pool of on-street parking spaces per unit area (which is held constant throughout the paper). The traffic technology is defined

[7] In unpublished work, Vickrey referred to isotropic models as "bathtub models". The density of traffic is analogous to the height of water in the bath, and remains the same if the water flowing from the tap (the entry rate of cars) equals the water flowing from the drain (the exit rate of cars). The model here can be interpreted as a bathtub model.

One may reasonably object to the assumption that the network is isotropic if cars are entering from outside the downtown area, since edge effects are then present. There are two ways of dealing with this objection. The first is to assume that the city is located on the outside of a sphere, and that entering cars are randomly parachuted in. The second is to assume that everyone lives in the downtown area and parks in his private off-street garage. A driver then exits his private garage, drives to his destination, parks there on street, and at the end of his visit returns to his private garage.

[8] Arnott and Rowse (2009) extends Arnott and Inci (2006) to allow for both on- and off-street parking.

[9] One might object to the assumption that upon completion of a visit, a car just exits the downtown area. If we had assumed instead that, upon completion of a visit, a car returns to the same point at which it entered the downtown area and exits there, the steady-state equilibrium conditions would be unchanged but the transient dynamics would be different.

by an in-transit travel time function $t(T, C, P)$, where t is per unit distance.[10] Let $P_{\max}$ be the maximum possible number of on-street parking spaces per unit area. We assume that the technology satisfies $t_T > 0$, $t_C > 0$, $t_P > 0$, $t(0, 0, P) > 0$, $\lim_{P \to P_{\max}} t(T, C, P) = \infty$, and t is convex in T, C, and P.

This is a convenient point at which to introduce a distinction that will prove important in the subsequent analysis. We define *flow* per unit area to be $(T+C)/(mt)$ and *throughput* per unit area to be $T/(mt)$.[11] Since cars cruising for parking just circle the block, they contribute to flow but not to throughput.

Denoting the rate of entry into the network per unit area-time by Δ and the exit rate from the pool of cars in transit by E, we can write the rate of change in the pool of cars in transit as follows:

$$\dot{T}(u) = \Delta(u) - E(u) \quad , \tag{1}$$

where u is the time. This trivially describes the evolution of the pool of cars in transit at every instant. Describing the evolution of downtown parking is less trivial. As noted earlier, there are two parking régimes, and along the path of adjustment from the initial condition to a steady state the possibility of a switch from one to the other must be accounted for.

In the first régime, downtown parking is *saturated*, meaning that a vacant parking space is immediately taken by a car cruising for parking. In this régime, all parking spaces are filled at any given time but the pool of cars cruising for parking evolves over time. When parking is saturated, which we term régime 1, the rate of change in the pool of cars cruising for parking is simply the difference between the entry rate into the pool of cars cruising for parking and the exit rate from it, or simply

$$\dot{C}(u) = E(u) - Z(u) \quad , \tag{2}$$

[10] Note that we assume that P enters the in-transit travel time function even when parking is unsaturated. The rationale is that even one car parked curbside on a city block precludes the use of that lane for traffic flow over the entire block.

[11] We define steady-state throughput per unit area to be the entry rate per unit area or the exit rate per unit area. The exit rate per unit area equals the rate at which cars exit the stock of cars in transit per unit area, which equals the stock of cars in transit per unit area divided by the time each car spends in transit. We define flow per unit area as $(C + T)/T$ times throughput per unit area, so defined.

where E now denotes the entry rate into the pool of cars cruising for parking, which equals the exit rate from the in-transit pool, and Z the exit rate from the pool of cars cruising for parking. In this régime, the pool of occupied parking spaces, S, remains fixed at $S = P$ (so that $\dot{S} = 0$), but the pool of cars cruising for parking evolves.

In the second régime, parking is *unsaturated*, meaning that there are empty parking spaces so that cars in transit can find a parking space upon arrival at their destinations. In this régime, the stock of cars cruising for parking is zero (so that trivially $\dot{C} = 0$, too) but the pool of occupied parking spaces evolves. The evolution of S is given by

$$\dot{S}(u) = E(u) - X(u) \quad , \tag{3}$$

where E is now the entry rate into the pool of occupied parking spaces, which equals the exit rate from the in-transit pool, and X the exit rate from the pool of occupied parking spaces.

We assume that the (flow) demand function for trips is stationary, with the quantity of trips demanded at time u depending on the common *perceived* full trip price at time u, $F(u)$. We also assume that the perceived full trip time at time u depends on traffic conditions at time u and on mean trip length and visit duration. The perceived trip price equals the perceived in-transit travel time cost plus the perceived cruising-for-parking time cost plus the perceived cost of on-street parking. The perceived in-transit travel time cost at time u is calculated as the value of time, ρ, times perceived in-transit travel time, which equals the time to traverse m miles at the travel velocity at time u; the perceived cruising-for-parking time cost at time u equals the value of time times the expected cruising-for-parking time based on the stock of cars cruising for parking at time u;[12] and the perceived cost of on-street parking equals mean parking time times the per-unit-time parking fee of λ. Thus,[13]

$$F(u) = \rho \left(mt(T(u), C(u), P) + \frac{C(u)\, l}{P} \right) + \lambda l \quad . \tag{4}$$

We also assume that the demand function, $D(F)$ satisfies $D(0) = \infty$, $D(\infty) = 0$, and

[12] The number of parking spaces vacated per unit time divided by the number of cars cruising for parking, $(P/l)/C$, gives the probability that a person exits the cruising-for-parking pool per unit time. As a result, the expected time cruising for parking is Cl/P.

[13] One could define the full price of a trip to include the time cost of a visit, as is done in Arnott and Inci (2006). λ would then be defined as the time and money cost of a visit per unit time.

$D' < 0$. Since the entry rate at time u equals the quantity of trips demanded at time u, we have

$$\Delta(u) = D(F(u)) \quad . \tag{5}$$

The exit rate from the in-transit pool equals the stock of cars in the in-transit pool multiplied by the probability that a car will exit the in-transit pool per unit time:[14]

$$E(u) = \frac{T(u)}{mt(T(u), C(u), P)} \quad . \tag{6}$$

Due to the assumption that visit durations are generated by a Poisson process, the probability that an occupied parking space is vacated per unit time is $1/l$. Thus, when parking is saturated, the exit rate from the cruising-for-parking pool equals that probability multiplied by the number of parking spaces, P:

$$Z(u) = \frac{P}{l} \quad . \tag{7}$$

When parking is unsaturated, the exit rate from the pool of occupied parking spaces is defined similarly. X is the probability that a particular parking space is vacated, $1/l$, times the number of occupied parking spaces at that particular time, $S(u)$:

$$X(u) = \frac{S(u)}{l} \quad . \tag{8}$$

After substituting out the variables Δ, E, Z, and X, downtown traffic is characterized by the following autonomous differential equation system with two régimes.

$$\text{Régime 1} \quad : \quad \begin{cases} \dot{T}(u) = D\left(\rho\left(mt(T(u), C(u), P) + \frac{C(u)l}{P}\right) + \lambda l\right) - \frac{T(u)}{mt(T(u),C(u),P)} \\ \dot{C}(u) = \frac{T(u)}{mt(T(u),C(u),P)} - \frac{P}{l} \\ \dot{S}(u) = 0 \end{cases} \tag{9}$$

$$\text{Régime 2} \quad : \quad \begin{cases} \dot{T}(u) = D(\rho mt(T(u), 0, P) + \lambda l) - \frac{T(u)}{mt(T(u),0,P)} \\ \dot{C}(u) = 0 \\ \dot{S}(u) = \frac{T(u)}{mt(T(u),0,P)} - \frac{S(u)}{l} \quad . \end{cases} \tag{10}$$

[14] Appendix A.1 derives this equilibrium condition.

Remember also that $S(u) = P$ in régime 1 and $C(u) = 0$ in régime 2. In Section 4, we shall focus on these two régimes in turn. That the differential equation system is autonomous (since u does not appear as a separate argument on the right-hand sides) allows us to employ phase plane analysis to investigate the stability of the traffic system, converting what would otherwise be an essentially intractable problem into one that is straightforward to analyze.

To achieve "autonomy", we made three essential simplifying assumptions: *i*) trip length is Poisson distributed; *ii*) visit duration is Poisson distributed; and *iii*) the entry rate at time u is a function only of the state variables, C and T, at time u. The former Poisson assumption makes the exit rate of cars in transit at time u dependent on only the stock of cars in transit and cruising for parking at time u. The latter Poisson assumption makes the exit rate from the pool of parked cars at time u dependent only on the stock of parked cars at time u. The three assumptions together imply that the dynamics of the traffic system depend only on the system's state variables, T, C, and S, and not separately on time. Put alternatively, the history of the traffic system is fully captured by the values of the state variables.

None of these assumptions is realistic. The assumption that the perceived trip price depends only on current traffic conditions entails a form of myopic expectations. And the assumption that demand depends only on the means, and not other properties, of the trip length and visit duration distributions, is hard to rationalize.[15] Our justification for making these assumptions is that together generate a model that both allows a rigorous stability analysis and fully respects the physics of traffic flow. The model is however highly particular.

4 Analysis of Steady-state Equilibrium

In this section, we characterize the steady-state equilibria of the model and display them graphically. In any steady-state equilibrium, the entry rate into each pool equals the corresponding exit rate from it, so that the size of each pool is time invariant.

[15] One consistent but unrealistic rationale is that individuals do not know their trip lengths and visit durations when they make their trip decisions and are risk-neutral expected utility maximizers. Another is that individual demand functions sum to form an aggregate demand function with this property.

We have the following definitions:

Definition 1 (Saturated equilibrium) *A saturated steady-state equilibrium is a triple* $\{T, C, S\}$ *such that* $\dot{T}(u) = 0$, $\dot{C}(u) = 0$, $\dot{S}(u) = 0$, *and* $S = P$.

Definition 2 (Unsaturated equilibrium) *An unsaturated steady-state equilibrium is a triple* $\{T, C, S\}$ *such that* $\dot{T}(u) = 0$, $\dot{C}(u) = 0$, $\dot{S}(u) = 0$, *and* $C = 0$.

4.1 Régime 1: Saturated steady-state equilibria

We shall start by investigating the saturated steady-state equilibria of régime 1, for which the equations of motion are given in (9). There are cars cruising for parking in any traffic equilibrium in which parking is saturated. The parking spots are completely full at any given time and once a spot is vacated it is immediately filled by a car that is currently cruising for parking. We make two additional assumptions regarding the traffic technology and the street architecture. First, we assume that cars cruising for parking contribute to congestion at least as much as cars in transit.

Assumption 1 $t_C \geq t_T$.

Now define throughout capacity to be the maximum throughput consistent with the congestion technology, which is obtained when there are no cars cruising for parking. Throughput capacity equals $\max_T\{T \div (mt(T, 0, P)\}$. The second assumption is that throughput capacity exceed the exit rate from saturated parking, P/l, since otherwise parking would never be saturated in a steady-state equilibrium.

Assumption 2 $\max_T\{T/(mt(T, 0, P))\} > P/l$.

This assumption, along with the assumptions on t, implies that $T/(mt(T, 0, P)) = P/l$ has two roots. For the existence of a saturated steady-state equilibrium, the entry rate in the absence of cruising for parking must lie between these roots.[16] Arnott

[16] Too low an entry rate results in parking never being saturated; too high an entry rate results in the street system not being able to accommodate the demand, as a result of which no equilibrium exists.

and Inci (2006) proved that, when this condition, as well as Assumptions 1 and 2, hold, there is a unique saturated steady-state equilibrium. The unique equilibrium is characterized by two equations (in addition to $\dot{S}(u) = 0$), $\dot{T} = 0$ or

$$D(F) = \frac{T}{mt(T, C, P)} \quad , \tag{11}$$

where F is given in (4) and $\dot{C} = 0$ or

$$\frac{T}{mt(T, C, P)} = \frac{P}{l} \quad . \tag{12}$$

Under reasonable assumptions on the demand function and congestion function, a solution to (11) and (12) exists and is unique. A thorough analysis is provided in Arnott and Inci (2006). Here we present some intuition. To start, we substitute (12) into (11) to obtain the alternative pair of equations, $D(F) = P/l$ and (12). $D(F) = P/l$ indicates that, in steady-state equilibrium with saturated parking, the full trip price must clear the market. The flow supply of trips equals the parking turnover rate, P/l. As long as the demand curve is downward sloping and intersects the supply curve, this full price is unique.

In $T - C$ space, the equilibrium full price line is negatively sloped since more cars in transit and more cars cruising for parking both increase the full trip price. Eq. (12) specifies that, in equilibrium, throughput must equal the parking turnover rate. Now imagine moving southeast along the equilibrium full price line. Where the equilibrium full price line intersects the $C-$axis, throughput is zero. If throughput increases continuously with movement southeast along the equilibrium full price line, and if throughput exceeds P/l where the equilibrium full price line intersects the $T-$axis, then there is a unique (T, C) for which full trip price clears the market and for which throughput equals the parking turnover rate. The former condition is satisfied if a car cruising for parking contributes at least as much to traffic congestion as a car in transit, and the latter if, in addition, demand is not too high relative to downtown street capacity.

Figure 3 draws these equations in $T - C$ space with reasonable functional specifications taken from Arnott and Inci (2006) that we specify below. Note that the $\dot{T} = 0$ locus includes the jam density line since with jam density, F and t are infinite and D is zero.

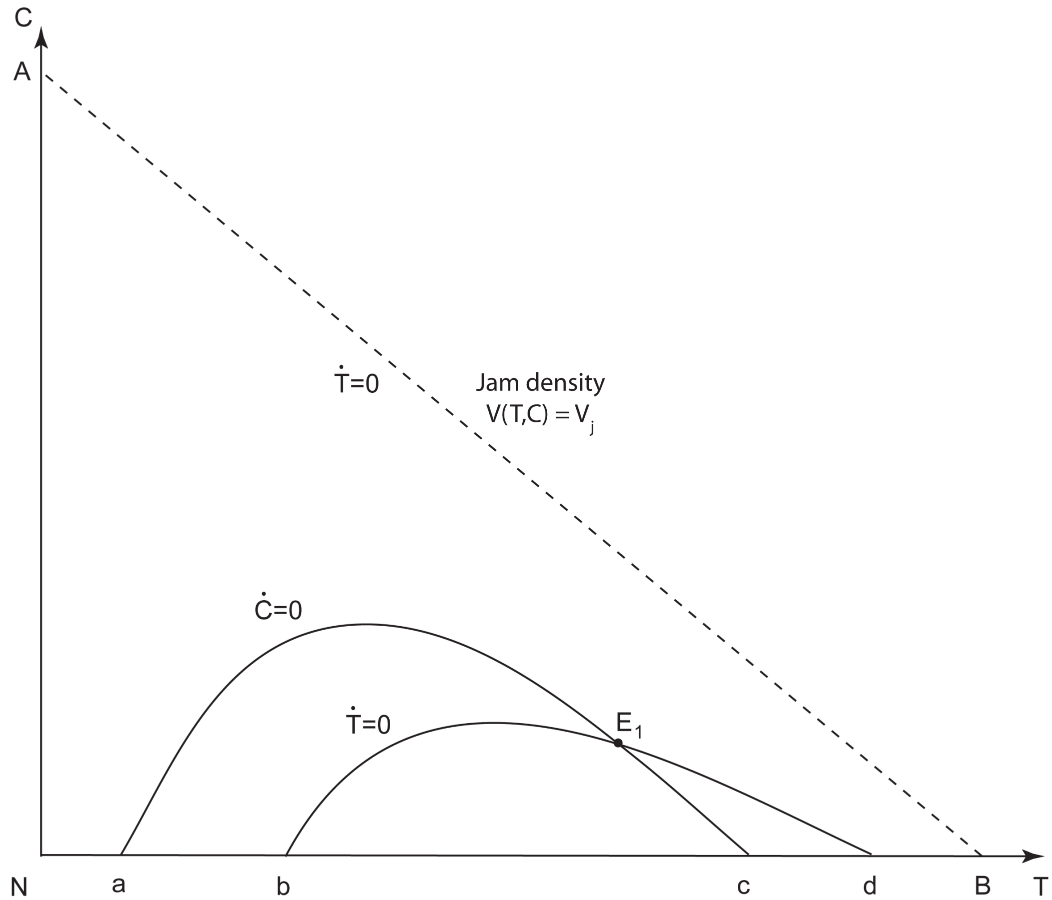


Figure 3: Saturated, steady-state equilibrium in $T - C$ space

Suppose that travel time t is weakly separable between (T, C) and P; refer to the sub-function $V(T, C)$ as the *effective density function*, and V_j as *effective jam density*. As usual, suppose also that t depends on the ratio of effective density and capacity, so that $t = t(V(T, C)/V_j)$. We measure effective density in terms of in-transit car equivalents, and assume it to take the following form:

$$V(T, C) = T + \theta C; \qquad \theta \geq 1 \quad , \tag{13}$$

so that a car cruising for parking contributes θ times as much to congestion as a car in transit.

Finally, we assume that Greenshield's Relation (1935) holds, so that the velocity of cars is a decreasing linear function of effective density. We therefore have

$$t = \frac{t_0}{1 - \frac{V(T,C)}{V_j}} \quad , \tag{14}$$

where t_0 is free-flow travel time. We also assume that demand is iso-elastic so that

$$D\left(F\right) = D_0 F^a \quad , \tag{15}$$

where $D_0 > 0$ is *demand intensity* and $a < 0$ the constant elasticity of demand.

Given these assumptions, as shown in Figure 3, the implicit function $C(T)$ defined by $\dot{C}(u) = 0$ (see (12)) is a concave function having two roots at $C = 0$, both of which are greater than zero and less than V_j. The $\dot{T} = 0$ locus (see (11)) has two parts. The first intersects $C = 0$ potentially multiple times between zero and less than V_j. The second is the jam density line. As already noted, if the $\dot{C}(u) = 0$ and $\dot{T}(u) = 0$ loci intersect they do so once, establishing the unique saturated equilibrium, E_1, shown in Figure 3. In section 4.3, we shall define congestion and hypercongestion. According to the definitions there, whether E_1 is congested or hypercongested depends on where the $\dot{C} = 0$ and $\dot{T} = 0$ loci intersect. The "qualitative" curvature of the figures in this paper can be obtained with the following parametric specifications: $m = 2$ miles, $l = 2$ hours, $\rho = \$20$ per hour, $t_0 = 0.05$ hours per mile, $P = 3712$ parking spaces per square mile, $V_j = 1778.17$ per square mile, $\lambda = \$1$ per hour, $D_0 = 3190.04$, $\theta = 1.5$, and $a = -0.2$.

For future reference, note that the $\dot{C} = 0$ locus cuts the $T-$axis at points a and c, and that with the assumed functional forms, the $\dot{T}(u) = 0$ locus cuts the $T-$axis three times, at points b, d and B. There can be no equilibrium above the jam density line AB ($T + \theta C = V_j$). Hence, the relevant subspace for the analysis of saturated equilibria is inside the triangle ANB (where N is the point where $C = 0$, $T = 0$[17]). With the assumed functional forms and parameters, we shall see that, in addition to the saturated equilibrium E_1, there are two unsaturated equilibria, one of which corresponds to gridlock. If the amount of on-street parking is increased sufficiently, there are three unsaturated equilibria.

[17] Later we work in (T, C, S) space, for which the origin is $(0, 0, 0)$. We do not refer to the point N as the origin since its coordinates in this space are $(0, 0, P)$.

4.2 Régime 2: Unsaturated steady-state equilibria

Unsaturated equilibria correspond to régime 2 whose equation system is given in (10). The stock of cars cruising for parking is zero so that a driver finds a parking space immediately upon reaching his destination. The stock of occupied parking spaces adjusts until the system reaches a steady state. Apart from $C(u) = 0$ (and $\dot{C}(u) = 0$), two equations characterize an unsaturated steady-state equilibrium. The first is again that the entry rate into the in-transit pool equals the exit rate from it:

$$D(F) = \frac{T}{mt(T, 0, P)} \quad , \tag{16}$$

where F is given in (4). The second is that the entry rate into the pool of occupied parking spaces equals the exit rate from it.

$$\frac{T}{mt(T, 0, P)} = \frac{S}{l} \quad . \tag{17}$$

Eqs. (16) and (17) represent $\dot{T}(u) = 0$ and $\dot{C}(u) = 0$, respectively.

Figure 4 draws these equations in $T - S$ space with the functional specifications indicated above. Eq. (16) is the same as (11) with $C = 0$. Thus, in $T - S$ space one part of the $\dot{T} = 0$ locus is vertical at the T coordinates corresponding to the points b and d, the other part is vertical at jam density. Eq. (17) has an inverted U-shape, passes through the origin and $(V_j, 0)$, and intersects $S = P$ at the points a and c, which are the same as the points a and c in Figure 3 where the $\dot{C} = 0$ locus intersects $C = 0$. Thus, each of the vertical lines associated with $\dot{T} = 0$ intersects $\dot{S} = 0$ exactly once, leading to three potential unsaturated equilibria, E_2, E_3, and E_4.

Above $S = P$, parking becomes saturated so that (17) ceases to apply.[18] This is indicated in the diagram by the dashes along $\dot{S} = 0$ for $S > P$. Thus, the parking capacity constraint rules out E_4 as an equilibrium.[19] For future reference, the relevant subspace for our analysis in the $T - S$ plane is the rectangle $ONBV_j$.

Figure 5 displays the model's equilibria in a diagram similar to Figure 2, but mod-

[18] Appendix A.2 briefly discusses the case in which there is no parking capacity constraint.

[19] Imagine gradually increasing the amount of on-street parking. In terms of Figure 5, this corresponds to a rightward movement of the parking capacity constraint; in terms of Figure 4, to an upward movement of the $S = P$ line. Above some level of parking capacity, E_1 transitions into E_4.

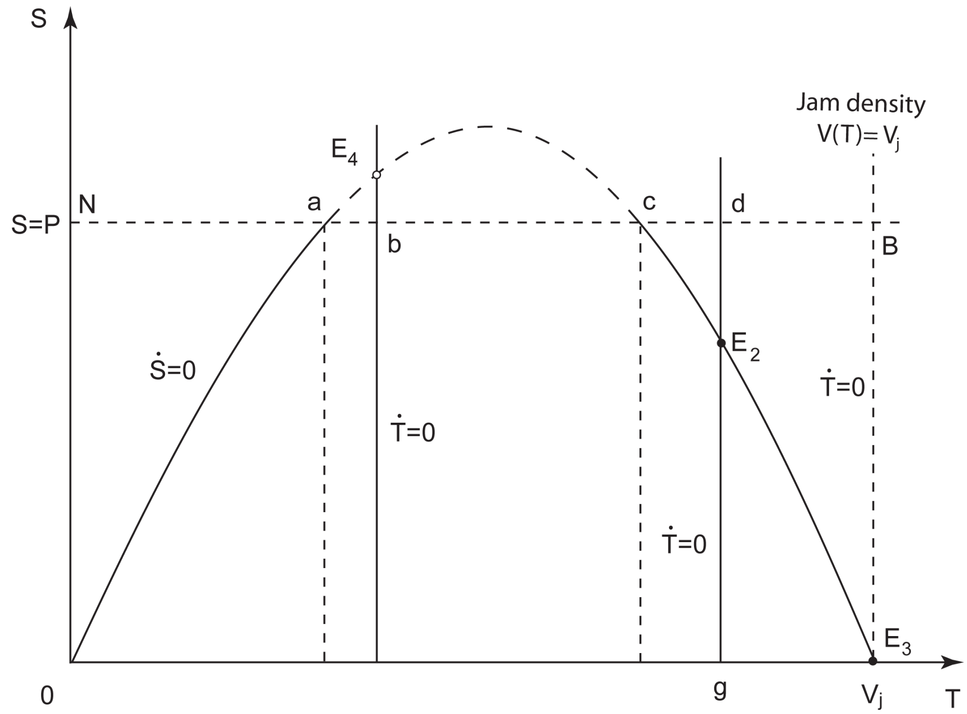


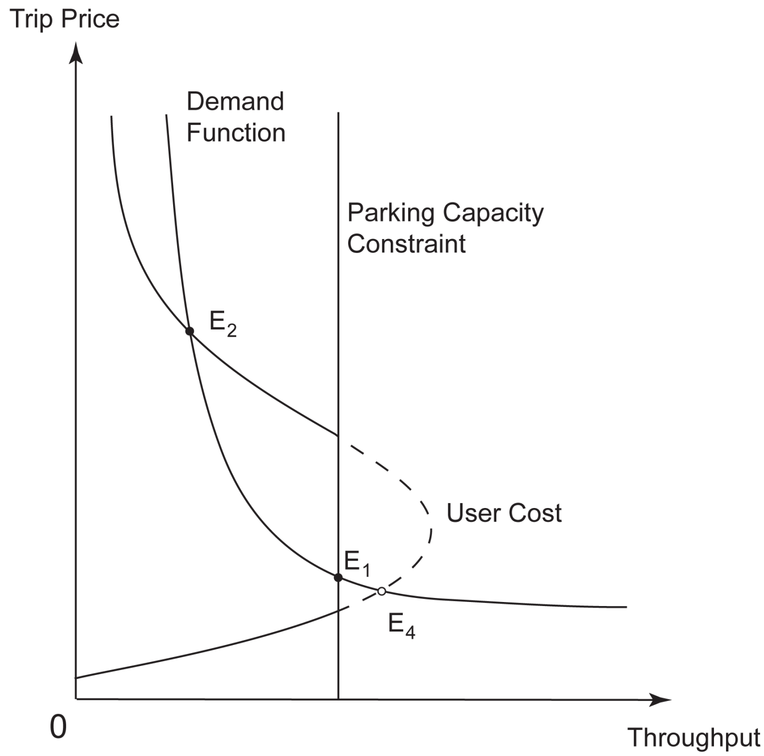
Figure 4: Unsaturated, steady-state equilibrium in $T-S$ space

ified by replacing flow with throughput and adding the parking capacity constraint (which by Assumption 2 is less than capacity throughput). The equilibrium E_3 is not shown since it corresponds to the intersection point of the demand function and the user cost function at zero throughput and infinite trip price. The figure also shows clearly why the parking capacity constraint rules out E_4 as an equilibrium.

In the next subsection we shall investigate whether traffic flow corresponding to each of these equilibria is congested or hypercongested, and in the next section the stability properties of the three equilibria.

4.3 Identifying hypercongestion

Recall that we have made a distinction between the (physical) density of traffic measured in cars per unit area, and the effective density of traffic measured in in-transit car-equivalents per unit area, which takes into account that a car cruising for parking generates at least as much congestion as a car in transit. The fundamental identity of traffic flow holds if flow and density are both defined in terms of physical cars. It



Note: There is another equilibrium, E_3, in which flow is zero and trip price is infinite.

Figure 5: Steady-state equilibria in throughput-trip price space

also holds if flow and density are both defined in terms of car equivalents. We have however chosen to work with effective density, since that is what traffic congestion is a function of, but to use the term flow to refer to the physical flow of cars, since that is what a bystander would observe. Thus, we must proceed with care.

We have assumed that traffic congestion is described by Greenshield's Relation, adapted to take into account cars cruising for parking. In particular, we have assumed eq. (i) in Section 2. We define congestion and hypercongestion in the following way:

Definition 3 (Congestion, hypercongestion) *Congestion occurs when traffic velocity is greater than that associated with capacity throughput, hypercongestion when traffic velocity is less than that associated with capacity throughput.*

Capacity throughput, which equals capacity flow, has been defined as $\max_T T/(mt(T,0,P))$. With Greenshield's Relation, capacity throughput equals $\max_T T(V_j - T)/(mt_0 V_j) =$

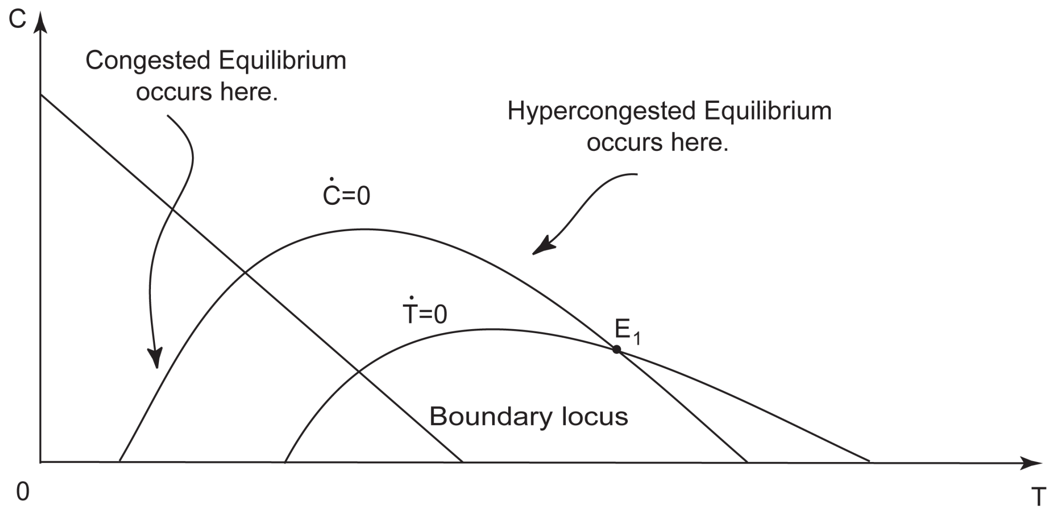
$V_j/(4mt_0)$, associated with which are in-transit traffic density $T^* = V_j/2$ and velocity $v^* = v_f/2$. Thus, we say that travel is congested when velocity is less than $v_f/2$ and hypercongested when velocity is greater than $v_f/2$. Since velocity and effective density are negatively related, we may equivalently define traffic to be hypercongested if $V(T,C) > V_j/2$, and congested when the inequality is reversed. For the particular effective density function we have assumed in (13), we obtain that travel is hypercongested if $T + \theta C > V_j/2$ and is congested otherwise.

We refer to the equation $T + \theta C = V_j/2$ as the *boundary locus*, since it separates the region of congested travel from the region of hypercongested travel. Figure 6 plots the boundary locus, as well as the $\dot{T} = 0$ and $\dot{C} = 0$ loci in $T - C$ space. The boundary locus has the same slope as the jam density line, $-1/\theta$. Travel below the locus is congested, and above the locus is hypercongested. We define equilibria to be congested or hypercongested accordingly. In particular:

Definition 4 (Congested equilibrium, hypercongested equilibrium) *An equilibrium is congested when congestion according to Definition 3 occurs, and hypercongested otherwise.*

As drawn in Figure 6, the equilibrium E_1 is hypercongested. Return to Figure 4. The peak of the $\dot{S} = 0$ locus corresponds to T^*. Travel on the left side of the peak is congested, and to the right side is hypercongested. Thus, travel is also hypercongested at both E_2 and E_3.

This – before starting the stability analysis – is a useful point to summarize our results. With the qualitative configuration of the phase plane we have derived, based on specific functional forms and parameters, we have identified three equilibria, E_1, E_2, and E_3. E_1 has saturated parking and may be either congested or hypercongested (in our example, it is hypercongested). E_2 has unsaturated parking and is hypercongested. E_3 has unsaturated parking and gridlock – the most extreme form of hypercongestion.



Note: The figure is drawn choosing parameters such that the saturated equilibrium is hypercongested. With a different choice of parameters the saturated equilibrium can instead be congested.

Figure 6: Identifying hypercongested travel in $T-C$ space

5 Stability

This section carries out the formal stability analysis by combining the two régimes followed by a discussion of the results.

5.1 Analysis

We start our analysis by stating our stability criteria.

Definition 5 (Stability) *(i) A steady-state equilibrium is said to be* ***locally stable***[20] *if it can be reached from all initial traffic conditions in its neighborhood; (ii) a steady-state equilibrium is said to be* ***saddle-path stable*** *if it can be reached only from initial traffic conditions on one of its arms; (iii) A steady-state equilibrium is said to be* ***dynamically stable*** *if it can be reached from at least one initial traffic condition other than itself.*

[20] This is sometimes called *asymptotically stable*.

Saddle-path stability and dynamic stability are both global concepts. Local stability is of course local. Both saddle-path stability and local stability imply dynamic stability.

For a complete stability analysis, we need to take into account not only transition between the two régimes but also the possibility that traffic might get stuck at jam density. Régime 1 (the saturated régime) is shown in Figure 3 in $T - C$ space, and régime 2 (the unsaturated régime) in Figure 4 in $T - S$ space. The two régimes may be analyzed simultaneously in the three-dimensional figure in $T - C - S$ space displayed in Figure 7. One should consider only points on the two illustrated planes and not any other points as initial traffic conditions. We shall explain how to read this figure before analyzing the stability of the equilibria.

The vertical $T - C$ plane reproduces Figure 3 with some added detail. The horizontal $T - S$ plane reproduces Figure 4 with some added detail. The fold where the two planes join is along $C = 0$ and $S = P$, with N representing the point $(T, C, S) = (0, 0, P)$ and B the point $(V_j, 0, P)$.

Consider the $T - C$ plane, which corresponds to the saturated régime whose dynamics are given in (9). The line AB corresponds to jam density. Since densities above jam density are infeasible, the feasible region of the plane is the triangle NAB. The $\dot{T} = 0$ and the $\dot{C} = 0$ loci divide the plane into four areas, labeled x_1, x_2, x_3, and x_4. Within a particular region, the direction of motion of C and T – shown by the arrows – is the same; for example, in region x_1, C is decreasing and T is increasing. The point M is the point on the jam density locus whose trajectory leads to the point d.

Consider the $T - S$ plane, which corresponds to the unsaturated régime whose dynamics are given in (10). The line BV_j corresponds to jam density. Since densities above jam density are infeasible, the feasible region of the plane is the rectangle $ONBV_j$. The $\dot{S} = 0$ locus and the three parts of the $\dot{T} = 0$ locus divide the plane into six areas, z_1, z_2, z_3, z_4, z_5, and z_6. Within a particular region, the direction of motion is the same; for example in region z_1, T is increasing and S is decreasing.

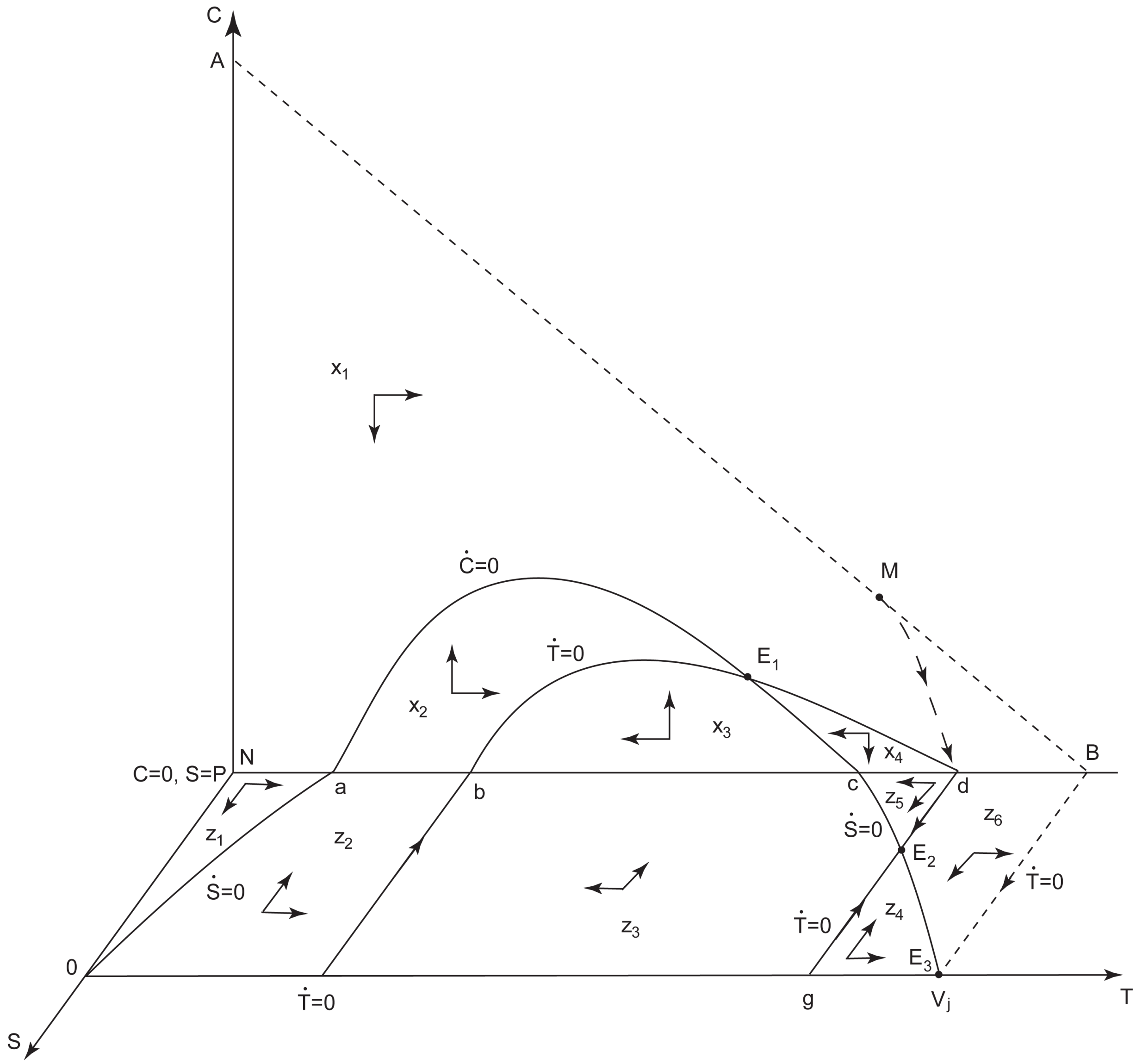


Figure 7: Saturated and unsaturated steady-state equilibrium in $T{-}C{-}S$ space

The direction of motion in the $T - C$ plane are obtained by combining (1), (2), (4), (5), and (6) to give

$$\dot{T}(u) = D\left(\rho\left(mt\left(T(u), C(u), P\right) + \frac{C(u)\,l}{P}\right) + \lambda l\right) - \frac{T(u)}{mt\left(T(u), C(u), P\right)} \quad (18)$$

$$\dot{C}(u) = \frac{T(u)}{mt\left(T(u), C(u), P\right)} - \frac{P}{l} \quad . \quad (19)$$

The direction of motion in the $T-S$ plane are obtained by combining (1), (3), (4), (6), and (8) to give

$$\dot{T}(u) = D\left(\rho\left(mt\left(T(u), C(u), P\right) + \frac{C(u)\,l}{P}\right) + \lambda l\right) - \frac{T(u)}{mt\left(T(u), C(u), P\right)} \quad (20)$$

$$\dot{S}(u) = \frac{T(u)}{mt\left(T(u), C(u), P\right)} - \frac{S(u)}{l} \quad . \quad (21)$$

Figure 7 is drawn using the parameters and functional forms (eqs. (13)-(15) given earlier) but modified for visual presentation.

In summary: x_i and z_j (where $i \in \{1, ..., 4\}$ and $j \in \{1, ..., 6\}$) denote areas of the phase plane; in each area the direction of the arrows indicates the direction of motion there; the point 0 is the origin of the 3D figure; the points a, b, c, and d are as defined before; the dotted lines indicate jam density situations; E_1 is the saturated equilibrium; E_2 is an unsaturated equilibrium; E_3 is another unsaturated equilibrium in which there is gridlock; as drawn, all of these equilibria are hypercongested.[21]

We shall now state the only proposition of the paper, from which we deduce our main findings.

Proposition 1 *Any starting point on the locus MdE_2g moves to E_2. Any starting point to the left of the locus moves to E_1, and any starting point to the right of the locus moves to E_3.*

Proof. We shall prove this proposition in three steps.

Step 1. Areas in the triangle ANB :

Area x_1 (excluding the adjustment path Md) : The direction of motion in this area is south east. Any initial condition in x_1 will either hit E_1 or E_3.

For sufficiently high values of C, the trajectories will reach the equilibrium E_1. They will approach the $\dot{C} = 0$ locus before reaching E_1 since the direction of motion right below the locus (in area x_2) is north east.

[21] One might argue the possibility of a limit cycle. However, it is ruled out by Bendixson's Nonexistence Criterion for the equations of motion of régime 1. Direction of motion shows that it cannot happen for the equations of motion of régime 2, either. We conjecture that there cannot be a limit cycle circling between the régimes. Figure 11 in Appendix A.3 displays sample trajectories for our numerical example.

The trajectories for low values of T and C will hit the line segment Na. Once they hit Na, C cannot further decrease since it cannot go below zero. Thus, parking becomes unsaturated and the direction of motion in area z_1 will apply. The trajectory will pass through $a0$ and enter the area z_2. The direction of motion in this area will then carry the trajectories toward the line segment ab either via area z_2 or via the $\dot{T} = 0$ locus in the $T - S$ plane. On the line segment ab, S cannot further decrease since it has to be nonnegative. Thus, parking becomes saturated again. Then, the direction of motion shown in x_2 will apply and therefore the trajectory will once again hit E_1.

For some intermediate values of T, the trajectories may hit the curve E_1d and pass through the area x_4. At this time, the trajectory may either hit E_1c and move into x_3 (and maybe x_2 after that) before reaching E_1 or it may hit the line segment cd. If it hits cd, parking will have to become unsaturated. Then, the trajectory moves into the area z_5 followed by area z_3. Once it is in z_3, the trajectory will move toward the line segment bc and then parking becomes saturated again before reaching E_1 from the area x_3.

Yet another possibility occurs for sufficiently large values of T and sufficiently small values of C. There has to be an initial condition M such that the trajectory initiated from M passes through the point where the $\dot{T} = 0$ locus cuts the NB line, namely point d. Given that trajectories in this differential equation system cannot intersect unless one is on the same trajectory as the other, for any initial point on the right hand side of the path Md, the trajectory will hit the line segment dB. Once it hits there, C cannot further decrease, parking becomes unsaturated, and the trajectory will move into the area z_6. Given the direction of motion there, it is then obvious that the trajectory will move towards E_3 to establish a gridlock of cars on the network of streets. The direction of motion in z_6 cannot carry a trajectory towards E_2.

Area x_2 : The direction of motion in this area is north east. Any trajectory from any initial condition in this area will trivially reach E_1. Parking never becomes unsaturated along the adjustment path as C will increase at all times.

Area x_3 : The direction of motion in this area is north west. There are two possibilities in this area. For lower values of T, the trajectories will enter x_2 (or move along the border of x_2 and x_3) before reaching E_1. For higher values of T, they hit the equilibrium E_1 from the area x_3. Parking never becomes unsaturated along the adjustment path as C will continuously increase until it reaches a steady-state equilibrium according to the direction of motion.

Area x_4 : The direction of motion in this area is south west. There are two possibilities in this region. First, the trajectory may hit E_1c and enter the area x_3 before hitting E_1. The other possibility is that the trajectory may hit the line segment cd. Once it hits cd, parking becomes unsaturated. The trajectory then moves into the area z_5 followed by area z_3 before reaching E_1, as previously explained. ■

Step 2. Areas in the rectangle $ONBV_j$:

Area z_1 : The direction of motion in this area is down east. Therefore, any trajectory in this area hits the curve $a0$ and passes into the area z_2. The direction of motion in this area will then carry the trajectories toward the line segment ab either via area z_2 or via the $\dot{T} = 0$ locus. On the line segment ab, S cannot increase further since it cannot exceed P. Thus, parking becomes saturated. Then, the direction of motion shown in x_2 will apply and therefore the trajectory will hit E_1, as previously explained.

Area z_2 : The direction of motion in this area is up east. They will carry the trajectories toward the line segment ab either via area z_2 or via the $\dot{T} = 0$ locus. On the line segment ab, S cannot further decrease since it has to be nonnegative. Thus, parking becomes saturated and the direction of motion of the area x_2 will apply. Consequently, the trajectory will reach E_1.

Area z_3 : The direction of motion in this area is up west. Any trajectory in this area will hit bc and reach E_1, as previously explained.

Area z_4 : The direction of motion in this area is up east. Any trajectory in this area will first hit E_2V_j and then follow this curve until it reaches the gridlock equilibrium E_3.

Area z_5 : The direction of motion in this area is down west. As previously explained, any trajectory here will first hit cE_2 and then enter into z_3 before reaching E_1, as previously explained.

Area z_6 : The direction of motion in this area is down east. Any trajectory here will either directly hit E_3 or follow BV_j before doing so. ■

Step 3. Points on the locus MdE_2g :

Points on the line segment E_2g : Since $\dot{T} = 0$ but $\dot{S} > 0$, any trajectory initiated on this line segment will follow the $\dot{T} = 0$ locus until it reaches E_2.

Points on the line segment E_2d : Since $\dot{T} = 0$ but $\dot{S} < 0$, any trajectory initiated on this line segment will follow the $\dot{T} = 0$ locus until it reaches E_2. This, along with other parts of the proof, implies that E_2 is a saddle point.

Points on the adjustment path Md : Any trajectory initiated from this curve will first hit d. However, since C cannot be negative, parking will become unsaturated. Thereafter, the trajectory will follow the $\dot{T} = 0$ locus until it reaches E_2, as previously explained. ■

This completes the proof. ■

There are two important corollaries to Proposition 1.

Corollary 1 *The hypercongested saturated equilibrium* E_1*, the hypercongested unsaturated equilibrium* E_2*, and the hypercongested gridlock equilibrium* E_3 *are all dynamically stable.*

This corollary follows directly from Proposition 1 and Definition 5. We should point out here that E_1 and E_3 are both locally stable equilibria. E_2 is not locally stable but is saddle-path stable.

Corollary 2 *There is no gridlock equilibrium with cruising for parking.*

The intuition for this result is straightforward. Start with a situation with traffic gridlock and cruising for parking. Since traffic is gridlocked, the exit rate from the in-transit pool and hence the entry rate into the cruising-for-parking pool is zero. Since parking is saturated, the exit rate from the cruising-for-parking pool is P/l. The cruising-for-parking pool therefore shrinks, so the initial situation cannot have been an equilibrium.

5.2 Discussion

Having completed our formal stability analysis, we shall now provide some intuition for the traffic system's dynamics. We first consider starting in area z_4 and investigate how traffic and parking adjust along the path to the gridlock equilibrium E_3. In area z_4, travel is so slow that the exit rate from the in-transit pool is lower than the inflow, so that the size of the in-transit pool increases. Since S is low, the exit rate from the in-transit pool is still larger than the rate at which parking is vacated, so that S increases.

Eventually, however, as travel gets slower and slower and the exit rate from the in-transit pool decreases, a point is reached where the exit rate from the parking pool equals the exit rate from the in-transit pool. As time proceeds, travel becomes even slower, the exit rate from the in-transit pool declines and falls short of the exit rate from the parking pool. The density of cars in transit continues to increase and the stock of parked cars decreases asymptotically towards the equilibrium E_3. Even though E_3 cannot be reached from the origin with a time-invariant demand function, a demand pulse may push traffic into the regions z_4 or z_6, or to the right of Md in the saturated régime, and once in those regions, with a time-invariant demand function, there is no way of escaping.

Another instructive exercise is to consider the adjustment dynamics in moving from inside the area dMB close to the jam density line to the gridlock equilibrium. This situation is interesting since traffic is initially almost gridlocked, then loosens up, and then becomes completely gridlocked. At the starting point, traffic is almost completely gridlocked and parking is saturated. Since traffic is almost completely gridlocked, the exit rate from the in-transit pool is lower than the exit rate from saturated parking. As a result, the stock of cars cruising for parking falls, and sufficiently rapidly that, even though the stock of cars in transit continues to rise, effective density falls – the traffic jam loosens.

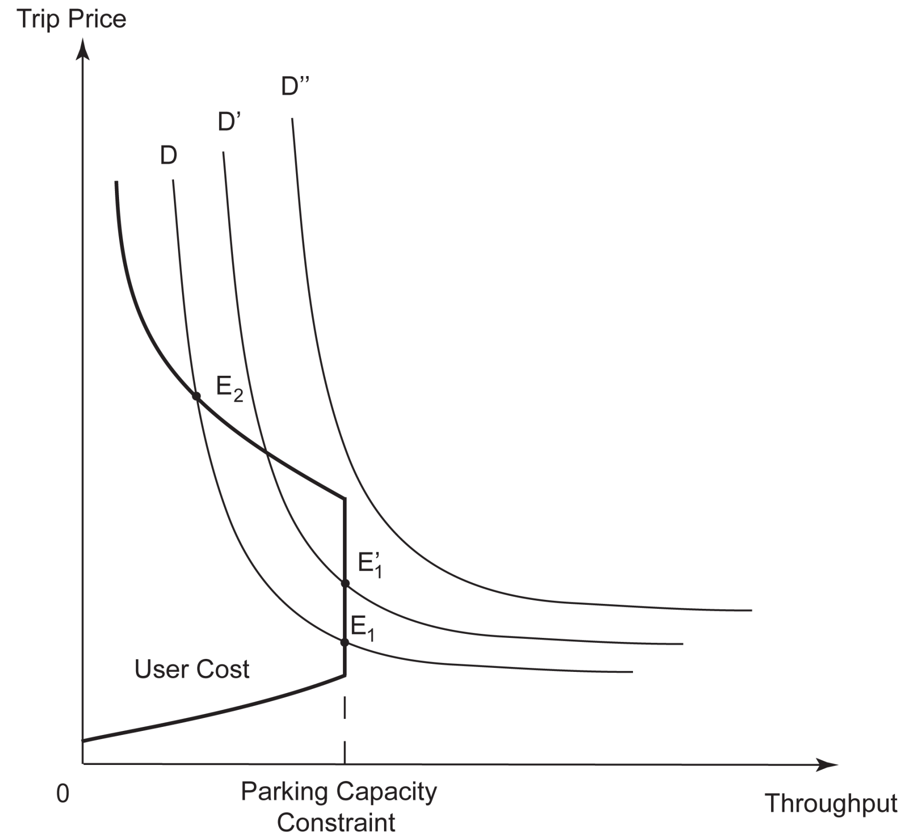
This process proceeds until the stock of cars cruising for parking reaches zero and parking becomes unsaturated. Traffic then moves into the area V_jE_2dB, where the number of occupied parking spaces falls (since the entry rate into parking continues to fall short of the exit rate), and where the stock of cars in transit continues to rise (since the entry rate into the in-transit pool exceeds the exit rate). Since there are now no cars cruising for parking, effective density rises. This process continues until the unsaturated gridlock equilibrium is reached.

Employing specific functional forms and parameter values, we have applied our analysis to examine the stability of steady-state equilibria. The analysis can also be applied to determine the comparative static properties of the set of steady-state equilibria. Return to Figure 7. Suppose that the traffic system is in steady-state equilibrium at E_1, and consider the effect of a moderate, once-and-for-all increase in travel demand. This results in a downward shift of the $\dot{T}=0$ locus, causing the corresponding equilibrium to relocate to a position on the $\dot{C}=0$ locus southeast of E_1 – call it E_1', for which T is higher and C lower. Since E_1 then lies in the interior of area x_1 to the left of the locus MdE_2g, the system moves directly from E_1 to E_1'.

Figure 8 displays the same result in throughput-trip-price space. Demand increases from D to D', which results in the saturated equilibrium moving from E_1 to E_1'. Parking remains saturated so that throughput remains unchanged. This requires that t increase, which requires that effective density increase. Since flow equals throughput times $(C+T)/T$, and since $(C+T)/T$ falls, flow decreases. Thus, the increase in demand results in reduced velocity and flow (so that velocity and flow move in the same direction, another indication of hypercongestion) and no change in throughput. Now consider the effect of a large, once-and-for-all increase in travel demand, that causes the $\dot{T}=0$ locus to move downward so far that no portion lies in the $T-C$ plane. Since the point E_1 is then located in the region x_1 to the right of the locus MdE_2g, the system moves from E_1 to the gridlock equilibrium (see the movement from D to D'' in Figure 8).

Figure 9 displays the bifurcation diagram of throughput plotted against demand intensity, D_0, with the assumed functional forms and parameters. The E_1 line corresponds to the interval of demand intensity where parking is saturated. The way Figures 8 and 9 are drawn is consistent with the assumed functional forms and parameter values.

Consider now raising parking capacity such that it ceases to bind (so that the analysis in Appendix A.2 applies). There is a level of demand intensity for which the demand curve is tangent to the backward-bending portion of the user cost curve. For somewhat lower levels of demand intensity, in the equilibrium E_4 (E_4 rather



Note: E_1 corresponds to E_1 in Figures 6 and 7, which is a saturated, hypercongested equilibrium. E_1' is the corresponding equilibrium when there is a moderate increase in demand. E_2 corresponds to E_2 in Figures 6 and 7.

Figure 8: The effects of an increase in demand when the initial steady-state equilibrium is saturated and hypercongested

than E_1 since the parking capacity constraint does not bind) flow and throughput are negatively related to demand intensity. Figure 10 displays the bifurcation diagram for this situation. We have confirmed numerically, for the numerical example of the paper (but with no parking constraint), that the equilibrium of type E_4 on the backward-bending portion of the user cost curve is stable under myopic expectations.[22]

[22] To further strengthen our argument, we also investigated its stability under perfect foresight. The way we did this was to adopt a simple learning process. Iteration (i) entails myopic expectations; each driver naively assumes that his velocity will be constant throughout his trip and equal to the velocity of traffic at the time he departs. Iteration (ii) takes the expected in-transit travel time of a driver departing at time u to be the realized travel time on iteration (i); and so on. That is, a driver departing at time u expects his in-transit travel time to be the in-transit travel time for a driver departing at time u on the previous day. In our numerical example, the learning process is

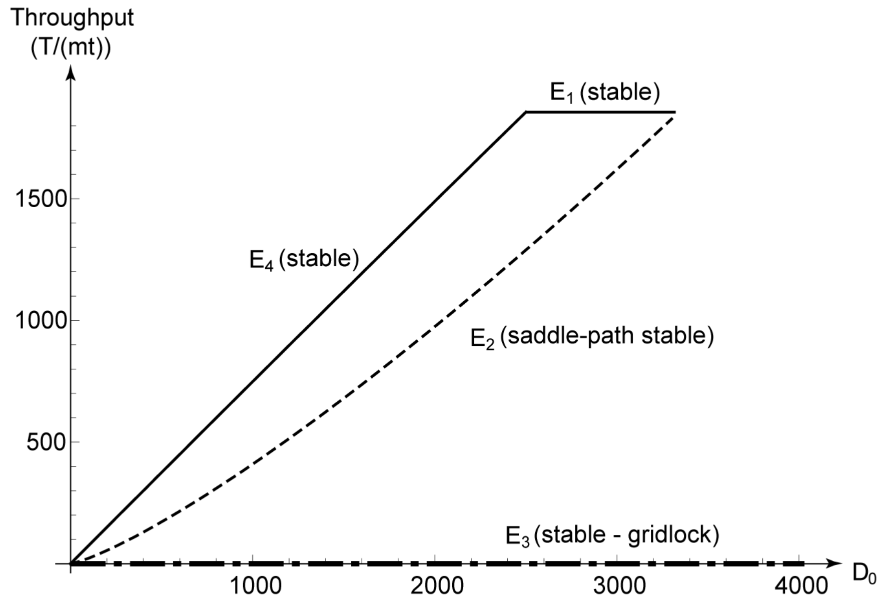


Figure 9: Bifurcation diagram: throughput against demand intensity (when the parking capacity constraint binds)

We have also proved that the gridlock equilibrium is stable under myopic expectations and confirmed this numerically for the example presented in the paper. Since traffic becomes increasingly congested as the gridlock equilibrium is approached, under myopic expectations drivers consistently underestimate in-transit travel time. Because entering drivers fail to take into account that traffic density increases over their route, some drivers will enter even though they will get gridlocked before reaching their destinations. Thus, gridlock occurs in finite time, but because parking duration is Poisson distributed, it takes an infinite time for the gridlock equilibrium, in which no parking spaces are occupied, to be reached.[23]

convergent, and converges to perfect foresight, so that the equilibrium of type E_4 is stable under perfect foresight. The numerical analysis is available on request.

[23] We conjecture that the gridlock equilibrium is stable under perfect foresight as well, but have been unable to come up with an algorithm to demonstrate this numerically. The problems that arise can be understood by considering why the algorithm employed to demonstrate stability under perfect foresight of the equilibrium of type E_4 on the backward-bending portion of the user cost curve does not work here (see footnote 22 for a brief description of the algorithm). On the first iteration, iteration (i), which corresponds to myopic expectations, drivers base their expected travel times on the density of traffic when they depart. Since traffic density increases over time as gridlock is being approached, realized in-transit travel times (for distance m) exceed the corresponding expected in-transit travel times, and traffic becomes gridlocked in finite time. Let u^{**} denote the time at which traffic becomes gridlocked and u^* $(< u^{**})$ denote the last entry time at which a car traveling distance

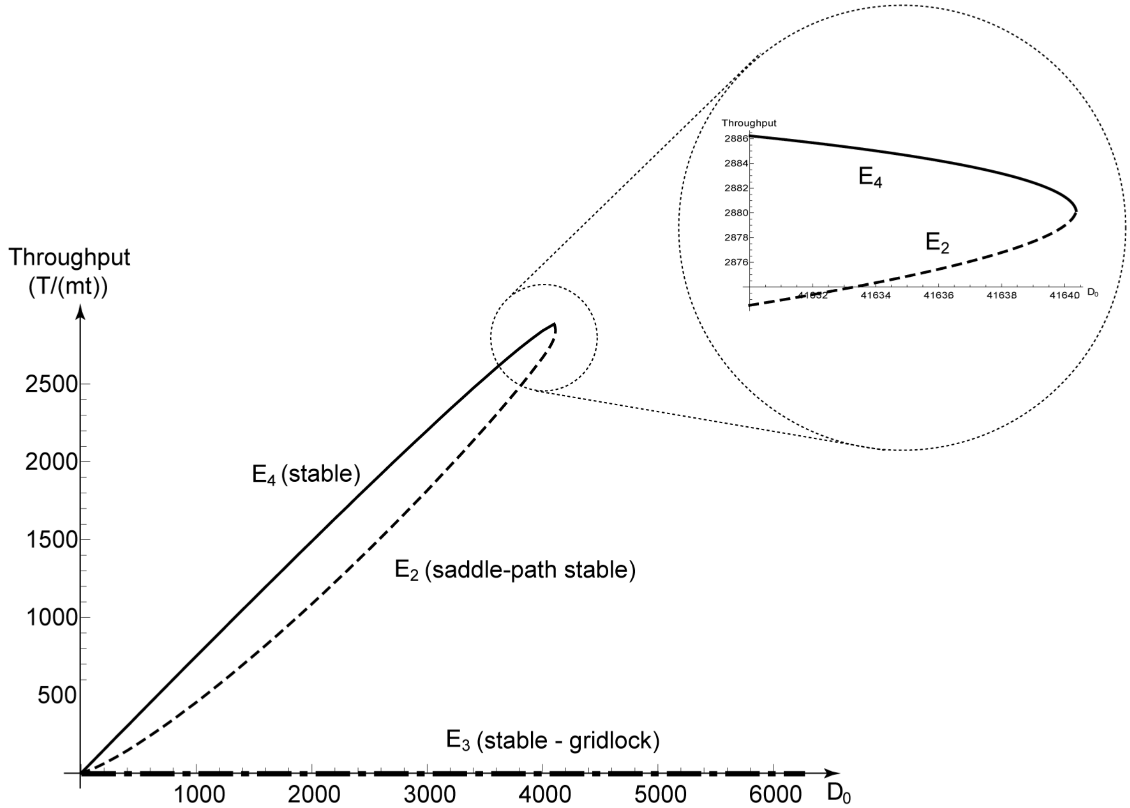


Figure 10: Bifurcation diagram: throughput against demand intensity (when the parking capacity constraint does not bind)

We now attempt to relate our method of stability analysis to those employed in other studies of the stability of hypercongested equilibria. We have applied global stability analysis, which examines where the traffic system will move to from any initial condition. Much of the analysis of the existence and stability of hypercongested equilibria has instead applied local stability analysis, which examines whether a traffic system that starts in an equilibrium will return to the same equilibrium after a small perturbation. Since the perturbation results in the starting point of the transient dynamics being close to an equilibrium, global analysis is more general.

---

m reaches its destination. On iteration (ii), the entry rate function for $u < u^*$ lies everywhere below that on iteration (i) and is zero beyond u^*. At time u^{**}, therefore, traffic is not gridlocked, and after u^{**} the entry rate is zero while cars continue to exit, so that density falls to zero. On iteration (iii), the entry rate function is high since the realized travel time function on iteration (ii) is low, and, as on iteration (i), traffic becomes gridlocked in finite time. Even when expectations are updated only very slowly, the learning process does not converge. We have tried several other numerical approaches to solving for the perfect foresight gridlock equilibrium trajectory, but without success. Thus, with reluctance we leave the stability of the gridlock equilibrium under perfect foresight unresolved.

The equilibrium E_1 is analogous to the quasi-queuing/high demand equilibrium of Verhoef (2005). The on-street parking capacity constraint in our model is analogous to the bottleneck constraint in that paper's model, and the hypercongested traffic flow on downtown streets in our model is analogous to the hypercongestion in the bottleneck queue in that paper's model. Since both E_1 with our adjustment dynamics and Verhoef's quasi-queueing/high demand equilibrium with his adjustment dynamics, are stable, in the discussion that follows we shall focus on the contentious equilibria E_2 and E_3 in our model, which occur when $C = 0$. The earliest reasonably formal stability analyses of hypercongested equilibria are provided by Else (1981) and Nash (1982). Both analyze local stability in the context of the flow-trip-price diagram. Else considers the effects of a perturbation entailing an increase in density with no change in flow, Nash the effects of a perturbation entailing an increase in flow with no change in density. These perturbations are inapplicable to our isotropic network for which flow and density are technologically related.

Verhoef (2005, p. 797) defines an equilibrium, E', to be dynamically stable if starting in some other equilibrium, E, there exists an entry rate, e', such that a once-and-for-all change to entry rate e' results in a path of adjustment to E'. According to this definition of stability, gridlock is not a stable equilibrium since if one starts from any other equilibrium and reduces the quantity of trips demanded to zero, the traffic system moves to the origin. In our model, since demand is not perfectly inelastic, the entry rate is endogenous. One can modify Verhoef's definition of dynamic stability, replacing the entry rate with a level of demand intensity. Accordingly, an equilibrium, E', would be defined to be dynamically stable if starting in some other equilibrium, E, there exists a demand intensity, D_0', such that a once-and-for-all change to demand intensity D_0' results in a path of adjustment to E'. According to this definition of equilibrium, gridlock *is* a stable equilibrium, since if one starts at E_2 and increases the demand intensity, the path of adjustment leads to the gridlock equilibrium.

An equilibrium is defined to be stable with respect to particular adjustment dynamics. The gridlock equilibrium is not dynamically stable with respect to adjustment dynamics that assume demand to be completely inelastic, but is dynamically stable with respect to adjustment dynamics that allow for any price sensitivity of demand. We believe that adjustment dynamics that allow for price sensitivity of demand are more realistic. Verhoef (2005) defines stability with reference to the user cost curve only, without indicating whether demand is price sensitive or completely inelastic. We judge his analysis to be inconsistent since it treats demand as exhibiting price sensitivity but then defines stability with respect to adjustment dynamics that assume demand to be completely inelastic.

6 Concluding Remarks

There has been an active debate for fifty years concerning the stability properties of equilibrium steady states of simple, congestible traffic systems. The debate has remained unresolved because of the technical difficulty of determining even the qualitative transient dynamic properties of traffic flow on even a single road. Analyzing the transient dynamics of an isotropic network of downtown streets is considerably easier than for a road, since out of steady state traffic is the same everywhere on the isotropic network but not on a road. Taking advantage of this simplification, this paper is the first to provide a comprehensive stability analysis of the economic (permitting the entry rate to be price-responsive) steady-state equilibria of a congestible traffic system. It employs the model of downtown parking and traffic congestion developed by Arnott and Inci (2006), modified to permit rigorous state-space analysis.

We started the paper by asking the question: How does a steady-state traffic system that is operating at capacity flow, with no queues, respond to a once-and-for-all increase in demand? The current wisdom, as represented in Small and Verhoef (2007), is that the system moves to a new steady state with an increase in travel time but the same capacity flow. This is possible through the formation of queues – either vertical queues at entry points or quasi-queues behind bottlenecks. The response of the traffic system scrutinized in this paper is more complex. For small and moderate increases in demand, the system may respond according to the current wisdom. But it may also respond with reduced flow. And for large increases in demand, the new steady state entails gridlock.[24] The discrepancy between our results

[24]How should one respond to the gridlock equilibrium? Since gridlock is rarely encountered in downtown traffic, even in the most heavily congested conditions, it would not be unreasonable to argue that the gridlock equilibrium should be dismissed as unrealistic or indicative of an unrealistic model. We think this would be a mistake, however. We conjecture that the gridlock equilibrium is but one example of stable, steady-state equilibria in which, under high demand conditions, throughput is less than capacity throughput. As we noted in the introduction, there are many other

and those in Verhoef's papers appears to stem from their use of a stability criterion that is either unrealistic (if completely inelastic demand is assumed) or inappropriate (if price-sensitive demand is assumed). We say "appears to" stem since we have established that applying Verhoef's stability criterion to our model renders equilibria on the backward-bending portion of the user cost curve unstable, that according to our criterion are stable, but not that applying our stability criterion to the model in Verhoef (2005) would result in the emergence of new stable equilibria. The reason we are unable to do so is that the network geometry in Verhoef (2005) is not amenable to the state-space analysis we applied.

Our results are potentially important for three reasons. First, they provide insight into the nature of economic equilibrium in congestible systems. Second, they suggest that the deadweight loss due to traffic congestion (and therefore the potential efficiency gain from congestion pricing) may be significantly higher than current estimates, because the excessive demand for travel caused by underpricing congestion may lead not only to increased travel time but also to reduced flow. And third, they suggest that in heavily congested conditions, traffic management policy should aim to prevent catastrophic transitions to reduced-flow equilibria.[25]

Since our model is particular, we cannot claim that its qualitative features are general. Nevertheless, by providing a counterexample, our results challenge the generality of the current wisdom. Furthermore, recent empirical and traffic microsimulation studies obtain results consistent with our main finding – that an increase in demand may lead to a decrease in flow.[26] Further analysis will be required to de-

congestible systems in which the analog of throughput falls when the system is very heavily loaded. A demand spike occurs, overloading the system and causing a catastrophic transition to a new part of the phase plane in which the attractor is a steady-state equilibrium with reduced throughput. Why should traffic systems not behave in the same way?

[25] Indeed, this appears to be the intention of ramp metering (Varaiya, 2008) and some other traffic calming measures.

[26] May *et al.* (2000) used the traffic microsimulation model NEMIS to simulate how average network velocity (veh-km/veh-hr) and average network travel (veh-km/hr) in hypothetical grid and ring radial networks respond to a fixed entry flow over a one-hour period. For both network configurations, they found that above critical levels of entry flow, average network travel declines. While this is not quite the appropriate experiment to investigate steady-state equilibria (an appropriate experiment would hold demand intensity, D_0, fixed, allowing for price-sensitive demand, for an extended period of time), the result suggests hypercongested travel at the aggregate level. The paper provided no explanation for the result.

Using fixed detectors, Geroliminis and Daganzo (2008) obtained data on traffic density and velocity for a 10 km^2 area of Yokohama, Japan. They then averaged traffic density and velocity across termine whether our model's challenge to the current wisdom is robust. Will our qualitative results extend to rush-hour equilibria[27] and to more complex and realistic networks?

A Appendix

A.1 Derivation of $E = T/(mt)$

This appendix derives the equilibrium condition $E = T/(mt)$. Let A (arrivals) denote the cumulative number of cars that have entered the downtown area, X (exits) denote the cumulative number of cars that have exited the downtown area, and S the stock of occupied parking spaces. We have the stock identity that $A = T + C + S + X$. Thus, $\dot{A} = \dot{T} + \dot{C} + \dot{S} + \dot{X}$. Moreover, D is the entry rate into downtown and thus $\dot{A} = D$. Since visit lengths are Poisson distributed, with mean l, we have $\dot{X} = S/l$. Letting E denote the exit rate from the in-transit pool and Z the exit rate from the cruising-for-parking pool, we have that $\dot{T} = D - E$, $\dot{C} = E - Z$, and $\dot{S} = Z - S/l$. We have two régimes to consider. In the saturated parking régime:

$$Z = \frac{P}{l} \tag{A-1}$$

$$\dot{P} = 0 \quad , \tag{A-2}$$

sensors for five-minute intervals for one weekday and one weekend day. The plot of the resulting points shows not only a tight macroscopic fundamental diagram between average velocity, average flow, and average density, but also numerous points in the hypercongested region. They then used taxis as probes, plotting for five-minute intervals taxi trips and taxi average velocities, and obtained very similar results. While the results apply to a situation of time-varying rather than time-invariant demand, they indicate that large subnetworks and not just individual links experience hypercongestion, which is consistent with the logic that in high demand situations there can be steady-state equilibria with throughput less than capacity throughput.

[27] In the simplest of the endogenous scheduling models in which identical individuals travel from a common origin to a common destination over the rush hour, and have a common desired arrival time, equilibrium can be represented as the intersection of a rush-hour travel demand function and a reduced-form supply function that relates the common trip cost to the number of commuters over the entire rush hour. With reasonable parameter values, it seems implausible that an increase in demand intensity could lead to a decrease in the equilibrium number of commuters over the entire rush hour, since the traffic system responds to an increase in demand intensity through a lengthening of the rush hour. We conjecture however that an increase in demand intensity can lead to reduced flow.

whereas in the unsaturated parking régime:

$$Z = E \tag{A-3}$$

$$\dot{S} = E - \frac{S}{l} \quad . \tag{A-4}$$

From these equations, it is evident that the entire evolution can be determined once E is calculated. Let M be the cumulative number of cars that have exited the in-transit pool, so that $E = \dot{M}$ and $T = A - M$. The technology of traffic congestion is captured by the function $t = t(T, C, P)$ or alternatively by $v = v(T, C, P)$. Consider a car that enters the in-transit pool at time u. By time w it has traveled a distance

$$x(u, w) = \int_u^w v(y)\, dy \quad . \tag{A-5}$$

Thus, of the cars that enter the in-transit pool at time u, the proportion that have exited it by time w, $\Xi(u, w)$, is

$$\Xi(u, w) = 1 - e^{-hx(u,w)} \quad , \tag{A-6}$$

where $h = 1/m$, and so the number that have exited it per unit-area by time w, $N(u, w)$, is

$$N(u, w) = \Delta(u) \left(1 - e^{-hx(u,w)}\right) \quad . \tag{A-7}$$

As a result, the total number of cars that have exited the in-transit pool per unit-area by time w, $M(w)$, is

$$M(w) = \int_0^w \Delta(u) \left(1 - e^{-hx(u,w)}\right) du \quad . \tag{A-8}$$

Differentiating this with respect to w, we obtain

$$\dot{M}(w) = \Delta(w) \left(1 - e^{-hx(w,w)}\right) + \int_0^w \Delta(u)\, hx_w(u, w)\, e^{-hx(u,w)} du \quad . \tag{A-9}$$

Since the first term on the right-hand side is zero, we have

$$\dot{M}(w) = \int_0^w \Delta(u)\, hx_w(u, w)\, e^{-hx(u,w)} du \quad . \tag{A-10}$$

From (A-5), $x_w(u, w) = v(w)$. This comes out of the integral, so that we have

$$\dot{M}(w) = v(w) \int_0^w \Delta(u)\, he^{-hx(u,w)} du \quad . \tag{A-11}$$

The total number of cars that have entered the in-transit pool per unit-area by time w is simply

$$A(w) = \int_0^w \Delta(u)\, du \quad . \tag{A-12}$$

Thus, the stock of cars in the in-transit pool per unit-area at time w is

$$T(w) = \int_0^w \Delta(u)\, e^{-hx(u,w)} du \quad . \tag{A-13}$$

Combining this with (A-11), we obtain $\dot{M}(w) = v(w)hT(w)$, and thus $E = vhT$ or

$$E = \frac{T}{mt} \quad . \tag{A-14}$$

A.2 Equilibrium and stability with no parking constraint

In the body of the paper, we focused on the situation where parking capacity is less than maximum throughput. Here we comment on how the results are altered when either there is no parking capacity constraint or parking capacity exceeds maximum throughput. One reason that analysis of this simplified model is interesting is that comparison of its properties with those of the text's model points to the role played by the parking capacity constraint. Another is that Small and Chu (2003) investigates the simplified model's dynamic response to a temporary demand spike and its endogenous scheduling equilibrium, but not its steady-state equilibria with price-sensitive demand.

Modifying the example so that the demand function here with $\lambda = 0$ is the same as before with $\lambda = 1$, Figures 4 and 5 continue to apply, except that the parking capacity constraint does not bind. There are then three equilibria E_4, E_3, and E_2. E_4 is locally stable and is reached starting from any initial point to the left of the locus dE_2g in Figure 4; E_3 is the locally stable, gridlock equilibrium, and is reached from any initial point to the right of the locus dE_2g; and E_2 is the saddle-path stable equilibrium, and is reached from any initial point on the locus dE_2g.

Since there are no cars cruising for parking, flow and throughput coincide, and traffic is congested when $T < V_j/2$ and is hypercongested when the inequality is reversed. Since $T = V_j/2$ at the peak of the $\dot{S} = 0$ locus, as Figure 4 is drawn E_4 is congested and E_2 is hypercongested. But from Figure 5, it can be seen that if demand is increased so that E_4 lies on the backward-bending portion of the user cost curve, all the equilibria are hypercongested. And above a critical level of demand, the equilibria E_2 and E_4 disappear and only the gridlock equilibrium E_3 remains.

A.3 Trajectories

This appendix presents some trajectories of the differential equation system. For expositional convenience, we shall make a transformation of variables and reduce the 3D system to a 2D system. The proper transformation is defined as follows. Define

$$R = R^{+} + R^{-} \quad , \tag{A-15}$$

where

$$R^{+} = \max\{R, 0\} = \frac{R + |R|}{2} \tag{A-16}$$

$$R^{-} = \min\{R, 0\} = \frac{R - |R|}{2} \quad . \tag{A-17}$$

Let $C(u) = R^{+}(u)$ and $S(u) = P + R^{-}(u)$. Therefore, R is defined to be the difference between the pool of cars cruising for parking per unit area, C, and the pool of unoccupied parking spaces per unit area, $P - S$. Note that when $R(u) \geq 0$, $C(u) = R(u)$ and $S(u) = P$, and when $R(u) \leq 0$, $C(u) = 0$ and $S(u) = P + R(u)$. The transformed autonomous differential equation system is given by

$$\dot{T}(u) = D\left(\rho\left(mt\left(T(u), \frac{1}{2}(|R| + R), P\right) + \frac{\frac{1}{2}(|R| + R)\,l}{P}\right) + \lambda l\right) - \frac{T(u)}{mt\left(T(u), \frac{1}{2}(|R| + R), P\right)} \tag{A-18}$$

$$\dot{R}(u) = \frac{T(u)}{mt\left(T(u), \frac{1}{2}(|R| + R), P\right)} - \frac{P + \frac{1}{2}(R - |R|)}{l} \quad . \tag{A-19}$$

Since this system is Leibnitz, all existence and uniqueness theorems apply. However, the system is only piecewise differentiable and there is a phase transition at $R^{+}(u) = 0$.

Geometrically, this transformation corresponds to folding the top plane of Figure 7 flat and relabeling the origin. The transformed system is shown in Figure 11, which assumes the parameter values stated in Section 4.1. In this figure, the $\dot{T}(u) = 0$ and $\dot{R}(u) = 0$ loci are shown with solid lines and trajectories with dotted lines. We do not show E_3 since doing so would entail loss of important detail, but one should note that the $\dot{R}(u) = 0$ locus cuts the $x-$axis at $\{T, R\} = \{1778.17, -3712\}$. However, we show one of the trajectories approaching E_3 at the far right of the figure.

References

[1] Agnew, C., 1977. The theory of congestion tolls, *Journal of Regional Science* 17, 381-393.

[2] Arnott, R., Inci, E., 2006. An integrated model of downtown parking and traffic congestion, *Journal of Urban Economics* 60, 418-442.

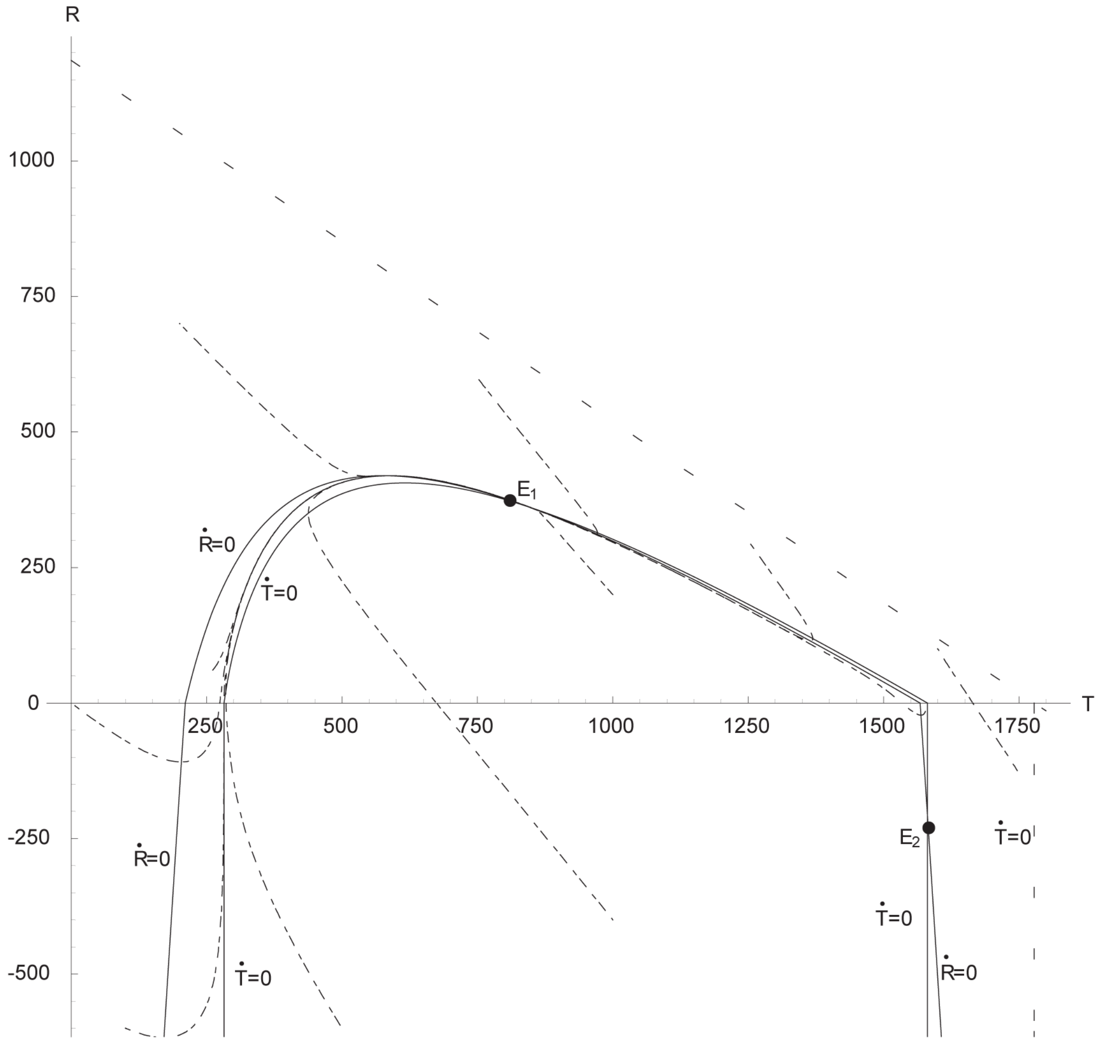


Figure 11: Trajectories of the transformed differential equation system

[3] Arnott, R., Rowse, J., 1999. Modeling parking, *Journal of Urban Economics* 45, 97-124.

[4] Arnott, R., Rowse, J., 2009. Downtown parking in auto city, *Regional Science and Urban Economics* 39, 1-14.

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

[6] Else, P., 1981. A reformulation of the theory of optimal congestion taxes, *Journal of Transport Economics and Policy* 15, 217-232.

[7] Daganzo, C., 1992. The cell transmission model. Part I: a simplified representation of highway traffic, *California Partners for Advanced Transit and Highways (PATH). Research Reports*: Paper UCB-ITS-PRR-93-7.

[8] Daganzo, C., 1997. *Fundamentals of Transportation and Traffic Operations*, Pergamon, Kidlington, UK.

[9] Geroliminis, N., Daganzo, C., 2008. Existence of urban-scale macroscopic fundamental diagrams: some experimental findings, *Transportation Research B* 42, 759-770.

[10] Greenshield, B., 1935. A study of traffic capacity, *Highway Research Board Proceedings* 14, 448-477.

[11] Henderson, J., 1981. The economics of staggered work hours, *Journal of Urban Economics* 9, 349-364.

[12] Lindsey, R., 1980. Non-steady-state traffic flow, Directed research project in partial fulfillment of the requirements for the degree of Ph.D., Department of Economics, Princeton University.

[13] May, A., Shepherd, S., Bates, J., 2000. Supply curves for urban road networks, *Journal of Transport Economics and Policy* 34, 261-290.

[14] Nash, C., 1982. A reformulation of the theory of optimal congestion taxes: a comment, *Journal of Transport Economics and Policy* 26, 295-299.

[15] Small, K., Chu, X., 2003. Hypercongestion, *Journal of Transport Economics and Policy* 37, 319-352.

[16] Small, K., Verhoef, E., 2007. *The Economics of Urban Transportation*, Routledge, Abingdon, England.

[17] Varaiya, P., 2008. Congestion, ramp metering and tolls, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 366, 1921-1930.

[18] Verhoef, E., 1999. Time, speeds, flows and densities in static models of road traffic congestion and congestion pricing, *Regional Science and Urban Economics* 29, 341-369.

[19] Verhoef, E., 2001. An integrated dynamic model of road traffic congestion based on simple car-following theory: exploring hypercongestion, *Journal of Urban Economics* 49, 505-542.

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

[21] Verhoef, E., 2005. Speed-flow relations and cost functions for congested traffic: theory and empirical evidence, *Transportation Research A* 39, 792-812.

[22] Vickrey, W., 1969. Congestion theory and transport investment, *American Economic Review* 59, 251-261.

[23] Walters, A., 1961. The theory and measurement of private and social cost of highway congestion, *Econometrica* 29, 676-699.