

TI 2014-077/VIII
Tinbergen Institute Discussion Paper



On the Existence and Uniqueness of Equilibrium in the Bottleneck Model with Atomic Users

Hugo Emilio Silva^a

Robin Lindsey^b

André de Palma^c

Vincent A.C. van den Berg^a

^a *Faculty of Economics and Business Administration, VU University Amsterdam, and Tinbergen Institute, the Netherlands;*

^b *University of British Columbia, Canada;*

^c *École Normale Supérieure de Cachan, Cachan, France;*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

On the Existence and Uniqueness of Equilibrium in the Bottleneck Model with Atomic Users

Hugo E. Silva^{a,b,*}, Robin Lindsey^c, André de Palma^d, Vincent A.C. van den Berg^{a,b}

This version: May 26th, 2014

^a*Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.*

^b*Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS, Amsterdam, The Netherlands.*

^c*Sauder School of Business, University of British Columbia, Canada.*

^d*Ecole Normale Supérieure de Cachan, Département Economie et Gestion, Cachan, France.*

Abstract

This paper investigates the existence and uniqueness of equilibrium in the Vickrey bottleneck model when each user controls a positive fraction of total traffic. Users simultaneously choose departure schedules for their vehicle fleets. Each user internalizes the congestion cost that each of its vehicles imposes on other vehicles in its fleet. We establish three results. First, a pure strategy Nash equilibrium (PSNE) may not exist. Second, if a PSNE does exist, identical users may incur appreciably different equilibrium costs. Finally, a multiplicity of PSNE can exist in which no queuing occurs but departures begin earlier or later than in the social optimum. The order in which users depart can be suboptimal as well. Nevertheless, by internalizing self-imposed congestion costs individual users can realize much, and possibly all, of the potential cost savings from either centralized traffic control or time-varying congestion tolls.

Keywords: Bottleneck model, Large users, Atomic users, Existence of Equilibrium, Uniqueness of Equilibrium

1. Introduction

Individual users of roads and other transportation facilities are usually assumed to be small in the sense that they control a negligible fraction of total traffic. Yet large users are prevalent in many settings. A commercial airline often accounts for a sizable fraction of total traffic at an airport. Postal services and major freight shippers such as Fedex, UPS, and Purolator Courier operate large vehicle fleets that travel long distances each

*Corresponding author

Email addresses: h.silvamentalva@vu.nl (Hugo E. Silva), robin.lindsey@sauder.ubc.ca (Robin Lindsey), andre.depalma@ens-cachan.fr (André de Palma), vberg@feweb.vu.nl (Vincent A.C. van den Berg)

day.¹ Major employers such as government departments and large corporations can add substantially to traffic on certain roads at peak times. Large users such as these suffer from the congestion delays their own aircraft, trucks, or other vehicles impose on each other. Thus, at airports, on congested roads, and on other transportation infrastructure networks, one would expect large users to internalize their self-imposed delays, and therefore to make different trip-related decisions than small users controlling the same aggregate traffic.

Following the terminology of game theory we will refer to small users as *non-atomic*, and large users that control a positive fraction of traffic as *atomic*.² There are several branches of literature on congestion with atomic users. In the aviation literature, Daniel (1995) was the first to recognize that airlines with market power and large shares of total traffic could internalize the delays their aircraft impose on each other. Brueckner (2002) showed that under Cournot competition airlines fully internalize self-imposed congestion. Further contributions in this line have been made by Pels and Verhoef (2004), Brueckner (2005), Zhang and Zhang (2006), Basso and Zhang (2007), Brueckner and Van Dender (2008), and Silva and Verhoef (2013). In the context of road transportation, route-choice decisions by atomic users have been studied (e.g., Devarajan, 1981; Marcotte, 1987; Harker, 1988; Catoni and Pallottino, 1991; Miller et al., 1991; Cominetti et al., 2009; de Palma and Engelson, 2012). There is also a literature in operations research and computer science on atomic congestion games (e.g., Fotakis et al., 2008; Hoefer and Skopalik, 2009).

The above-mentioned studies have used static models of congestion except for Daniel (1995) who uses a stochastic queuing model empirically. Except for a few studies described below, traffic congestion with atomic users has not been studied with deterministic dynamic models. This is surprising because the timing of trips matters a great deal for both passenger and freight transportation, and congestion is largely a consequence of peak-period loads.

The goal of this paper is to investigate the fundamental questions of existence and uniqueness of equilibrium in trip-timing decisions with atomic users. To focus the analysis on fundamentals while minimizing mathematical complications we use Vickrey's (1969) bottleneck model. The essence of the bottleneck model is that users trade off the costs of queuing delay at the bottleneck with the costs of schedule delay (i.e., arriving earlier or later than desired). The bottleneck model has been used to study many aspects of trip-timing decisions with congestion including: congestion pricing, route choice on simple road net-

¹FedEx also handles about 150 daily flights out of Memphis International Airport and its air-cargo operations support tens of thousands of jobs. UPS operates on a similar scale out of Louisville, Kentucky (The Economist, 2013).

²This terminology contrasts with the terminology used in the literature on airport congestion, beginning with Daniel (1995), in which users that control a negligible fraction of traffic and treat the congestion levels as parametric are called atomistic users. Somewhat confusingly, atomistic users are therefore non-atomic, and non-atomistic users are atomic.

works, mode choice, trip chaining, parking congestion, staggered work hours, and flextime. Existence and uniqueness of equilibrium in the bottleneck model have also been established under relatively general assumptions about trip-timing preferences and heterogeneity of non-atomic users (e.g., Newell, 1987; Lindsey, 2004). However, very little consideration has been given to atomic users in studies that use either the bottleneck model or other dynamic models.

A few studies have employed a variant of the bottleneck model in which time is discretized, and the number of users is finite, so that each user controls a positive measure of traffic (e.g., Levinson, 2005; Zou and Levinson, 2006; Otsubo and Rapoport, 2008). However, these studies assume that each user controls only one vehicle so that self-internalization of congestion does not come into play. To the best of our knowledge, only Daniel (2009) and Silva et al. (2014) have explored the scheduling decisions of atomic agents in the standard, continuous-time bottleneck model. These studies consider, in the context of aviation, a sequential competition between a Stackelberg leader with market power and a group of perfectly competitive airlines (non-atomic users). Both studies show that, when users have homogeneous preferences, non-atomic users schedule all their flights during the peak period when passengers prefer to arrive. The Stackelberg leader schedules a fraction of its flights during the peak as well. Queuing time at the bottleneck evolves at the same rate as in the standard model of non-atomic players. The leader schedules its remaining flights earlier and later in the off-peak (and less popular) periods and limits its departure rate to bottleneck capacity so that no queue develops.

The existence of a unique equilibrium in Daniel (2009) and Silva et al. (2014) hinges on the sequential nature of the game they consider, and the assumption that there is only one atomic agent. By contrast, we focus in this paper on settings with two atomic users who make scheduling decisions simultaneously. The solution concept we employ is pure strategy Nash equilibrium (PSNE). We establish three major results. First, we show that a PSNE may not exist. We demonstrate this for an example featuring two identical atomic users who each control half the total traffic. The trip-timing preferences of each vehicle in each fleet are described by parameters $\{\alpha, \beta, \gamma, t^*\}$, where α is the cost of travel time, β is the cost of schedule delay early, γ is the cost of schedule delay late, and t^* is the desired arrival time. We show that if $\gamma > \alpha$, a PSNE does not exist.

Second, for the same example, we show that if $\gamma \leq \alpha$ a multiplicity of PSNE exists in which no queuing occurs and the timing of departures is socially optimal. The PSNE differ according to the departure rates of individual users and the equilibrium costs they incur. Depending on parameter values, one of the two users can incur up to three quarters of total costs. The cases $\gamma > \alpha$ and $\gamma \leq \alpha$ are both of theoretical interest, and each may be relevant in particular settings. Most empirical studies of scheduling preferences for automobile drivers have obtained estimates that satisfy $\gamma > \alpha$ (e.g., Small, 1982; Wardman,

2001; Asensio and Matas, 2008). ³Daniel and Harback (2008) find that $\gamma < \alpha$ holds for most airlines at major US airports.

Third, we consider a variant of the example in which the two users differ in their desired arrival times t^* and can have fleets of different size. We show that — independent of the relative size of α and γ — a multiplicity of PSNE can exist in which no queuing occurs but the timing of departures is not optimal. Depending on parameter values, the PSNE may begin earlier than, later than, or at the same time as the social optimum. The order in which users depart can be suboptimal as well. Nevertheless, by internalizing self-imposed congestion costs the two users realize much, and possibly all, of the potential cost savings from either centralized traffic control or time-varying congestion tolls.

These examples demonstrate that neither the existence of equilibrium nor the uniqueness of an equilibrium (if one exists) is guaranteed under conditions where a unique PSNE does exist if all the traffic were controlled by non-atomic users. Given the central role of equilibrium models in the analysis of transportation systems, these results are troubling and highlight the need for further research.

The paper is organized as follows. Section 2 reviews the no-toll equilibrium and social optimum in the standard bottleneck model with non-atomic users. Section 3 demonstrates the possible non-existence of PSNE, and the non-uniqueness of individual departure rates and costs where a PSNE does exist. Section 4 demonstrates the possible nonuniqueness of PSNE in the timing of departures when no queuing occurs, and the degree of inefficiency relative to the social optimum. Section 5 concludes.

2. The bottleneck model with homogeneous non-atomic users

The bottleneck model was developed by Vickrey (1969) and extended by Arnott et al. (1990, 1993). It is reviewed in Arnott et al. (1998) and de Palma and Fosgerau (2011), and the summary here is brief. In the model, all users travel from a common origin to a common destination along a single link that has a bottleneck with fixed flow capacity, s . Without loss of generality, travel times from the origin to the bottleneck and from the bottleneck to the destination are normalized to zero. If there is no queue upstream of the bottleneck, travel time through the bottleneck is also zero and departure time from the origin coincides with arrival time at the destination. If the departure rate exceeds s , a queue develops. Let \hat{t} be the most recent time at which there was no queue, and $r(t)$ the aggregate departure rate from the origin at time t . The number of vehicles in the queue is then:

$$Q(t) = \int_{\hat{t}}^t (r(u) - s) du.$$

³ de Palma and Fontan (2001) list estimates from eleven studies. Including their own estimates, there are ten cases with $\gamma > \alpha$, and two cases with $\alpha > \gamma$.

Travel time through the bottleneck is $T(t) = Q(t)/s$, and a traveler who departs at time t arrives at time $t_a = t + T(t)$.

Following Small (1982) and de Palma et al. (1983), users are assumed to have a desired arrival time t^* . They incur a unit cost of $\beta > 0$ for arriving early, and a unit cost of $\gamma > 0$ for arriving late. Travel time is valued at α , with $\alpha > \beta$. The costs of schedule delay and travel time are additive so that the generalized cost of a trip, $c(t)$, is:

$$c(t) = \alpha \cdot T(t) + \begin{cases} \beta \cdot (t^* - t - T(t)), & t + T(t) \leq t^* \\ \gamma \cdot (t + T(t) - t^*), & t + T(t) \geq t^* \end{cases}. \quad (1)$$

The number (or measure) of users, N , is assumed to be exogenous (i.e., independent of trip cost). Each user decides when to depart from the origin by trading off schedule delay against travel delay. A pure strategy Nash equilibrium (PSNE) is a set of departure times such that no user can benefit (i.e., reduce trip cost) by unilaterally changing departure time while taking the departure times of all other users as given.

2.1. No-toll equilibrium

Let superscript n denote the no-toll non-atomic PSNE, and t_s^n and t_e^n denote the start and end of the travel period. Let \tilde{t} be the departure time for which a user arrives on time (i.e., $\tilde{t} + T(\tilde{t}) = t^*$). In a PSNE with no toll, $c(t)$ must be constant during the travel period $[t_s^n, t_e^n]$. Users who arrive closer to t^* must incur longer queuing delays in order to offset their lower schedule delay costs. The equilibrium aggregate departure rate is derived by differentiating Eq. (1) and setting the derivative to zero:

$$r^n(t) = \begin{cases} \frac{\alpha \cdot s}{\alpha - \beta}, & t \in (t_s^n, \tilde{t}) \\ \frac{\alpha \cdot s}{\alpha + \gamma}, & t \in (\tilde{t}, t_e^n) \end{cases}. \quad (2)$$

The assumption $\alpha > \beta$ assures that the departure rate for early arrivals is positive and finite.⁴ For ease of reference we will sometimes call the departure rate for early arrivals the *early departure rate*, and the departure rate for late arrivals the *late departure rate*.

There are two further equilibrium conditions. One is that the first and last users to depart, who encounter no queue, must incur equal schedule delay costs:

$$\beta \cdot (t^* - t_s^n) = \gamma \cdot (t_e^n - t^*). \quad (3)$$

The other condition is that the travel period lasts for N/s :

$$t_e^n - t_s^n = \frac{N}{s}. \quad (4)$$

⁴The condition $\alpha > \beta$ is plausible since a user who is destined to arrive early is likely to prefer arriving early to prolonging the trip by making a detour. The condition is also supported by Small's (1982) estimates for automobile commuting trips. de Palma and Fontan's (2001) list of estimates includes ten cases with $\beta < \alpha$, and four cases with $\beta > \alpha$.

Together, equilibrium conditions (2), (3) and (4) yield:

$$t_s^n = t^* - \frac{\gamma}{\beta + \gamma} \cdot \frac{N}{s}, \quad (5)$$

$$t_e^n = t^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N}{s}, \quad (6)$$

$$\tilde{t} = t^* - \frac{\beta}{\alpha} \cdot \frac{\gamma}{\beta + \gamma} \cdot \frac{N}{s},$$

$$c^n(t) = \delta \cdot \frac{N}{s}, \quad t \in [t_s^n, t_e^n] \text{ with } \delta \equiv \frac{\beta \cdot \gamma}{\beta + \gamma}.$$

Total costs are:

$$TC^n = \delta \cdot \frac{N^2}{s}. \quad (7)$$

2.2. Social optimum

Queuing delay at the bottleneck is a deadweight loss. The social optimum therefore avoids queuing and minimizes total schedule delay costs. The departure rate is maintained at s over a continuous time interval chosen so that the first and last users incur the same schedule delay cost. The departure period is therefore the same as in the laissez-faire PSNE (cf. Eqs. (5) and (6)). Using superscript o to denote the social optimum, these results are recorded for future reference as:

$$r^o(t) = s, \quad t \in (t_s^o, t_e^o), \quad r^o(t) = 0 \text{ otherwise}, \quad (8)$$

$$t_s^o = t^* - \frac{\gamma}{\beta + \gamma} \cdot \frac{N}{s}, \quad (9)$$

$$t_e^o = t^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N}{s}. \quad (10)$$

Total social costs are only half as large as in Eq. (7) for the no-toll equilibrium:

$$TC^o = \frac{\delta}{2} \cdot \frac{N^2}{s} = \frac{1}{2} \cdot TC^n. \quad (11)$$

The difference between total costs in the no-toll equilibrium and social optimum, $TC^n - TC^o$, serves as an upper bound on the benefits from self-internalization of congestion by atomic agents.

3. Existence and non-existence of equilibrium with homogeneous atomic users

In this section we study an example featuring two identical atomic users. We show that if $\gamma > \alpha$, a PSNE in departure schedules does not exist. We then show that if $\gamma \leq \alpha$, a PSNE does exist that entails no queuing and coincides with the social optimum. During early departures the two users can depart at somewhat different rates that add up to s . If

$\gamma < \alpha$, their late departure rates can also differ. Moreover, with $\gamma \leq \alpha$ the users can incur appreciably different total costs for their fleets. At the end of the section we briefly discuss how these results extend to more than two users.

Consider two atomic users, A and B . Each user controls a fleet of $N/2$ vehicles. Trip-timing preferences for each vehicle are defined by the same $\{\alpha, \beta, \gamma, t^*\}$ parameter values. Preferences are therefore independent in the sense that, for a given departure and arrival time, the cost incurred by one vehicle does not depend on when other vehicles in the same fleet make their trips.⁵ Users A and B simultaneously choose departure schedules for their fleets. The schedule for user i is a departure rate function, $r_i(t) \geq 0$.⁶ Each user recognizes that dispatching a vehicle at time t may delay vehicles in its fleet that depart after t . A delay occurs if there is a queue at time t that persists when the later vehicles depart. A delay also occurs if there is no queue prior to t , but the bottleneck is at capacity so that adding a vehicle to the departure schedule at t creates a (small) queue.

As noted above, the existence of a PSNE in this example depends on whether $\gamma > \alpha$, or $\gamma \leq \alpha$. The two cases are considered in the following two subsections.

3.1. Non-existence of PSNE with $\gamma > \alpha$

When $\gamma > \alpha$, a PSNE does not exist. This result is formalized in the following proposition.

Proposition 1. *If $\gamma > \alpha$, a PSNE in departure schedules does not exist.*

We prove Proposition 1 in four steps. First, we prove that a PSNE without queuing does not exist (Lemma 1). Second, we prove that mass departures cannot arise in equilibrium (Lemma 2). Third, we show that there is a unique departure pattern with queuing such that a user cannot reduce its fleet costs by rescheduling a single vehicle (Lemma 3). Finally, we show that this departure pattern is not a PSNE because a user can reduce its fleet costs by rescheduling a positive measure of vehicles in the fleet (Lemma 4). This establishes that a PSNE with queuing does not exist either.

Lemma 1. *When $\gamma > \alpha$, a PSNE without queuing does not exist.*

Proof: Consider a pair of departure schedules, $\{r_A(\cdot), r_B(\cdot)\}$, such that no queuing occurs. Some vehicles must arrive late since otherwise a user could reduce its fleet costs by rescheduling some vehicles to just after t^* . Consider a period (t_1, t_2) of late arrivals and

⁵This assumption seems realistic for delivery vans carrying merchandise or parcels to different customers. It may be inappropriate for vehicles in a military convoy or emergency vehicles traveling to an accident.

⁶Function $r_i(t)$ can be seen as a distribution function of the $N/2$ vehicles over some extended time interval such as a day. This function is not restricted to be continuous, and the possibility of mass departures will be considered.

assume that both users depart during this period (the case where only one user departs is considered later). The bottleneck must be used to capacity since otherwise a user could exploit the residual capacity by advancing departures for vehicles that are scheduled to depart later.

If user i removes a vehicle from the departure schedule at time t it saves a cost of

$$C_i^-(t) = \gamma \cdot (t - t^*). \quad (12)$$

Removing the vehicle saves the late-arrival cost incurred by the vehicle itself, but it has no effect on the rest of the fleet because there is no queue. If user i instead adds a vehicle to the departure schedule, it increases its fleet costs by a marginal private cost (MPC) of

$$C_i^+(t) = \gamma \cdot (t - t^*) + \frac{\alpha + \gamma}{s} \cdot \int_t^{\bar{t}} r_i(u) du, \quad (13)$$

where \bar{t} is the time when the queue created by the additional vehicle disappears.

The first term on the right-hand side of Eq. (13) matches the right-hand side of Eq. (12). The second term is the delay cost imposed on user i 's other vehicles that depart from t to \bar{t} . Each of them suffers an increase in travel time of $1/s$ valued at α , and an increase in late arrival of $1/s$ valued at γ .

The difference between the cost saved by removing a vehicle given by Eq. (12), and the cost of adding a vehicle to the same slot given by Eq. (13), arises when the bottleneck is at capacity but there is no queue. This asymmetry underlies the nonuniqueness of equilibrium considered later in the paper.

Suppose user i advances the departure of a vehicle from t to t' where $t_1 \leq t' < t \leq t_2$. User i 's fleet costs change by:

$$\Delta C_i = -C_i^-(t) + C_i^+(t') = -\gamma \cdot (t - t^*) + \gamma \cdot (t' - t^*) + \frac{\alpha + \gamma}{s} \int_{t'}^t r_i(u) du. \quad (14)$$

The queue induced by adding the vehicle at t' vanishes at t because a departure time slot opened up at t when the vehicle was removed then. Let $\lambda_{t,t'}^i \in [0, 1]$ be an auxiliary variable denoting the average fraction of capacity occupied by user i during the period $[t, t']$.⁷ The change in user i 's fleet costs can then be written as:

$$\Delta C_i = \frac{\alpha + \gamma}{s} \cdot \lambda_{t,t'}^i \cdot s \cdot (t - t') - \gamma \cdot (t - t') = ((\alpha + \gamma) \cdot \lambda_{t,t'}^i - \gamma) \cdot (t - t').$$

For a PSNE to exist, ΔC_i must be nonnegative for both users which requires:

$$\lambda_{t,t'}^i \geq \frac{\gamma}{\alpha + \gamma}, \quad t_1 \leq t' < t \leq t_2, \quad i = A, B. \quad (15)$$

⁷Formally, $\lambda_{t,t'}^i = \int_{t'}^t r_i(u) du / (s \cdot (t - t'))$.

Since the bottleneck is fully utilized, $\lambda_{t,t'}^A + \lambda_{t,t'}^B = 1$. This condition is least restrictive if $\lambda_{t,t'}^A = \lambda_{t,t'}^B = 1/2$ in which case it reduces to $\gamma \leq \alpha$ which is inconsistent with the assumption $\gamma > \alpha$.

Now consider the possibility that only one user, say user A , departs during (t_1, t_2) . This is not a PSNE if user B departs after t_2 since user B could reduce its costs by rescheduling some of its later vehicles into (t_1, t_2) . Doing so reduces their late-arrival costs, and the queue they create disappears during the time slots they vacated. Suppose user B does not depart after t_2 . If $t_1 = t^*$, user B does not depart late at all. But this cannot be a PSNE because user B could gain by rescheduling some of its early-arriving vehicles to just after t^* , thereby reducing their schedule delay costs without imposing any delay on its other vehicles. If $t_1 > t^*$, there must exist a late arrival period (t_0, t_1) during which both users depart. But this case has already been considered, and shown to be inconsistent with a PSNE when $\gamma > \alpha$. *QED*

To this point we have assumed that users depart at a finite rate. In theory, a user could schedule a positive measure of vehicles to depart at a given moment.⁸ Moreover, in the bottleneck model with non-atomic users and steps tolls a PSNE may exist only if mass departures (of non-cooperating vehicles) are possible (Arnott et al., 1990; Lindsey et al., 2012). We now show that in the model with atomic users mass departures cannot occur in a PSNE, regardless of whether $\gamma > \alpha$ holds or not.

Lemma 2. *A PSNE cannot exhibit mass departures.*

Proof: See Appendix A.

We now turn to the final possibility for a PSNE with $\gamma > \alpha$ in which departure rates remain finite and queuing occurs.

Lemma 3. *When $\gamma > \alpha$, there is a unique departure pattern with queuing in which a user cannot reduce its fleet costs by rescheduling a single vehicle.*

Proof: Consider a pair of departure schedules, $\{r_A(\cdot), r_B(\cdot)\}$, and let \tilde{t} denote the departure time for which a user arrives on time (i.e. $\tilde{t} = t^* - T(\tilde{t})$). Assume that a queue exists during a late-departure period (t_l, t_{qe}) , where t_l is an arbitrary time that satisfies $t_l > \tilde{t}$ and t_{qe} is the time when the queue disappears. Since users A and B are identical, it suffices to consider the best response of user A to $r_B(\cdot)$. The MPC to user A of scheduling a vehicle at time $t \in (t_l, t_{qe})$ is

$$C_A(t) = \alpha \cdot T(t) + \gamma \cdot (t + T(t) - t^*) + \frac{\alpha + \gamma}{s} \cdot \int_t^{t_{qe}} r_A(u) du. \quad (16)$$

⁸In practice this might be achieved by assembling a convoy of vehicles on a link that has right-of-way over other links.

Eq. (16) has a similar interpretation to Eq. (13). The first two terms on the right-hand side comprise the cost incurred by the vehicle itself, and the third term is the delay cost imposed on user A 's other vehicles that depart from t to t_{qe} .

User A could reschedule a vehicle from t to another time $t' \in (t_l, t_{qe})$. This would leave t_{qe} unchanged because the additional queuing time caused by inserting the vehicle at t' is offset by the reduction in queuing time due to removing the vehicle at t . Eq. (16) therefore holds if t is replaced by any $t' \in (t_l, t_{qe})$. Hence, a necessary condition for $r_A(\cdot)$ to be a best response to $r_B(\cdot)$ is that $C_A(t)$ in Eq. (16) is constant during the interval (t_l, t_{qe}) . Differentiating Eq. (16) with respect to t , and setting the derivative to zero, yields

$$\frac{\partial C_A(t)}{\partial t} = \gamma + (\alpha + \gamma) \cdot \frac{\partial T(t)}{\partial t} - \frac{\alpha + \gamma}{s} \cdot r_A(t) = 0.$$

Using the relationship $\partial T(t)/\partial t = (r_A(t) + r_B(t) - s)/s$, this condition simplifies to

$$r_B(t) = \frac{\alpha \cdot s}{\alpha + \gamma}, \quad t \in (t_l, t_{qe}). \quad (17)$$

According to Eq. (17), user A is willing to schedule a vehicle for late arrival when there is a queue if, and only if, user B is departing at exactly the rate $\alpha \cdot s / (\alpha + \gamma)$. This is none other than the equilibrium aggregate departure rate for the model with non-atomic users (cf. Eq. (2)). Eq. (17) also holds for user B with $r_A(t)$ in place of $r_B(t)$. The aggregate departure rate during an interval of late arrivals must therefore be $r(t) = r_A(t) + r_B(t) = 2 \cdot \alpha \cdot s / (\alpha + \gamma)$. With $\gamma > \alpha$, $r(t) < s$ and the queue must be shrinking for late arrivals. Consequently, a queue must exist at time \tilde{t} and it is possible to set $t_l = \tilde{t}$, where $\tilde{t} + T(\tilde{t}) = t^*$. This in turn implies

$$r_A(t) = r_B(t) = \frac{\alpha \cdot s}{\alpha + \gamma}, \quad t \in (\tilde{t}, t_{qe}).$$

Since a queue exists at time \tilde{t} , it must have built up during a period of early arrivals before \tilde{t} . Let t_q be the time at which queuing begins. The MPC to user A of scheduling a vehicle at any time $t \in (t_q, \tilde{t})$ is

$$C_A(t) = \alpha \cdot T(t) + \beta \cdot (t^* - t - T(t)) + \frac{\alpha - \beta}{s} \cdot \int_t^{\tilde{t}} r_A(u) du + \frac{\alpha + \gamma}{s} \cdot \int_{\tilde{t}}^{t_{qe}} r_A(u) du. \quad (18)$$

Again, the first two terms on the right-hand side of Eq. (18) comprise the cost borne by the vehicle itself. The third term is the cost imposed on user A 's other vehicles that depart after t but still arrive early. Each of them suffers an increase in travel time of $1/s$ valued at α , and benefits from a reduction in early arrival of $1/s$ valued at β . The last term in Eq. (18) is the cost imposed on user A 's other vehicles that arrive late.

A necessary condition for $r_A(\cdot)$ to be a best response to $r_B(\cdot)$ is for $C_A(t)$ to be constant during (t_q, \tilde{t}) . Setting the derivative of $C_A(t)$ to zero, one obtains a counterpart to Eq. (17):

$$r_B(t) = \frac{\alpha \cdot s}{\alpha - \beta}, \quad t \in (t_q, \tilde{t}). \quad (19)$$

An analogous necessary condition applies for user B . Hence, the aggregate early departure rate must be $r(t) = r_A(t) + r_B(t) = 2 \cdot \alpha \cdot s / (\alpha - \beta)$.

In summary, the unique departure rate of the candidate PSNE during the full period of queuing is

$$r_A(t) = r_B(t) = \frac{r(t)}{2} = \begin{cases} \frac{\alpha \cdot s}{\alpha - \beta}, & t \in (t_q, \tilde{t}) \\ \frac{\alpha \cdot s}{\alpha + \gamma}, & t \in (\tilde{t}, t_{qe}) \end{cases}, \quad (20)$$

where

$$\tilde{t} + T(\tilde{t}) = t^*. \quad (21)$$

Because queuing begins at t_q , and ends at t_{qe} , cumulative departures during the period (t_q, t_{qe}) match cumulative arrivals:

$$\int_{t_q}^{t_{qe}} (r(u) - s) du = 0. \quad (22)$$

Eqs. (20), (21), and (22) define evolution of the queue for the candidate PSNE with queuing.

By Lemma 1, no vehicles can depart without queuing after t_{qe} , so departures end at time $t_e = t_{qe}$. However, the cost of a vehicle trip at the beginning of the queuing period, $c(t_q)$, is less than the cost at the end of the period, $c(t_e)$, because the trip at t_q imposes a private delay cost on subsequent vehicles whereas the trip at t_e does not (if $c(t_q) = c(t_e)$, $C^+(t_q) > C^+(t_e)$ would hold and the user equilibrium condition would be violated). Departures must therefore occur during some time interval $[t_s, t_q]$ preceding t_q . This interval is defined by two conditions. First, vehicle trip cost must be the same at t_s and t_e since otherwise a user could reschedule vehicles from the time with higher cost to the time with lower cost and reduce its overall fleet costs without causing any queuing. Second, the full departure period $[t_s, t_e]$ must be long enough for all N vehicles to pass the bottleneck. Eqs. (3) and (4) for the non-atomic PSNE therefore hold for the candidate PSNE:

$$\beta \cdot (t^* - t_s) = \gamma \cdot (t_e - t^*), \quad (23)$$

$$t_e - t_s = \frac{N}{s}. \quad (24)$$

The two users each control $N/2$ vehicles and schedule the same number of vehicles during the queuing period. Therefore, they must also schedule the same number during $[t_s, t_q]$:

$$\int_{t_s}^{t_q} r_A(u) du = \int_{t_s}^{t_q} r_B(u) du, \quad (25)$$

where

$$r_A(u) + r_B(u) = s, \quad u \in (t_s, t_q). \quad (26)$$

The early departure schedule defined by Eqs. (25) and (26) is consistent with a PSNE as far as trip timing by individual vehicles. The first vehicle scheduled at t_s creates the same

MPC as all vehicles scheduled during $[t_q, t_e]$. Vehicles scheduled during (t_s, t_q) create a lower MPC, so that rescheduling them to any time outside (t_s, t_q) would increase total fleet costs. Rescheduling any vehicle into (t_s, t_q) would also increase fleet costs because it would impose a queuing delay on all vehicles departing later until t_e , and would therefore create a higher MPC than a vehicle departing at t_q .

Together, Eqs. (20)–(26) define the candidate PSNE with queuing. Using superscript c to denote this candidate, the critical times are:

$$t_s^c = t^* - \frac{\gamma}{\beta + \gamma} \cdot \frac{N}{s}, \quad (27)$$

$$t_e^c = t^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N}{s}, \quad (28)$$

$$\tilde{t}^c = t^* - \frac{\beta \cdot (\gamma - \alpha)}{2 \cdot \alpha \cdot (\beta + \gamma)} \cdot \frac{N}{s}, \quad (29)$$

$$t_q^c = t^* - \frac{\beta \cdot (\gamma - \alpha)}{(\alpha + \beta) \cdot (\beta + \gamma)} \cdot \frac{N}{s}. \quad (30)$$

We now show that the candidate PSNE just derived is not a PSNE because either user can reduce its fleet costs by rescheduling a positive fraction of its vehicles. This result is formalized as:

Lemma 4. *When $\gamma > \alpha$, the candidate departure pattern with queuing in which users cannot reduce costs by rescheduling a single vehicle is not a PSNE with respect to rescheduling a positive fraction of the fleet.*

Proof: The candidate PSNE is depicted in Figure 1. Cumulative departures during the whole travel period are shown by the piecewise linear schedule $OABC$. During the first interval (t_s^c, t_q^c) , the two users depart early at rates consistent with Eqs. (25). During the remaining interval (t_q^c, t_e^c) , user B contributes to cumulative departures the portion between schedule $ADGE$ and the horizontal line AF . User A contributes the equally big portion between schedule ABC and schedule $ADGE$. During the interval (t_q^c, \tilde{t}^c) , each user departs early at rate $\alpha \cdot s / (\alpha - \beta)$, and during (\tilde{t}^c, t_e^c) each user departs late at rate $\alpha \cdot s / (\alpha + \gamma)$. At time \hat{t} , each user's cumulative departures over the interval (t_q^c, \hat{t}) match cumulative bottleneck throughput over the same interval.

Suppose user A reschedules all its vehicles from period (t_q^c, \hat{t}) to period (\hat{t}, t_e^c) , and maintains a departure rate of $\gamma \cdot s / (\alpha + \gamma)$ during (\hat{t}, t_e^c) . As a consequence, the queue caused by B between t_q^c and \hat{t} fully dissipates at \hat{t} . Since user B departs at rate $\alpha \cdot s / (\alpha + \gamma)$ during (\hat{t}, t_e^c) , the total departure rate during this period equals s and the bottleneck is fully utilized without queuing. Moreover, since user B keeps the bottleneck fully utilized during (t_q^c, \hat{t}) , all vehicles in both fleets complete their trips by t_e^c . Consequently, with the deviant schedule, all of user A 's vehicles can complete their trips within the same time period as in

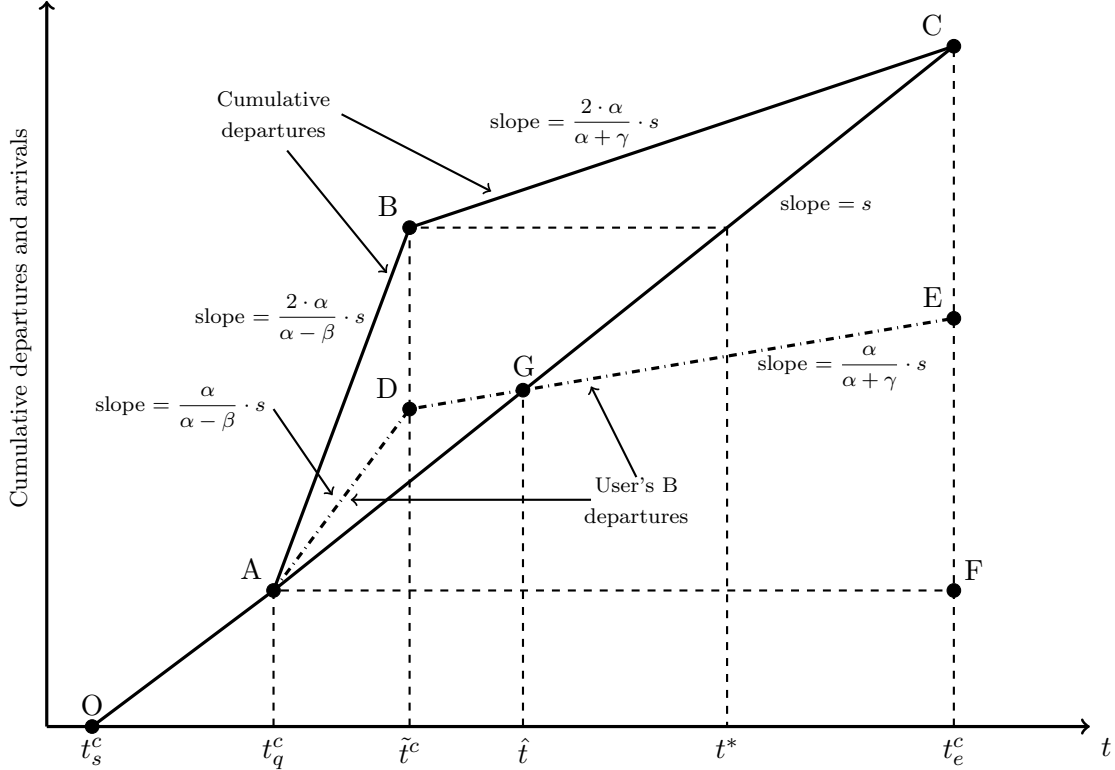


Figure 1: Candidate PSNE with $\gamma > \alpha$ and queuing.

the candidate PSNE, but without queuing. It is straightforward to show (see Appendix B) that user A 's fleet incurs lower total costs with the deviant schedule than with the candidate PSNE schedule. Hence the candidate PSNE is not a PSNE. *QED*

The reader may wonder why the candidate PSNE is robust to deviations in which a user reschedules a single vehicle, but not robust to rescheduling part of the fleet. The reason is that the first type of deviation comprises a zero measure of traffic whereas the second type comprises a positive measure. Rescheduling a positive measure of traffic can increase or decrease the MPC imposed by other vehicles in the fleet, and a judicious, coordinated, rescheduling of vehicles can result in a net reduction in fleet costs.

Proposition 1 establishes that a PSNE does not exist when $\gamma > \alpha$. As noted in the introduction, most empirical studies of scheduling preferences for automobile drivers have obtained estimates that satisfy this inequality. However, there is little evidence either on the trip-timing preferences of users that may control large shares of road traffic (e.g., freight shippers), or on preferences for travel by other modes of transportation. Daniel and Harback (2008) estimate that $\gamma < \alpha$ holds for many US airlines.⁹ Thus, it is of interest to

⁹They do not make explicit the role of the passengers' valuation of time.

study the existence and nature of PSNE when $\gamma \leq \alpha$.

3.2. Existence and Nature of PSNE with $\gamma \leq \alpha$

In this section we establish two results for the case $\gamma \leq \alpha$. First, we show that there exists a unique PSNE in the aggregate departure schedule that coincides with the social optimum (Proposition 2). Second, we show that the two users' individual departure rates are not uniquely defined in the PSNE (Proposition 3), and the users can incur different fleet costs (Section 3.2.3).

3.2.1. Existence of PSNE with $\gamma \leq \alpha$

Define $r_E \equiv \frac{2 \cdot \alpha \cdot (\beta + \gamma) + \gamma \cdot (\beta + 2 \cdot \gamma) - \sqrt{\beta^2 \cdot \gamma^2 + 4 \cdot \alpha \cdot (\beta + \gamma)^2 \cdot (\alpha + \gamma)}}{\gamma \cdot (\alpha + \gamma)} \cdot \frac{s}{2}$, $r_E \in \left(0, \frac{\gamma}{\alpha + \gamma} \cdot s\right)$.

Proposition 2. *If $\gamma \leq \alpha$, there exists a PSNE in which the aggregate departure schedule coincides with the social optimum. During late arrivals, each user's departure rate satisfies: $\frac{\gamma}{\alpha + \gamma} \cdot s \leq r_i(t) \leq \frac{\alpha}{\alpha + \gamma} \cdot s$, $t \in (t^*, t_e^o)$, $i = A, B$. During early arrivals, a sufficient condition for a PSNE is that each user's departure rate satisfies $r_i(t) \geq r_E$, $t \in (t_s^o, t^*)$, $i = A, B$.*

Proof: The proof entails establishing four results: 1) A PSNE with queuing does not exist. 2) All PSNE without queuing must coincide with the socially optimal departure pattern given by Eqs. (8), (9), and (10). 3) Given the lower bound on individual departure rates for late arrivals, neither user can gain by rescheduling a single vehicle. 4) Given the lower bounds on individual departure rates for early and late arrivals, neither user can gain by rescheduling part of its fleet.

Result 1. A PSNE with queuing does not exist.

By Lemma 2, mass departures cannot be part of a PSNE. Any candidate PSNE with queuing must satisfy conditions (20), (21), and (22). With $\gamma \leq \alpha$, these conditions cannot all be satisfied since the aggregate early departure rate exceeds capacity, and the aggregate late departure rate is no less than capacity. Hence any queue cannot dissipate while users are departing, which is inconsistent with a PSNE.

Result 2. The socially optimal departure pattern is the only possible PSNE in the aggregate departure schedule.

Departures must occur at rate s over a connected time interval since otherwise either user could reduce its fleet costs by rescheduling vehicles into “gaps” in the departure schedule when bottleneck capacity is not fully utilized. The departure period, $[t_s, t_e]$, must be as given by Eqs. (9) and (10) since otherwise trip costs at t_s and t_e would differ, and at least one user could reduce its fleet costs by rescheduling vehicles from the higher-cost endpoint to the lower-cost endpoint.

Result 3. A user cannot gain by rescheduling a single vehicle if each user's departure rate satisfies the conditions in Proposition 2.

In the candidate PSNE with a socially optimal aggregate departure pattern, there is no queuing but the bottleneck is used to capacity. As in the proof of Lemma 1, it is therefore necessary to distinguish between the cost saved by removing a vehicle from the departure schedule (which does not affect other vehicles' costs) and the cost of adding a vehicle (which creates a queue unless the vehicle is added at t_e^o). The respective costs are:

$$C_i^-(t) = \begin{cases} \beta \cdot (t^* - t), & t \in [t_s^o, t^*] \\ \gamma \cdot (t - t^*), & t \in [t^*, t_e^o] \end{cases},$$

$$C_i^+(t) = \begin{cases} \beta \cdot (t^* - t) + \frac{\alpha - \beta}{s} \cdot \int_t^{t^*} r_i(u) du + \frac{\alpha + \gamma}{s} \cdot \int_{t^*}^{t_e^o} r_i(u) du, & t \in [t_s^o, t^*] \\ \gamma \cdot (t - t^*) + \frac{\alpha + \gamma}{s} \cdot \int_t^{t_e^o} r_i(u) du, & t \in [t^*, t_e^o] \end{cases}.$$

A vehicle can be rescheduled in four ways: (i) late to late, (ii) late to early, (iii) early to late, and (iv) early to early. Consider each possibility in turn.

i. Rescheduling late to late: Rescheduling a late vehicle to a later time is never beneficial because the vehicle's trip cost increases, and other vehicles do not gain. Suppose a vehicle is rescheduled earlier from t to t' where $t^* \leq t' < t$. The change in fleet costs is given by Eq. (14):

$$\begin{aligned} \Delta C_i &= -C_i^-(t) + C_i^+(t') = -\gamma \cdot (t - t') + \frac{\alpha + \gamma}{s} \cdot \int_{t'}^t r_i(u) du \\ &\stackrel{s}{=} \lambda_{t',t}^i - \frac{\gamma}{\alpha + \gamma} \end{aligned}$$

where $\stackrel{s}{=}$ means identical in sign. Given $r_i(t) \geq \gamma \cdot s / (\alpha + \gamma)$ for $t \in (t^*, t_e^o)$, $\lambda_{t',t}^i \geq \gamma / (\alpha + \gamma)$, $\Delta C_i \geq 0$, and the deviation is not beneficial.

ii. Rescheduling late to early: Rescheduling a late vehicle to an early time is clearly inferior to rescheduling it to t^* because the vehicle incurs an early-arrival cost and creates a queue for a longer period. But rescheduling it to t^* is not beneficial as per case i.

iii. Rescheduling early to late: The best option in this case is to reschedule a vehicle from t_s^o . However, the gain is the same as for rescheduling a vehicle from t_e^o , and this is not beneficial as per case i. Rescheduling early to late therefore cannot be beneficial.

iv. Rescheduling early to early: The best option in this case is to reschedule a vehicle from t_s^o to t^* , but this, too, is not beneficial for the same reason as in case iii.

This establishes that the candidate PSNE in Proposition 2 is robust to deviations in which a single vehicle is rescheduled.

Result 4. A user cannot gain by rescheduling a positive measure of its fleet.

If user A reschedules vehicles to depart outside $[t_s^o, t_e^o]$, its fleet costs necessarily increase since the rescheduled vehicles experience greater schedule delay costs without benefiting the rest of the fleet. If user A instead reschedules vehicles to depart inside $[t_s^o, t_e^o]$, queuing

occurs. The optimal departure rate in the presence of a queue was derived in Section 3.1 and, although with $\gamma \leq \alpha$ a PSNE cannot exhibit queuing, we need to check if a deviation from the candidate to a setting with queuing is cost reducing. For early arrivals, user A is willing to depart at a positive and finite rate only if condition (19) is satisfied; i.e. $r_B(t) = \alpha \cdot s / (\alpha - \beta) > s$. Since user B departs at a rate less than s in the candidate PSNE, user A is better off scheduling the vehicles later. For late arrivals, user A is willing to depart at a positive and finite rate only if condition (17) is satisfied so that $r_B(t) = \alpha \cdot s / (\alpha + \gamma)$. Proposition 2 stipulates that $r_B(t) \leq \alpha \cdot s / (\alpha + \gamma)$. If $r_B(t) = \alpha \cdot s / (\alpha + \gamma)$, and therefore $r_A(t) = \gamma \cdot s / (\alpha + \gamma)$, user A is indifferent between departing or not so that the deviation does not reduce its fleet costs. If $r_B(t) < \alpha \cdot s / (\alpha + \gamma)$, user A is better off scheduling the vehicles later. The only remaining possibility for gainful deviation is one that involves mass departures. In Appendix C we prove that a deviation with mass departures is not gainful if the lower bounds on early and late departure rates stated in Proposition 2 are both satisfied. *QED*

The intuition for Proposition 2 is as follows. The only way for user A to gainfully deviate from the candidate PSNE by rescheduling a single vehicle is to advance its departure during the late arrival period. Doing so reduces the vehicle's late-arrival cost, but imposes queuing delay on user A 's vehicles that depart later. The tradeoff is not worthwhile if enough vehicles in A 's fleet have yet to depart, and queuing is sufficiently costly relative to late arrival. The lower bound on the late departure rate stated in Proposition 2 assures this condition is met. If user A deviates by rescheduling vehicles in mass departures and imposes queuing delays on its other vehicles, the same argument holds. The tradeoff is not worthwhile because the lower bound on the late departure rate ensures that there are enough vehicles in the fleet yet to depart that will be negatively affected.

To avoid delaying other vehicles in its fleet, user A must advance departures for vehicles that participate in a mass. Since vehicles in the mass suffer queuing delay, user A can benefit from such a deviation only if the vehicles' schedule delay costs are reduced enough. This requires that the vehicles were departing late over a period longer than the time they take to pass the bottleneck when in the mass. This, in turn, is possible only if user B occupies a large enough share of bottleneck capacity during the candidate PSNE. The lower bound on the late departure rate identified in Proposition 2 ensures that this condition is not met.

To see why a lower bound on the early departure rate is also required, suppose that user A does not depart during some time interval (t, t^*) . User A can then reschedule its vehicles departing in some interval (t^*, t') by launching them in a mass at a time $t_m \in (t, t^*)$. None of its fleet departing after t' will be delayed by the mass. If t_m is chosen to minimize the total schedule delay costs of vehicles in the mass, their total costs will fall. However, if user A is scheduling enough departures during (t_m, t^*) , the deviation will not be gainful since

the mass departure either imposes queuing delays on its other vehicles, or it has to include vehicles that were arriving early, which will suffer higher early arrival costs and queuing costs. If the early departure rate is high enough, such a mass departure is unfavorable. Proposition 2 identifies a minimum early departure rate to guarantee this when the late departure rate is held fixed at its minimum value: $\gamma \cdot s / (\alpha + \gamma)$.

3.2.2. Non-uniqueness of PSNE with $\gamma \leq \alpha$

Although the aggregate departure schedule in the PSNE described in Proposition 2 is unique when $\gamma \leq \alpha$, many pairs of departure schedules $\{r_A(\cdot), r_B(\cdot)\}$ are consistent with the aggregate pattern. Thus, the PSNE is not unique in terms of individual departure rates. This result is formalized in the following proposition:

Proposition 3. *During the late departure period $t \in (t^*, t_e^o)$, the two users depart at the same rate $r_A(t) = r_B(t) = s/2$ if $\alpha = \gamma$, but they can depart at different rates if $\gamma < \alpha$. During the early departure period $t \in (t_s^o, t^*)$, a continuum of departure schedules $\{r_A(\cdot), r_B(\cdot)\}$ are consistent with the PSNE.*

Proof: By Proposition 2, for $t \in (t^*, t_e^o)$, $\gamma / (\alpha + \gamma) \leq r_A(t) / s \leq \alpha / (\alpha + \gamma)$ and $r_B(t) = s - r_A(t)$. If $\gamma = \alpha$, $\gamma / (\alpha + \gamma) = 1/2$ and therefore $r_A(t) = r_B(t) = s/2$. If $\gamma < \alpha$, there is a continuum of pairs of $\{r_A(t), r_B(t)\}$ that satisfy the equilibrium condition while assuring that all vehicles depart during the interval $[t_s^o, t_e^o]$. For $t \in (t_s^o, t^*)$, the constraint $r_i(t) \geq r_E$ is less strict because $r_E < s/2$. Any pair of departure schedules $\{r_A(\cdot), r_B(\cdot)\}$ that satisfies $r_E \leq r_A(t) \leq s - r_E$ and $r_B(t) = s - r_A(t)$ satisfies the aggregate PSNE condition. *QED*

The non-uniqueness of individual departure rates in a PSNE implies that, unlike in the model for identical non-atomic users, users can experience different fleet costs in a PSNE. This prospect is examined further in the next subsection.

3.2.3. Asymmetric equilibrium costs with $\gamma \leq \alpha$

Clearly, if users A and B depart at the same rate throughout the departure period they incur the same fleet costs. In addition, there are many asymmetric PSNE departure schedules that result in the same average schedule delay costs for the two users and hence the same fleet costs. However, there are also many asymmetric departure schedules that result in different fleet costs.

Figure 2 depicts an illustrative example of a PSNE in which user A incurs lower fleet costs than user B . In the example, user A 's departures are concentrated near t^* so that user A has lower average schedule delay costs than B . User A 's departure rate is shown by solid lines and user B 's by broken lines. During the interval (t_s^o, t_{BA}) , user A departs at the minimum rate r_E defined in Proposition 2. During the next interval (t_{BA}, t^*) , user

B departs at rate r_E and user A at the complementary rate $s - r_E$. During the first part (t^*, t_{AB}) of the late-arrival period, user A departs at the maximum rate consistent with a PSNE for user B , $\alpha \cdot s / (\alpha + \gamma)$. During the remaining part (t_{AB}, t_e^o) of the departure period, user A departs at the minimum rate consistent with a PSNE for itself, $\gamma \cdot s / (\alpha + \gamma)$. The transition times t_{BA} and t_{AB} are such that each user dispatches a total of $N/2$ vehicles. Expressed in terms of user A 's fleet, the requisite condition is

$$r_E \cdot (t_{BA} - t_s^o) + (s - r_E) \cdot (t^* - t_{BA}) + \frac{\alpha}{\alpha + \gamma} \cdot s \cdot (t_{AB} - t^*) + \frac{\gamma}{\alpha + \gamma} \cdot s \cdot (t_e^o - t_{AB}) = \frac{N}{2}. \quad (31)$$

Total costs in the PSNE, TC^o , are a given; i.e. independent of users A 's and B 's individual departure rates. The maximum difference between the users' costs, $TC_B - TC_A$, can therefore be found by minimizing TC_A with respect to t_{BA} and t_{AB} . As described in Appendix D, user A 's costs as a fraction of total costs work out to

$$f \equiv \frac{TC_A}{TC^o} = \frac{3 + z}{4} - \frac{1}{2} \sqrt{\frac{\alpha}{\alpha + \gamma} + z^2},$$

where $z \equiv \beta \cdot \gamma / ((\alpha + \gamma) \cdot (\beta + \gamma))$. The fraction f depends on α , β , and γ only through the ratios $\beta/\alpha < 1$ and $\gamma/\alpha \leq 1$. It is readily shown that f is a monotonically increasing function of β/α and γ/α . At the upper limits with $\beta/\alpha \cong 1$ and $\gamma/\alpha = 1$, $f \cong 0.4535$. As β/α and γ/α approach their lower limits of 0, f approaches $1/4$.

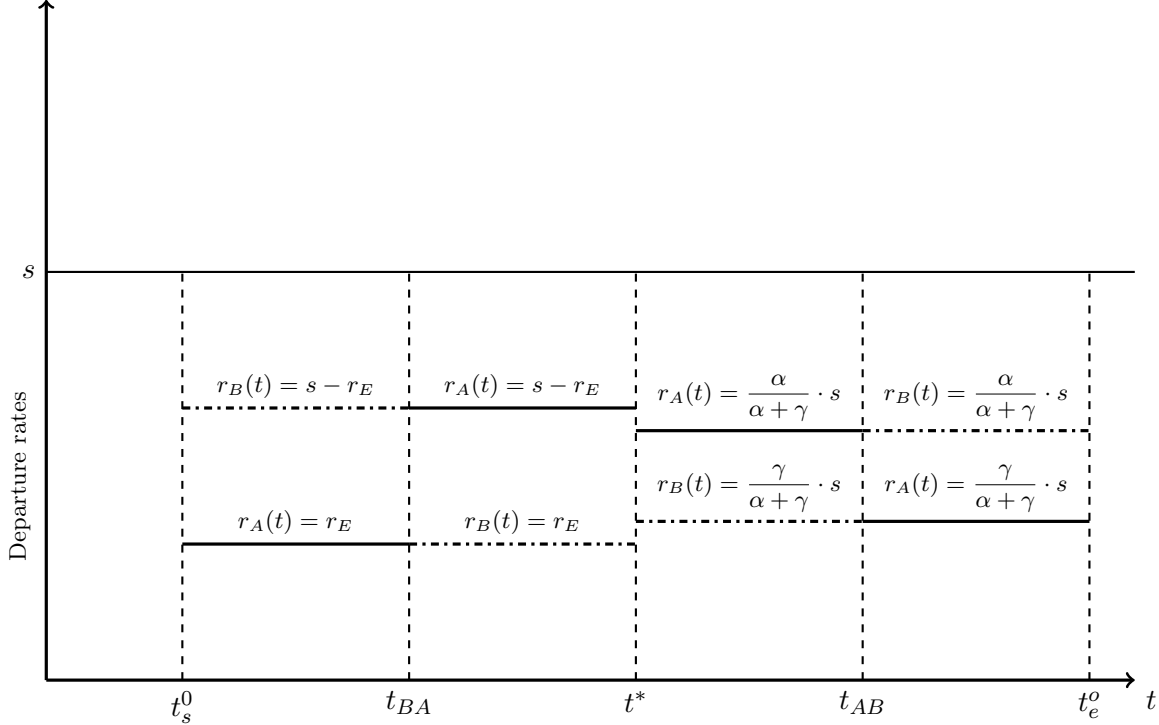


Figure 2: A PSNE with asymmetric costs: $\gamma \leq \alpha$, no queuing.

This example illustrates that although there exists a unique PSNE in terms of the aggregate departure rate, individual user departure rates can differ substantially and so can their costs. If the unit schedule delay cost parameters, β and γ , are small compared to the cost of travel time, α , one user's fleet costs can be as little as one third the other user's costs. This is because one user can concentrate its departures around t^* without the other user wanting to reschedule its fleet because the gain from reducing schedule delay costs would be outweighed by the high costs of queuing delay. This suggests that equity of access to a bottleneck can be an issue with atomic users.

3.3. Extension to multiple users

The analysis in this section can be generalized to $m > 2$ users. Propositions 1 and 2 can be extended in a straightforward manner by following similar lines of reasoning. For example, condition (15) for existence of a PSNE without queuing, $\lambda_{t,t'}^i \geq \gamma / (\alpha + \gamma)$, still holds. In the least restrictive case in which all m users depart at equal rates, this condition implies that $1/m \geq \gamma / (\alpha + \gamma)$, or $(m - 1) \cdot \gamma \leq \alpha$. The intuition is similar to the case with $m = 2$. In a symmetric candidate PSNE, each user departs at rate s/m and contributes a fraction $1/m$ of the traffic. If user i reschedules one vehicle ΔT units of time earlier, it reduces the vehicle's late arrival cost by $\gamma \cdot \Delta T$. The vehicle imposes a delay of $1/s$ on the $(s/m) \cdot \Delta T$ of user i 's other vehicles that depart during the time interval. The unit cost of the delay is $\alpha + \gamma$. User i benefits from rescheduling the vehicle unless $\gamma \cdot \Delta T \leq (\alpha + \gamma) \cdot \Delta T / m$, which is equivalent to $(m - 1) \cdot \gamma \leq \alpha$.

The nonexistence of a PSNE with queuing can also be established using the same approach for $m > 2$ as for $m = 2$. Candidate equilibrium departure rates for other users during early and late departures are still given by Eqs. (19) and (17) respectively. The candidate equilibrium is well-defined and unique, and it breaks down because any user can reduce its costs by rescheduling a portion of its fleet. Clearly, the condition $(m - 1) \cdot \gamma \leq \alpha$ becomes more stringent as m increases so that a PSNE without queuing becomes progressively less plausible. Indeed, given any fixed values of $\gamma > 0$ and $\alpha > 0$, condition $(m - 1) \cdot \gamma \leq \alpha$ necessarily fails if m is large enough. Thus, as $m \rightarrow \infty$ the bottleneck model with atomic users does not converge in behavior to the non-atomic model which has a unique PSNE. The reason for this divergence is that even for large values of m , users in the atomic model take into account how they affect queuing times. By contrast, in the standard bottleneck model users treat queuing times as given.

4. Nonuniqueness of equilibrium with heterogeneous atomic users

In this section we consider a modified version of the model in which users A and B differ in their preferred arrival times and can have fleets of different size. Arnott et al. (1987)

studied this setting for the case of non-atomic users and we draw on some of their results. We show that, regardless of the relative size of parameters α and γ , a multiplicity of PSNE without queuing can exist with two atomic users if preferred arrival times differ within a certain range. The PSNE differ both in the timing of aggregate departures and total costs. Departures can begin earlier than, later than, or at the same time as in the social optimum. Thus, a PSNE can be inefficient even if there is no queuing.

Suppose user i , $i = A, B$, has a fleet of N_i vehicles, each with a preferred arrival time of t_i^* . Assume without loss of generality that $t_B^* > t_A^*$. To begin, assume $t_B^* \gg t_A^*$ so that the users can schedule their fleets at individually optimal time windows that do not overlap. Let t_{is} and t_{ie} denote the first and last departure times for user i , and let superscript d denote the PSNE with disjoint arrival times. Each user departs at a rate s during the interval that minimizes the total schedule delay costs of its fleet:

$$\left(t_{As}^d, t_{Ae}^d\right) = \left(t_A^* - \frac{\gamma}{\beta + \gamma} \cdot \frac{N_A}{s}, t_A^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s}\right), \quad (32)$$

$$\left(t_{Bs}^d, t_{Be}^d\right) = \left(t_B^* - \frac{\gamma}{\beta + \gamma} \cdot \frac{N_B}{s}, t_B^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N_B}{s}\right). \quad (33)$$

To rule out this uninteresting case we hereafter assume $t_{Bs}^d < t_{Ae}^d$, which is equivalent to assuming that:

$$t_B^* - t_A^* < \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s} + \frac{\gamma}{\beta + \gamma} \cdot \frac{N_B}{s}. \quad (34)$$

Define the auxiliary variable

$$x \equiv \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s} + \frac{\gamma}{\beta + \gamma} \cdot \frac{N_B}{s} - (t_B^* - t_A^*) > 0. \quad (35)$$

Variable x measures how much the users' preferred departure schedules overlap, and thus the degree of "conflict" between them.

In addition to condition (34) we assume

$$t_B^* - t_A^* > \left\| \frac{\gamma}{\beta + \gamma} \cdot \frac{N_B}{s} - \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s} \right\|. \quad (36)$$

Condition (36) assures that in the social optimum, derived next, some of user A 's fleet arrives late and some of user B 's fleet arrives early.

4.1. Social optimum

In the social optimum, vehicles depart at an aggregate rate of s over a connected time interval. In general, the optimal individual departure rates, $r_A^*(t)$ and $r_B^*(t)$, are not unique. However, given condition (36), there is a unique optimum in which all of user A 's fleet is dispatched before any of user B 's fleet. Some of user A 's fleet arrives late, and some of user B 's fleet arrives early. Let t_{As} be the time at which user A starts to depart.

User A departs over the interval $(t_{As}, t_{As} + N_A/s)$, and user B departs over the interval $(t_{As} + N_A/s, t_{As} + (N_A + N_B)/s)$. Total costs are:

$$TC = \frac{\beta \cdot s}{2} \cdot (t_A^* - t_{As})^2 + \frac{\gamma \cdot s}{2} \cdot \left(t_{As} + \frac{N_A}{s} - t_A^* \right)^2 + \frac{\beta \cdot s}{2} \cdot \left(t_B^* - \left(t_{As} + \frac{N_A}{s} \right) \right)^2 + \frac{\gamma \cdot s}{2} \cdot \left(t_{As} + \frac{N_A + N_B}{s} - t_B^* \right)^2. \quad (37)$$

The first-order condition for minimizing TC with respect to t_{As} yields:¹⁰

$$t_{As}^o = \frac{t_A^* + t_B^*}{2} - \frac{\gamma}{2 \cdot (\beta + \gamma)} \cdot \frac{N_A + N_B}{s} - \frac{1}{2} \cdot \frac{N_A}{s}, \quad (38)$$

where superscript o refers to the social optimum. Given condition (36), $t_A^* < t_{Ae}^o < t_B^*$.

Substituting Eq. (38) into Eq. (37), total costs in the social optimum can be written as:

$$TC^o = \frac{\delta}{2} \cdot \frac{N_A^2 + N_B^2}{s} + \frac{\beta + \gamma}{4} \cdot s \cdot x^2, \quad (39)$$

where x is defined in Eq. (35). The first term in Eq. (39) equals total costs if condition (34) did not hold and the two users traveled in disjoint intervals. The second term in Eq. (39) is the additional costs incurred due to overlap in the users' preferred arrival times. This term is an increasing, quadratic function of the degree of conflict, x .

4.2. No-toll equilibrium with heterogeneous non-atomic users

Before considering the PSNE with users A and B , we briefly discuss the analogous no-toll PSNE with heterogeneous non-atomic users. In this case there is a measure N_A of non-atomic users with a preferred arrival time t_A^* , and a measure N_B of non-atomic users with a preferred arrival time t_B^* . Arnott et al. (1987) show that when conditions (34) and (36) both hold, the equilibrium queuing pattern has two peaks. The first peak corresponds to on-time arrival for users with preferred arrival time t_A^* , and the second peak to on-time arrival for users with preferred arrival time t_B^* . Total costs are

$$TC^n = \delta \cdot \frac{N_A^2 + N_B^2}{s} + \frac{\beta N_A + \gamma N_B}{2} \cdot x. \quad (40)$$

The first term in (40) gives total costs if the equilibrium departure schedules of the two user groups do not conflict. This term is twice the first term in (39), just as total costs with homogeneous users in Eq. (7) are twice total costs in the social optimum in Eq. (11). The second term in (40) is a linear function of x . This contrasts with the second term in (39) which is a quadratic function of x . Thus, a small degree of conflict between the two user groups raises total costs disproportionately more in the no-toll equilibrium than the

¹⁰This formula is derived in Arnott et al. (1987).

social optimum. In the next subsection we use the difference between total costs in the no-toll equilibrium and social optimum as a metric to assess the efficiency achieved from self-internalization of congestion by atomic users.

4.3. PSNE

A PSNE without queuing exists if $t_{Ae} < t_B^*$ since the two users then do not arrive late at the same time and do not compete to complete their trips as soon as possible. Indeed, a continuum of PSNE exists. To see this, assume that both users depart in the period $[t_A^*, t_{Ae}]$ and capacity is fully used. Suppose user A advances a vehicle's departure from t_2 to t_1 , where $t_A^* \leq t_1 < t_2 \leq t_{Ae}$. User A 's fleet costs change by:

$$\begin{aligned} \Delta C_A &= -C_A^-(t_2) + C_A^+(t_1) = -\gamma \cdot (t_2 - t_A^*) + \gamma \cdot (t_1 - t_A^*) + \frac{\alpha + \gamma}{s} \cdot \int_{t_1}^{t_2} r_A(u) du \\ &= ((\alpha + \gamma) \cdot \lambda_{t_1, t_2}^A - \gamma) \cdot (t_2 - t_1). \end{aligned}$$

User A does not benefit from the rescheduling if

$$\lambda_{t_1, t_2}^A \geq \frac{\gamma}{\alpha + \gamma}.$$

This condition is satisfied for any choice of t_1 and t_2 as long as:

$$r_A(t) \in \left[\frac{\gamma}{\alpha + \gamma} \cdot s, s \right], \quad t \in [t_A^*, t_{Ae}]. \quad (41)$$

User B 's departure rate is then $r_B(t) = s - r_A(t) \in [0, \alpha \cdot s / (\alpha + \gamma)]$. There is no restriction on $r_B(t)$ because user B arrives early throughout $[t_A^*, t_{Ae}]$. Thus, any departure profile satisfying Eq. (41) is consistent with a PSNE. However, different departure profiles generally result in different total costs.

An exhaustive treatment of all PSNE in this example would be tedious. Attention is limited to two illustrative cases. In Case 1, user A schedules its entire fleet during the same interval it would choose if user B did not exist. Hence, $r_A(t) = s$ during the interval (t_{As}^d, t_{Ae}^d) given by Eq. (32). User B departs at rate s during the ensuing interval (t_{Ae}^d, t_{Be}) . In Case 2, user A departs at rate s during the interval (t_{As}, t_A^*) . During the next interval (t_A^*, t_{Ae}) , user A departs at the minimum rate consistent with Eq. (41), $r_A(t) = \gamma \cdot s / (\alpha + \gamma)$. User B departs at the complementary rate $r_B(t) = \alpha \cdot s / (\alpha + \gamma)$. During the final interval (t_{Ae}, t_{Be}) , user B departs at rate s . Cases 1 and 2 are now examined in turn.

4.3.1. PSNE for Case 1

Figure 3 depicts a PSNE conforming with Case 1.¹¹

¹¹Although there is no queuing in the PSNE, the value of parameter α is relevant because it determines the maximum rate at which User B can depart during the central period.

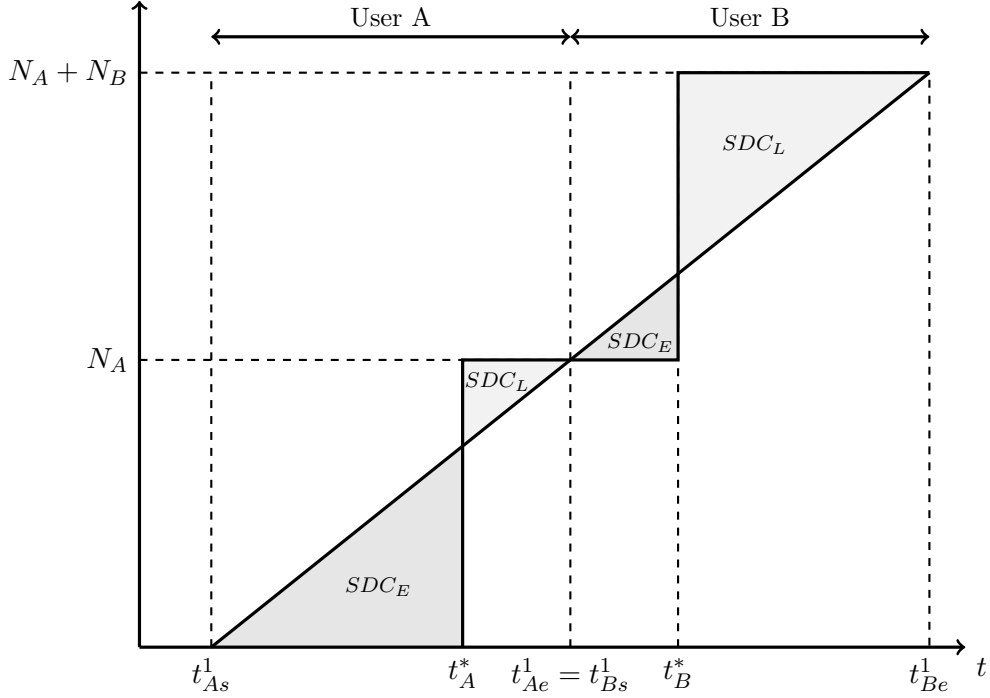


Figure 3: PSNE for Case 1 with $t_B^* > t_A^*$, $N_A = 1$, $N_B = 1$, $s = 1$, $\alpha = 2$, $\beta = 1$, $\gamma = 2$, $t_A^* = 0$, and $t_B^* = 2/3$.

Since all of user A 's fleet departs before any of user B 's fleet,

$$t_{Ae}^1 - t_{As}^1 = \frac{N_A}{s}, \quad (42)$$

where superscript 1 denotes Case 1. User A can schedule a vehicle either just before t_{As}^1 or just after t_{Ae}^1 without imposing a congestion delay on other vehicles in its fleet. The first and last vehicles must therefore incur equal schedule delay costs:

$$\beta \cdot (t_A^* - t_{As}^1) = \gamma \cdot (t_{Ae}^1 - t_A^*). \quad (43)$$

Together, Eqs. (42) and (43) imply:

$$t_{As}^1 = t_A^* - \frac{\gamma}{\beta + \gamma} \cdot \frac{N_A}{s}, \quad (44)$$

$$t_{Ae}^1 = t_A^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s}. \quad (45)$$

As noted above, the departure interval (t_{As}^1, t_{Ae}^1) coincides with the interval (t_{As}^d, t_{Ae}^d) given by Eq. (32). User B departs in a connected time interval immediately after user A . Hence,

$$t_{Bs}^1 = t_{Ae}^1 = t_A^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s}, \quad (46)$$

$$t_{Be}^1 = t_A^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s} + \frac{N_B}{s}. \quad (47)$$

A notable feature of the departure schedule given by Eqs. (44)-(47) is that it does not depend on t_B^* . To be consistent with the assumption that some of user B 's fleet arrives early we require $t_{Ae}^1 < t_B^*$, or:

$$t_B^* - t_A^* > \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s}. \quad (48)$$

Case 1 is a PSNE if neither user has an incentive to deviate from it. User A clearly has no incentive to deviate because its fleet costs are the minimum possible. Two conditions must be satisfied for user B . First, user B cannot gain by rescheduling its last vehicle from t_{Be}^1 to some time \hat{t} between t_{As}^1 and t_{Ae}^1 . This condition is satisfied since rescheduling would impose queuing delay on all vehicles that depart after \hat{t} without reducing schedule delay costs for user B 's fleet.

Second, user B cannot gain by rescheduling its last vehicle from t_{Be}^1 to just before the beginning of the travel period at t_{As}^1 . The condition is $\beta \cdot (t_B^* - t_{As}^1) > \gamma \cdot (t_{Be}^1 - t_B^*)$, which reduces to:

$$t_B^* - t_A^* \geq \frac{\gamma}{\beta + \gamma} \cdot \frac{N_B}{s}. \quad (49)$$

Conditions (48) and (49) together guarantee that condition (36) is satisfied.

The timing of departures in Case 1 and the social optimum can be compared using Eqs. (44) and (38):

$$t_{As}^1 - t_{As}^o = \frac{1}{2} \cdot \left(\frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s} + \frac{\gamma}{\beta + \gamma} \cdot \frac{N_B}{s} - (t_B^* - t_A^*) \right) = \frac{x}{2}.$$

Since $x > 0$, $t_{As}^1 > t_{As}^o$: departures in Case 1 begin later than in the social optimum. To see why, note that while early and late arrivals are balanced optimally for user A , user B 's arrivals begin inefficiently late because user A occupies the bottleneck until t_{Ae}^1 and effectively squeezes user B out. As shown in Appendix E, total costs in Case 1 can be written as:

$$TC^1 = \frac{\delta}{2} \cdot \frac{N_A^2 + N_B^2}{s} + \frac{\beta + \gamma}{2} \cdot s \cdot x^2 = TC^o + \frac{\beta + \gamma}{4} \cdot s \cdot x^2. \quad (50)$$

Total costs are an increasing quadratic function of the degree of conflict, x . This is similar to the social optimum (cf. Eq. (39)), but unlike the non-atomic equilibrium where the dependence is linear (cf. Eq. (40)). The last expression in Eq. (50), $(\beta + \gamma) s x^2 / 4$, measures inefficiency of the PSNE in Case 1. This inefficiency can be expressed as a price of anarchy, PA , using the ratio

$$PA = \frac{TC^1 - TC^o}{TC^o}.$$

PA is bounded below by 0. As shown in Appendix E, it is bounded above by 1/5. Thus, despite the fact that there is no queuing in the PSNE for Case 1, total costs can exceed the socially optimal level by up to 20 percent.

Efficiency of the PSNE can also be measured relative to the non-atomic equilibrium using the index

$$w \equiv \frac{TC^n - TC^1}{TC^n - TC^o}. \quad (51)$$

The numerator of (51) is the reduction in total costs achieved by self-internalization of congestion by the atomic users. The denominator is the reduction in total costs realized at the social optimum. Index w reaches its maximum value of 1 when $x = 0$. Appendix E establishes that it is bounded below by $6/7$. Self-internalization of congestion by the atomic users independently therefore achieves most of the potential benefits from coordinating departure times centrally.

4.3.2. PSNE for Case 2

Figure 4 depicts an example of a PSNE in Case 2 using the same parameter values as for Figure 3. During the initial interval, (t_{As}^2, t_A^*) , user A departs at rate s . User B begins to depart immediately after t_A^* , and during the interval (t_A^*, t_{Ae}^2) the two users depart simultaneously with an aggregate rate of s . User A departs at rate $r_A(t) = \gamma \cdot s / (\alpha + \gamma)$, and user B departs at rate $r_B(t) = \alpha \cdot s / (\alpha + \gamma)$. During the final interval, (t_{Ae}^2, t_{Be}^2) , user B departs at rate s .

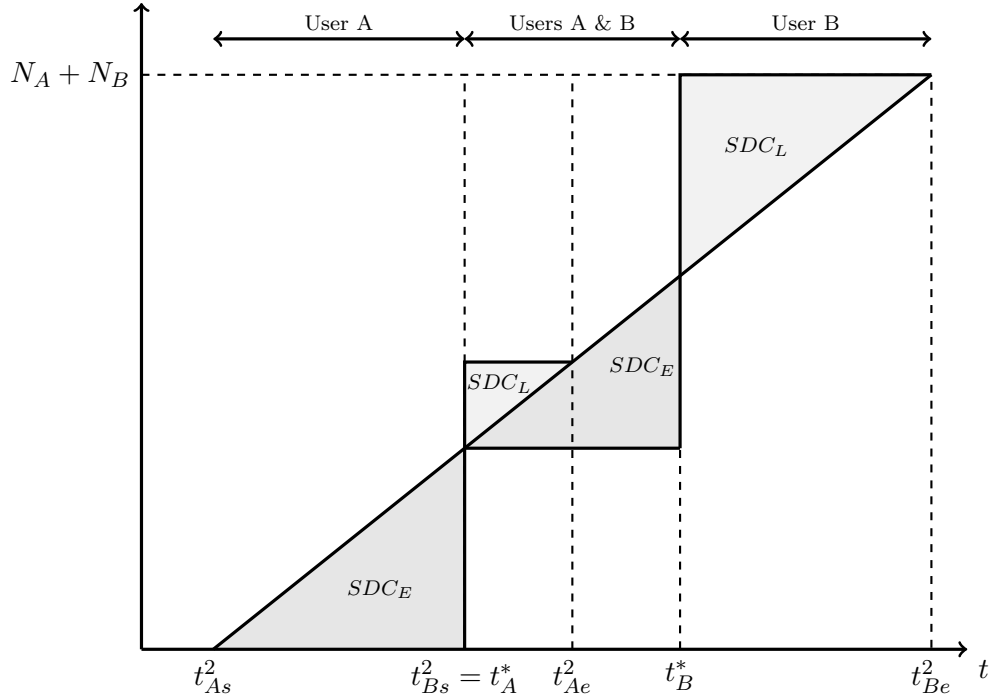


Figure 4: PSNE for Case 2 with $t_B^* > t_A^*$, $N_A = 1$, $N_B = 1$, $s = 1$, $\alpha = 2$, $\beta = 1$, $\gamma = 2$, $t_A^* = 0$, and $t_B^* = 2/3$.

The PSNE for Case 2 is solved in Appendix E. The three transition times work out to:

$$t_{As}^2 = t_A^* - \frac{\alpha + \gamma}{\alpha + \beta + \gamma} \cdot \frac{N_A}{s}, \quad (52)$$

$$t_{Ae}^2 = t_{Bs}^2 = t_A^* + \frac{\beta \cdot (\alpha + \gamma)}{\gamma \cdot (\alpha + \beta + \gamma)} \cdot \frac{N_A}{s}, \quad (53)$$

$$t_{Be}^2 = t_A^* + \frac{\beta}{\alpha + \beta + \gamma} \cdot \frac{N_A}{s} + \frac{N_B}{s}. \quad (54)$$

Similar to Case 1, the timing of the PSNE given by Eqs. (52)-(54) does not depend on t_B^* . As shown in Appendix E, departures in the PSNE of Case 2 can begin either earlier or later than in the social optimum. As t_B^* decreases, the social optimum begins earlier (cf. Eq. (38)), whereas the PSNE does not (cf. Eq. (52)). Hence, as user B 's preferred schedule moves closer to user A 's and the conflict between them grows, the PSNE in Case 2 shifts from beginning too early to beginning too late.

Unlike for Case 1, there is no simple formula for total costs in Case 2. Also unlike Case 1, total costs do not reach the socially optimal minimum value at the point where t_{As}^2 coincides with t_{As}^o . Thus, the price of anarchy is always positive. This is because both users depart during the interval (t_A^*, t_{Ae}^2) , with user A arriving late and user B arriving early. To see that simultaneous departure is inefficient, consider a vehicle of user B arriving at t_1 and a vehicle of user A arriving at $t_2 > t_1$. If the two vehicles exchanged slots, user A 's vehicle would arrive less late and user B 's vehicle would arrive less early. The two users might agree to such an exchange if they cooperated, but the switch cannot occur in a Nash equilibrium.

We conclude this section with a numerical example featuring the same parameter values as those used to construct Figures 3 and 4: $N_A = 1$, $N_B = 1$, $s = 1$, $\alpha = 2$, $\beta = 1$, $\gamma = 2$, $t_A^* = 0$, and $t_B^* = 2/3$. Departures and arrivals last for two hours. The social optimum begins at $t_{As}^o = -5/6$, and total costs are $TC^o = 3/4$. The PSNE for Case 1 begins at $t_{As}^1 = -2/3$, and total costs are $TC^1 = 5/6$ which is 11.1% higher than TC^o . User A 's total costs are $1/3$, and user B 's total costs are $1/2$. User B 's total costs are 50% higher than user A 's total costs even though the two users have the same size of fleet and the same unit costs α , β , and γ . The PSNE for Case 2 begins at $t_{As}^2 = -4/5$, and total costs are $TC^2 = 0.81$ which is 8.4% higher than TC^o but 2.4% lower than TC^1 .¹² User A 's total costs are $2/5$, and user B 's total costs are 0.413 . User B 's total costs are just 3.3% higher than user A 's total costs so that the PSNE in Case 2 is both more efficient and more equitable than in Case 1.

¹²If user B 's desired arrival time is increased from $t_B^* = 2/3$ to $t_B^* = 4/5$, t_{As}^o increases to $-23/30$ while t_{As}^2 remains at $-4/5$. The PSNE in Case 2 then begins later than the social optimum.

5. Conclusion

In this paper we have explored the existence and uniqueness of pure strategy Nash equilibrium (PSNE) in the bottleneck model with atomic users. We consider a simple case featuring two users with piecewise linear and independent trip-timing preferences for each vehicle in their fleets. We show that if users are identical, and $\gamma > \alpha$, a PSNE does not exist. By contrast, if $\gamma \leq \alpha$, multiple PSNE do exist in which no queuing occurs. The aggregate departure profile is unique. However, there exists a continuum of departure-time schedules for the two users that differ in the share of total costs borne by each user. We also consider a setting in which the two users differ in their preferred arrival times, t_A^* and t_B^* . For a range of values of $t_B^* - t_A^*$, there exists a continuum of PSNE that differ in the timing of trips, total costs, and the relative burden of costs borne by each user.

The potential nonexistence of PSNE with atomic users is, we believe, the most significant of the results. In the standard bottleneck model a PSNE exists under relative general assumptions. It fails to exist only if trip-timing preferences have discontinuities such as discrete penalties for late arrival or discontinuities in the cost of travel time as a function of duration (Lindsey, 2004). In the model with non-atomic users nonexistence is a problem even with smooth preferences because atomic users face conflicting incentives. On the one hand they prefer to spread out departures of their fleets in order to avoid self-imposed queuing delays. If there is only one user this results in a socially-optimal departure schedule with no queuing. But when other users are present, an atomic user has an incentive to force vehicles into the departure stream near the peak in order to reduce its fleet's schedule delay costs. A user's best response can be to preempt other users by scheduling a mass departure near t^* and hogging bottleneck capacity. Yet mass departures are inconsistent with PSNE, as we have shown.

The out-of-equilibrium behaviour in the bottleneck model bears some resemblance to firm behaviour in one-dimensional (Hotelling) spatial competition models.¹³ In location-choice games with parametric prices firms leapfrog each other to gain market share. The outcome depends on the number of firms. With two or four firms a unique PSNE exists. But with five firms there are multiple PSNE, and with three firms a PSNE does not exist. In simultaneous location- and price-choice games with two firms a PSNE does not exist. Firms face conflicting location-choice incentives that are broadly similar to the incentives of atomic users picking departure times in the bottleneck model. Moving closer to the centre (analogous to departing close to t^*) gains firms market share, but moving away softens price competition.

Preliminary analysis suggests that the difference between the standard bottleneck model

¹³See Anderson et al. (1992).

and the model with atomic users remains as the number of atomic users increases because each user still recognizes how its scheduling choices affect queuing delays. Thus, simple congestion-prone systems with atomic users can exhibit fundamentally different dynamic behavior than with non-atomic users. Dynamic systems with non-atomic users have been well studied in the literature using the bottleneck model as well as various flow-congestion models. Yet, to our knowledge, this paper is the first to examine the existence and uniqueness of equilibrium in dynamic systems with atomic users that make scheduling decisions simultaneously.

Our analysis is exploratory and can be extended in various directions. One obvious priority is to determine how the bottleneck model can be modified to restore existence of equilibrium without restricting parameter values. Possibilities include: extension of the solution concept to mixed strategies or sequential decision-making (e.g., Stackelberg equilibrium), and a more general specification of scheduling preferences based on activity analysis.

Another direction for future research is to investigate more fully the benefits of coordination or other forms of cooperation. Self-internalization of congestion externalities by an atomic agent is one form of coordination, and its potential for welfare gains is apparent from the examples in Section 4. Cooperation between atomic agents can also be beneficial as shown by the examples in which the price of anarchy is positive. Major freight shippers might cooperate by harmonizing their delivery schedules although doing so without triggering concerns about illegal collusion might be tricky. Governments can also implement or encourage cooperative behavior. One example are staggered work hours and flextime programs that entail coordination on target arrival times (i.e., t^*). Staggered work hours were introduced in the 1970s in Washington, D.C., and Ottawa, Canada. Gutiérrez-i Puigarnau and Van Ommeren (2012) report that one in three German firms uses staggered working hours. Mun and Yonekawa (2006) note that in Japan, in 1998, 8% of employees worked at a firm with flexible working hours, and in the US, in 1994-1997, less than 6% of employees had a formal flexible working arrangement, but 28% of all full-time workers varied their working times to some degree.

A recent example of a government-sponsored program is the Dutch Spitsmijden experiment in which drivers are offered rewards for avoiding peak-hour travel by retiming their trips, choosing another mode, or teleworking (Ben-Elia and Ettema, 2009, 2011). The Spitsmijden experiment differs from other travel-demand measures such as road pricing in that participation is voluntary. Large employers in the test regions have shown interest in participating as a way to curb congestion and to attract or retain employees (Ben-Elia and Ettema, 2009). However, it is an open question whether behavioral changes induced by a reward system will persist after a program is terminated (Ben-Elia and Ettema, 2011).

Acknowledgments

We are very grateful to Erik Verhoef for his constructive and incisive comments which have significantly improved the paper. Financial support from ERC Advanced Grant OPTION (#246969) is gratefully acknowledged. We would also like to thank participants at the Kuhmo Nectar Conference on Transportation Economics 2013 at Northwestern University. Any errors are ours.

Appendix A. Proof of Lemma 2

Suppose that user A schedules a mass departure at time t . Clearly, user B will not depart either at t or immediately after t since it would be better to depart either just before the mass or some time after t when the queue caused by the mass departure has diminished or disappeared. To complete the proof we will show that user A will never schedule a mass departure at t unless user B is scheduling vehicles either at t or immediately after t .

Thus, suppose that user A is considering a mass departure of M vehicles at time t , and assume that user B does not depart during the period $[t, \check{t}]$, where $\check{t} > t$. Variables M , t , and \check{t} are held fixed throughout the proof. The logic of the proof depends on whether or not a queue exists at t .

Appendix A.1. A queue exists at t

Assume $Q(t) > 0$. Pick any time $t' \in (t, \check{t}]$ such that $Q(t') > 0$. We will show that user A is better off postponing the mass departure from t to t' . Let $m \in [0, M]$ index vehicles in the order they are positioned in the mass, and let $c(m, t)$ denote the cost incurred by vehicle m when the mass departs at t :

$$c(m, t) = \alpha \cdot \left(T(t) + \frac{m}{s} \right) + \begin{cases} \beta \cdot \left(t^* - t - T(t) - \frac{m}{s} \right) & \text{if } t + T(t) + \frac{m}{s} \leq t^* \\ \gamma \cdot \left(t + T(t) + \frac{m}{s} - t^* \right) & \text{if } t + T(t) + \frac{m}{s} \geq t^* \end{cases} .$$

If the mass is rescheduled to depart at t' instead, vehicle m incurs a cost:

$$c(m, t') = \alpha \cdot \left(T(t') + \frac{m}{s} \right) + \begin{cases} \beta \cdot \left(t^* - t' - T(t') - \frac{m}{s} \right) & \text{if } t' + T(t') + \frac{m}{s} \leq t^* \\ \gamma \cdot \left(t' + T(t') + \frac{m}{s} - t^* \right) & \text{if } t' + T(t') + \frac{m}{s} \geq t^* \end{cases} .$$

Now, vehicle m arrives at the same time whether the mass is scheduled at t or t' since a queue remains at t' and user B does not depart between t and t' . Hence $t' + T(t') = t + T(t)$, and $c(m, t')$ can be written:

$$\begin{aligned} c(m, t') &= \alpha \cdot \left(T(t) + \frac{m}{s} \right) - \alpha \cdot (t' - t) + \begin{cases} \beta \cdot \left(t^* - t - T(t) - \frac{m}{s} \right) & \text{if } t + T(t) + \frac{m}{s} \leq t^* \\ \gamma \cdot \left(t + T(t) + \frac{m}{s} - t^* \right) & \text{if } t + T(t) + \frac{m}{s} \geq t^* \end{cases} \\ &= c(m, t) - \alpha \cdot (t' - t) < c(m, t). \end{aligned}$$

Postponing the mass departure therefore reduces costs for all M vehicles in the mass.

Appendix A.2. No queue exists at t

Assume $Q(t) = 0$ and set $t' = \check{t}$. If $M \leq s \cdot (t' - t)$, vehicles provisionally scheduled in the mass can instead be rescheduled to depart at rate s for a period of duration M/s . The M vehicles arrive at the same time as with the mass departure, but without incurring a queuing delay. If $M > s \cdot (t' - t)$, then $s \cdot (t' - t)$ of the vehicles can be rescheduled to depart at rate s for a period $t' - t$, and the remaining $M - s \cdot (t' - t)$ vehicles can be rescheduled to depart in a mass at t' . All M vehicles will arrive at the same time as they do in the mass departure at t . The first $s \cdot (t' - t)$ vehicles do not queue at all, and the last $M - s \cdot (t' - t)$ vehicles do not queue during $[t, t']$. Total queuing costs are therefore reduced, and total schedule delay costs are unchanged. Postponing the mass departure again reduces fleet costs, so that the mass departure is not optimal. *QED*

Appendix B. Proof of Lemma 4

It suffices to show that user A 's vehicles that depart after t_q incur lower total costs in the revised schedule than in the candidate PSNE schedule.

Appendix B.1. Costs in the candidate equilibrium

Total queuing time in the candidate PSNE is measured by area ABC in Figure 1. User A incurs half of this delay. User A 's queuing time costs are therefore $\alpha \cdot (t^* - \tilde{t}) \cdot s \cdot (t_e - t_q) / 4$. The number of user A 's vehicles that queue and arrive early is $s \cdot (t^* - t_q) / 2$, and their average schedule delay cost is $\beta \cdot (t^* - t_q) / 2$. The number of vehicles that queue and arrive late is $s \cdot (t_e - t^*) / 2$, and their average schedule delay cost is $\gamma \cdot (t_e - t^*) / 2$. User A 's queuing vehicles therefore incur total costs of

$$TC^e = \frac{\alpha \cdot s}{4} \cdot (t^* - \tilde{t}) \cdot (t_e - t_q) + \frac{\beta \cdot s}{4} \cdot (t^* - t_q)^2 + \frac{\gamma \cdot s}{4} \cdot (t_e - t^*)^2, \quad (\text{B.1})$$

where superscript e denotes the candidate PSNE. Substituting Eqs. (28), (29), and (30) into Eq. (B.1) gives

$$TC^e = \frac{\beta^2 \cdot (\alpha^2 + \alpha \cdot (\gamma - \beta) + 3 \cdot \beta \cdot \gamma)}{8 \cdot s \cdot (\alpha + \beta)^2 \cdot (\beta + \gamma)} \cdot N^2. \quad (\text{B.2})$$

Appendix B.2. Costs in the revised schedule

Time \hat{t} in Figure 1 is defined by the condition

$$\frac{\alpha}{\alpha - \beta} \cdot s \cdot (\tilde{t} - t_q) + \frac{\alpha}{\alpha + \gamma} \cdot s \cdot (\hat{t} - \tilde{t}) = s \cdot (\hat{t} - t_q).$$

This solves to yield

$$\hat{t} = t^* - \frac{\beta \cdot (\gamma - \alpha) \cdot (\gamma - \beta)}{2 \cdot \gamma \cdot (\alpha + \beta) \cdot (\beta + \gamma)} \cdot \frac{N}{s} < t^*, \quad (\text{B.3})$$

where the inequality follows from $\gamma > \alpha > \beta$.

The number of user A 's vehicles that depart after \hat{t} and arrive early in the revised schedule is $\gamma \cdot s \cdot (t^* - \hat{t}) / (\alpha + \gamma)$. These vehicles incur an average schedule delay cost of $\beta \cdot (t^* - \hat{t}) / 2$. The number of user A 's vehicles that depart after \hat{t} and arrive late is $\gamma \cdot s \cdot (t_e - t^*) / (\alpha + \gamma)$ and their average schedule delay cost is $\gamma \cdot (t_e - t^*) / 2$. Total costs are therefore

$$TC^d = \frac{\beta \cdot \gamma \cdot s}{2 \cdot (\alpha + \gamma)} \cdot (t^* - \hat{t})^2 + \frac{\gamma^2 \cdot s}{2 \cdot (\alpha + \gamma)} \cdot (t_e - t^*)^2, \quad (\text{B.4})$$

where superscript d denotes the deviant schedule. Substituting Eqs. (B.3) and (28) into Eq. (B.4) gives

$$TC^d = \frac{\alpha^2 \cdot (\beta^2 - 3 \cdot \beta \cdot \gamma + 4 \cdot \gamma^2) + 2 \cdot \alpha \cdot \beta \cdot \gamma \cdot (3 \cdot \gamma - \beta) + \beta \cdot \gamma^2 \cdot (\beta + \gamma)}{8 \cdot \gamma \cdot (\alpha + \beta)^2 \cdot (\alpha + \gamma) \cdot (\beta + \gamma) \cdot s} \cdot \beta^2 \cdot N^2. \quad (\text{B.5})$$

Given Eqs. (B.2) and (B.5), the costs saved by deviating from the candidate PSNE work out to

$$TC^e - TC^d = \frac{\beta^2 \cdot (\alpha - \gamma)^2 \cdot (\alpha \cdot \gamma - \beta^2 + 2 \cdot \beta \cdot \gamma)}{8 \cdot s \cdot \gamma \cdot (\alpha + \beta)^2 \cdot (\alpha + \gamma) \cdot (\beta + \gamma)} \cdot N^2 > 0.$$

Since the cost saving is positive, the candidate PSNE is not a PSNE. *QED*

Appendix C. Proposition 2: Proof of Result 4 for mass departures

Suppose that user A deviates from the candidate PSNE by scheduling multiple mass departures. We prove that such a deviation cannot reduce user A 's costs. The proof is done in two steps. We first show that any deviation with multiple mass departures does not achieve lower costs than a deviation with a single mass departure launched before t^* (Appendix C.1). We then show that this deviation is not gainful if the bounds on departure rates stated in Proposition 2 are satisfied (Appendix C.2).

Appendix C.1. Optimality of the single mass departure deviation

We show that any deviation from the PSNE involving multiple mass departures is dominated by a single mass departure. The proof involves establishing the following three results for either user: (i) Fleet costs can be (weakly) reduced by rescheduling any vehicles that suffer queuing delay, but are not part of a mass, to a period without queuing. (ii) Fleet costs can be (weakly) reduced by rescheduling any vehicles in a mass departure after t^* to a period without queuing. (iii) Any deviation with multiple mass departures launched before t^* entails strictly higher fleet costs than a deviation with a single mass departure. These three results establish that the candidate PSNE need only be tested against a single mass departure launched before t^* .

Result i: When a queue exists, user A is willing to depart at a positive and finite rate only if condition (20) is satisfied. For early arrivals this requires $r_B(t) = \alpha \cdot s / (\alpha - \beta) > s$. Since $r_B(t) < s$ in the candidate PSNE, condition (20) is violated, and user A will not deviate by departing early when there is a queue. For late arrivals, user A is willing to depart at a positive and finite rate when there is a queue only if $r_B(t) = \alpha \cdot s / (\alpha + \gamma)$. Proposition 2 stipulates that $r_B(t) \leq \alpha \cdot s / (\alpha + \gamma)$. If $r_B(t) < \alpha \cdot s / (\alpha + \gamma)$, user A is better off scheduling its vehicles later. If $r_B(t) = \alpha \cdot s / (\alpha + \gamma)$, user A is indifferent between departing or not, and would not gain by rescheduling vehicles late when there is a queue.

Result ii: Assume that the last mass departure is launched at a time of late arrivals. We show that rescheduling vehicles in the mass to a later period in which they do not incur queuing delay is beneficial. By induction, it then follows that all mass departures launched at times of late arrivals can be gainfully rescheduled.

Suppose the last mass departure is launched at time t_L and comprises M vehicles. Assume first that there is a queue at t_L . We show that postponing the mass departure to the moment when the queue disappears (weakly) reduces fleet costs. Let $m \in [0, M]$ index vehicles in the order they are positioned in the mass, and let $c(m, t_L)$ denote the cost incurred by vehicle m :

$$c(m, t_L) = \alpha \cdot \left(T(t_L) + \frac{m}{s} \right) + \gamma \cdot \left(t_L + T(t_L) + \frac{m}{s} - t^* \right).$$

Now suppose the mass is postponed to $t'_L > t_L$ when a queue still exists. Vehicle m now incurs a cost of:

$$c(m, t'_L) = \alpha \cdot \left(T(t'_L) + \frac{m}{s} \right) + \gamma \cdot \left(t'_L + T(t'_L) + \frac{m}{s} - t^* \right). \quad (\text{C.1})$$

Since only user B departs during (t_L, t'_L) ,

$$T(t'_L) = T(t_L) + \int_{t_L}^{t'_L} \frac{r_B(u) - s}{s} du = T(t_L) - (t'_L - t_L) \cdot \left(1 - \lambda_{t_L, t'_L}^B \right). \quad (\text{C.2})$$

Substituting (C.2) in (C.1), we get:

$$c(m, t'_L) = c(m, t_L) - \alpha \cdot (t'_L - t_L) \cdot \left(1 - \lambda_{t_L, t'_L}^B \right) + \gamma \cdot (t'_L - t_L) \cdot \lambda_{t_L, t'_L}^B.$$

Postponing the mass departure changes costs by:

$$c(m, t'_L) - c(m, t_L) = (t'_L - t_L) \cdot \left[\lambda_{t_L, t'_L}^B \cdot (\alpha + \gamma) - \alpha \right] \leq 0,$$

where the inequality follows from the condition $r_B(t) \leq \alpha \cdot s / (\alpha + \gamma)$ in Proposition 2.

We conclude that if there is a queue when the last mass departs, the mass departure can be postponed to the time where the queue just disappears without increasing costs (later vehicles are not affected by postponing the mass).

Now assume there is no queue at t_L when the last mass is launched. We show that vehicles in the mass can be rescheduled later to a time when they do not face a queue, and that doing so does not increase user A 's fleet costs. Consider the cost of the last (i.e., M -th) vehicle in the mass:

$$c(M, t_L) = \gamma \cdot (t_L - t^*) + (\alpha + \gamma) \cdot \frac{M}{s} .$$

Because user A does not depart until the queue has disappeared (result i), the queue produced by the mass departure disappears at time t'_L where:

$$\frac{M}{s} + \int_{t_L}^{t'_L} \frac{r_B(u) - s}{s} du = \frac{M}{s} - (t'_L - t_L) \cdot (1 - \lambda_{t_L, t'_L}^B) = 0 . \quad (\text{C.3})$$

A vehicle that departs at t'_L incurs a cost:

$$c(t'_L) = \gamma \cdot (t'_L - t^*) = \gamma \cdot (t'_L - t_L) + \gamma \cdot (t_L - t^*) = \gamma \cdot (t_L - t^*) + \gamma \cdot \frac{M}{s \cdot (1 - \lambda_{t_L, t'_L}^B)} ,$$

where the last equality uses Eq. (C.3). As removing a vehicle from the mass opens up a (queue-free) slot at t'_L , rescheduling the last vehicle to t'_L yields a change in cost of:

$$c(t'_L) - c(M, t_L) = \frac{M}{s} \cdot \left[\frac{\gamma}{1 - \lambda_{t_L, t'_L}^B} - (\alpha + \gamma) \right] \leq 0 ,$$

where the inequality follows from condition $r_B(t) \leq \alpha \cdot s / (\alpha + \gamma)$ in Proposition 2.

This shows that, in any deviation from the candidate PSNE, the last mass departure launched at a time of late arrivals can be eliminated without increasing fleet costs. By induction, any mass departure launched at a time of late arrivals can be rescheduled without increasing fleet costs.

Next, we show that any deviation entailing multiple mass departures at times of early arrivals is dominated by scheduling a single mass departure before t^* .

Result iii: Suppose that more than one mass departure is scheduled before t^* . Assume the first mass is launched at time t_E with M vehicles, and the second mass is launched at time $t'_E > t_E$. If the queue from the first mass disappears before t'_E , fleet costs can be reduced by rescheduling vehicles in the first mass to depart at a rate equal to residual capacity $(s - r_B(t))$. Since user A does not depart in the original deviation until the queue has dissipated, the rescheduled vehicles in the alternative deviation escape queuing and arrive less early – thereby reducing both their queuing and schedule delay costs. If the queue from the first mass does not disappear before t'_E , user A can still reduce its fleet costs by rescheduling $s \cdot (t'_E - t_E) \cdot (1 - \lambda_{t_E, t'_E}^B)$ vehicles at a rate $s - r_B(t)$ during (t_E, t'_E) , and letting the remaining $M - s \cdot (t'_E - t_E) \cdot (1 - \lambda_{t_E, t'_E}^B)$ vehicles join the head of the second mass at t'_E . The first set of vehicles rescheduled at residual capacity avoid queuing and

incur lower early arrival costs because they arrive closer to t^* . The second set of vehicles also incur lower schedule delay costs since they arrive later. They also incur lower queuing costs as well since they no longer queue between t_E and t'_E . Vehicles in the original mass that departs at t'_E still depart and arrive at the same time because the same number of vehicles depart before them and the bottleneck operates at capacity throughout.

By induction, all but one of the mass departures launched before t^* can be eliminated in a way that decreases fleet costs. Thus, results (i)–(iii) show that a deviation with a single mass departure launched before t^* is the most viable deviation, of deviations entailing mass departures, from the candidate PSNE. In the following section we show that the cost-minimizing deviation does not reduce fleet costs with respect to the candidate PSNE.

Appendix C.2. The optimal mass departure

Following the logic of Appendix C.1, suppose user A launches a single mass departure of M vehicles at $t_m < t^*$. User A will not include in the mass vehicles that were scheduled to depart before t_m because this would create a queue that lasts until after t_e , and the cost of the delay imposed on A 's later vehicles would outweigh any benefit. User A will include in the mass all vehicles scheduled to depart after t_m that would be delayed by the mass. However, as discussed above, this would be counterproductive if all these vehicles were scheduled to depart before t^* because they would suffer not only queuing delay but also greater early-arrival costs (the queue would disappear before t^*).

The mass departure is potentially beneficial only if it includes vehicles scheduled to depart both early and late. Therefore, the mass must be launched at a time $t_m < t^*$, and the queue must persist until a time $t_M > t^*$, where t_M is defined by the condition:

$$\int_{t_m}^{t_M} r_A(u) du = M, \quad (\text{C.4})$$

where $r_A(\cdot)$ is user A 's departure schedule in the candidate PSNE. Since the queuing costs incurred by the mass do not depend on when it is launched, t_m should be chosen to minimize total schedule delay costs for vehicles in the mass. The first and last vehicles should therefore incur the same schedule delay cost:

$$\beta \cdot (t^* - t_m) = \gamma \cdot \left(t_m + \frac{M}{s} - t^* \right).$$

This implies

$$t_m = t^* - \frac{\gamma}{\beta + \gamma} \cdot \frac{M}{s}.$$

Vehicles in the mass incur total queuing time costs of $\alpha \cdot M^2 / (2 \cdot s)$, and total schedule delay costs of $\delta \cdot M^2 / (2 \cdot s)$. Total costs for the mass are therefore:

$$TC^m = \frac{\alpha + \delta}{2} \cdot \frac{M^2}{s}. \quad (\text{C.5})$$

In the candidate PSNE, where $r_A(t) + r_B(t) = s$ for $t \in (t_m, t_M)$, the M vehicles incur total costs of

$$TC^e = \beta \cdot \int_{t_m}^{t^*} r_A(u) \cdot (t^* - u) du + \gamma \cdot \int_{t^*}^{t_M} r_A(u) \cdot (u - t^*) du.$$

To complete the proof we must show that $TC^e \leq TC^m$. User A 's departure rate over the period (t_m, t_M) must be consistent with condition (C.4). As shown in proving Result 3 of Proposition 2, during late arrivals $r_A(u)$ is bounded below by $\gamma \cdot s / (\alpha + \gamma)$. Let r_E denote the minimum departure rate of user A prior to t^* . It follows by straightforward algebra that

$$TC^e \leq M^2 \cdot \frac{\left(\beta \cdot \gamma^2 \cdot r_E \cdot s + (\alpha + \gamma) \cdot \left((\beta + \gamma)^2 \cdot s^2 + r_E^2 \cdot \gamma^2 - 2 \cdot (\beta + \gamma) \cdot \gamma \cdot s \cdot r_E \right) \right)}{2 \cdot (\beta + \gamma)^2 \cdot s^3}. \quad (\text{C.6})$$

Setting the right-hand side of inequality (C.6) equal to Eq. (C.5) one obtains

$$r_E = \frac{s}{2} \cdot \frac{2 \cdot \alpha \cdot (\beta + \gamma) + \gamma \cdot (\beta + 2 \cdot \gamma) - \sqrt{\beta^2 \cdot \gamma^2 + 4 \cdot \alpha \cdot (\beta + \gamma)^2 \cdot (\alpha + \gamma)}}{\gamma \cdot (\alpha + \gamma)}.$$

If user A 's departure rate during early arrivals is at least r_E , departing in a mass cannot reduce its fleet costs. The candidate PSNE is therefore robust to mass departure deviations.

Straightforward algebra leads to:

$$0 < r_E < \frac{\gamma}{\alpha + \gamma} \cdot s,$$

which shows that in the limit $\gamma \rightarrow 0$, $r_E \rightarrow 0$. *QED*

Appendix D. Asymmetric equilibrium costs with $\gamma \leq \alpha$

User A 's fleet costs in the PSNE depicted in Figure 2 are

$$\begin{aligned} TC_A &= r_E \cdot (t_{BA} - t_s^o) \cdot \beta \cdot \left(t^* - \frac{t_s^o + t_{BA}}{2} \right) + (s - r_E) \cdot (t^* - t_{BA}) \cdot \beta \cdot \frac{t^* - t_{BA}}{2} + \\ &\quad \frac{\alpha}{\alpha + \gamma} \cdot s \cdot (t_{AB} - t^*) \cdot \gamma \cdot \frac{t_{AB} - t^*}{2} + \frac{\gamma}{\alpha + \gamma} \cdot s \cdot (t_e^o - t_{AB}) \cdot \gamma \cdot \left(\frac{t_{AB} + t_e^o}{2} - t^* \right) \\ &= \frac{\beta \cdot r_E}{2} \cdot \left((t^* - t_s^o)^2 - (t^* - t_{BA})^2 \right) + \frac{\beta}{2} \cdot (s - r_E) \cdot (t^* - t_{BA})^2 \\ &\quad + \frac{\alpha \cdot \gamma \cdot s}{2 \cdot (\alpha + \gamma)} \cdot (t_{AB} - t^*)^2 + \frac{\gamma^2 \cdot s}{2 \cdot (\alpha + \gamma)} \cdot \left((t_e^o - t^*)^2 - (t_{AB} - t^*)^2 \right). \end{aligned} \quad (\text{D.1})$$

Differentiating TC_A with respect to t_{BA} and t_{AB} , and using Eqs. (9), (10) and condition (31) in the text, one obtains for the cost-minimizing transition times:

$$t_{BA} = t^* - \frac{\gamma}{2 \cdot (\beta + \gamma)} \cdot \frac{N}{s}, \quad t_{AB} = t^* + \frac{\beta}{2 \cdot (\beta + \gamma)} \cdot \frac{N}{s}. \quad (\text{D.2})$$

Substituting Eqs. (9), (10), and (D.2) into Eq. (D.1) gives

$$TC_A = \left(\frac{3+z}{4} - \frac{1}{2} \cdot \sqrt{\frac{\alpha}{\alpha+\gamma} + z^2} \right) \cdot TC^o,$$

where $z \equiv \beta \cdot \gamma / ((\alpha + \gamma) \cdot (\beta + \gamma))$.

Appendix E. Heterogeneous desired arrival times

Appendix E.1. Case 1

User A 's fleet costs in Case 1 are:

$$TC_A^1 = \frac{\beta \cdot s}{2} \cdot (t_A^* - t_{As}^1)^2 + \frac{\gamma \cdot s}{2} \cdot \left(t_{As}^1 + \frac{N_A}{s} - t_A^* \right)^2 = \frac{\beta \cdot \gamma}{2(\beta + \gamma)} \cdot \frac{N_A^2}{s}.$$

User B 's fleet costs are:

$$\begin{aligned} TC_B^1 &= \frac{\beta \cdot s}{2} \cdot (t_B^* - t_{Be}^1)^2 + \frac{\gamma \cdot s}{2} \cdot (t_{Be}^1 - t_B^*)^2 \\ &= \frac{\beta \cdot s}{2} \cdot \left(t_B^* - t_A^* - \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s} \right)^2 + \frac{\gamma \cdot s}{2} \cdot \left(t_A^* + \frac{\beta}{\beta + \gamma} \cdot \frac{N_A}{s} + \frac{N_B}{s} - t_B^* \right)^2 \\ &= \frac{\beta \cdot s}{2} \cdot \left(\frac{\gamma}{\beta + \gamma} \cdot \frac{N_B}{s} - x \right)^2 + \frac{\gamma \cdot s}{2} \cdot \left(x + \frac{\beta}{\beta + \gamma} \cdot \frac{N_B}{s} \right)^2 \\ &= \frac{\beta \cdot \gamma}{2 \cdot (\beta + \gamma)} \cdot \frac{N_B^2}{s} + \frac{\beta + \gamma}{2} \cdot s \cdot x^2, \end{aligned}$$

where $x \equiv t_A^* - t_B^* + (\beta/(\beta + \gamma)) \cdot (N_A/s) + (\gamma/(\beta + \gamma)) \cdot (N_B/s)$. Total system costs are:

$$TC^1 = TC_A^1 + TC_B^1 = \frac{\beta \cdot \gamma}{2 \cdot (\beta + \gamma) \cdot s} \cdot (N_A^2 + N_B^2) + \frac{\beta + \gamma}{2} \cdot s \cdot x^2.$$

The price of anarchy is

$$PA = \frac{TC^1 - TC^o}{TC^o} = \frac{\frac{\beta + \gamma}{4} \cdot s \cdot x^2}{\frac{\beta \cdot \gamma}{2 \cdot (\beta + \gamma) \cdot s} \cdot (N_A^2 + N_B^2) + \frac{\beta + \gamma}{4} \cdot s \cdot x^2}. \quad (\text{E.1})$$

PA is an increasing function of x . Condition (49) in the text, $t_B^* - t_A^* \geq (\gamma/(\beta + \gamma)) \cdot (N_B/s)$, implies that $x \leq (\beta/(\beta + \gamma)) \cdot (N_A/s)$. Substituting this value for x into Eq. (E.1), and simplifying, yields

$$PA \leq \frac{\beta \cdot N_A^2}{2 \cdot \gamma \cdot (N_A^2 + N_B^2) + \beta \cdot N_A^2}. \quad (\text{E.2})$$

Condition (48) in the text, $t_B^* - t_A^* > (\beta/(\beta + \gamma)) \cdot (N_A/s)$, implies that $x \leq (\gamma/(\beta + \gamma)) \cdot (N_B/s)$. This in turn implies $(\beta/(\beta + \gamma)) \cdot (N_A/s) \leq (\gamma/(\beta + \gamma)) \cdot (N_B/s)$, or $N_A \leq (\gamma/\beta) \cdot N_B$. Substituting this inequality into Eq. (E.2) yields

$$PA \leq \frac{\frac{\beta}{\gamma}}{2 + \frac{\beta}{\gamma} + 2 \cdot \frac{\beta^2}{\gamma^2}}. \quad (\text{E.3})$$

In the model variant in Section 4.1 the relative magnitudes of parameters β and γ are unrestricted so there is no upper or lower bound on β/γ . The formula in Eq. (E.3) reaches a maximum at $\beta/\gamma = 1$, for which $PA = 1/5$.

The lower bound on the efficiency index w is derived in the same way as the upper bound on PA . Imposing conditions (48) and (49) as equalities yields $x = (\beta/(\beta + \gamma)) \cdot (N_A/s) = (\gamma/(\beta + \gamma)) \cdot (N_B/s)$. Substituting these equalities into Eq. (51) gives

$$w \geq \frac{\beta \cdot \gamma + \beta^2 + \gamma^2}{\frac{3}{2}\beta \cdot \gamma + \beta^2 + \gamma^2}. \quad (\text{E.4})$$

Eq. (E.4) achieves a minimum value of $6/7$ with $\beta = \gamma$.

Appendix E.2. Case 2

The three transition times in Case 2, t_{As}^2 , t_{Ae}^2 , and t_{Be}^2 , are solved using three conditions. First, Eq. (43) for Case 1 continues to apply, as otherwise user A could decrease its fleet costs by rescheduling a vehicle from t_{As}^2 to t_{Ae}^2 , or vice versa (cf Eq. (E.5a) below). Second, the bottleneck is fully utilized from t_{As}^2 to t_{Be}^2 (cf Eq. (E.5b)). Finally, user B 's entire fleet must depart during the period $[t_A^*, t_{Be}^2]$ (cf Eq. (E.5c)).

$$\beta \cdot (t_A^* - t_{As}^2) = \gamma \cdot (t_{Ae}^2 - t_A^*) \quad (\text{E.5a})$$

$$t_{Be}^2 - t_{As}^2 = \frac{N_A + N_B}{s} \quad (\text{E.5b})$$

$$\frac{\alpha}{\alpha + \gamma} \cdot s (t_{Ae}^2 - t_A^*) + s \cdot (t_{Be}^2 - t_{Ae}^2) = N_B. \quad (\text{E.5c})$$

Equations (E.5) resolve to Eqs. (52)-(54) in the text. Three consistency conditions must be satisfied. First, as in Case 1, user B cannot gain by rescheduling its last vehicle to t_{As}^2 . The requisite condition is:

$$t_B^* - t_A^* \geq \frac{\gamma}{\beta + \gamma} \cdot \frac{N_B}{s} - \frac{\alpha \cdot \beta}{(\beta + \gamma) \cdot (\alpha + \beta + \gamma)} \cdot \frac{N_A}{s}. \quad (\text{E.6})$$

Second, user B cannot gain by rescheduling a vehicle at t_A^* to t_{Be}^2 . Thus, $\beta \cdot (t_B^* - t_A^*) \leq \gamma \cdot (t_{Be}^2 - t_B^*)$ which reduces to:

$$t_B^* - t_A^* \leq \frac{\beta \cdot \gamma}{(\beta + \gamma) \cdot (\alpha + \beta + \gamma)} \cdot \frac{N_A}{s} + \frac{\gamma}{\beta + \gamma} \cdot \frac{N_B}{s}. \quad (\text{E.7})$$

Condition (E.7) is more stringent than condition (34). Third, user A must stop departing before user B starts to arrive late (i.e., $t_{Ae}^2 < t_B^*$) since otherwise the PSNE does not exist unless $\gamma \leq \alpha$.¹⁴ Given Eq. (53) this implies

$$t_B^* - t_A^* \geq \frac{\beta \cdot (\alpha + \gamma)}{\gamma \cdot (\alpha + \beta + \gamma)} \cdot \frac{N_A}{s}. \quad (\text{E.8})$$

¹⁴This follows from the proof of Proposition 1. In particular, Lemma 1 holds and the only possible PSNE entails queuing. Using a similar reasoning as in Lemmas 3 and 4 it is possible to show that a PSNE will not exist.

Using Eqs. (38) and (52), the difference in timing of the social optimum and the PSNE in Case 2 can be written as

$$t_{As}^2 - t_{As}^o = \frac{-\beta \cdot \alpha}{(\beta + \gamma) \cdot (\alpha + \beta + \gamma)} \cdot \frac{N_A}{s} + \frac{x}{2}. \quad (\text{E.9})$$

Condition (E.7) implies a minimum value of $x = (\beta \cdot (\alpha + \beta) / ((\beta + \gamma) \cdot (\alpha + \beta + \gamma))) \cdot (N_A/s)$. Condition (E.6) implies a maximum value of $x = (\beta \cdot (2\alpha + \beta + \gamma) / ((\beta + \gamma) \cdot (\alpha + \beta + \gamma))) \cdot (N_A/s)$. Finally, Condition (E.8) implies a maximum value of $x = (\gamma / (\beta + \gamma)) \cdot (N_B/s) - (\alpha \cdot \beta^2 / (\gamma \cdot (\beta + \gamma) \cdot (\alpha + \beta + \gamma))) \cdot (N_A/s)$. Applying these values to Eq. (E.9) yields the feasible range for $t_{As}^2 - t_{As}^o$:

$$t_{As}^2 - t_{As}^o \in \left[\text{Min} \left(\frac{\beta}{2 \cdot (\alpha + \beta + \gamma)} \cdot \frac{N_A}{s}, \frac{\gamma}{2 \cdot (\beta + \gamma)} \cdot \frac{N_B}{s} - \frac{\alpha \cdot \beta \cdot (\beta + 2\gamma)}{2 \cdot \gamma \cdot (\beta + \gamma) \cdot (\alpha + \beta + \gamma)} \cdot \frac{N_A}{s} \right) \right].$$

The lower bound applies with $t_B^* - t_A^*$ at its maximum value consistent with condition (E.7). The upper bound applies with $t_B^* - t_A^*$ at its minimum value consistent with conditions (E.6) and (E.8).

References

- Anderson, S. P., de Palma, A. and Thisse, J. F. (1992), *Discrete choice theory of product differentiation*, MIT press, Cambridge, Mass.
- Arnott, R., de Palma, A. and Lindsey, R. (1987), Bottleneck congestion with elastic demand, Discussion Paper 690, Institute for Economic Research, Queen's University.
- Arnott, R., de Palma, A. and Lindsey, R. (1990), 'Economics of a bottleneck', *Journal of Urban Economics* **27**(1), 111–130.
- Arnott, R., de Palma, A. and Lindsey, R. (1993), 'A structural model of peak-period congestion: A traffic bottleneck with elastic demand', *The American Economic Review* **83**(1), 161–179.
- Arnott, R., de Palma, A. and Lindsey, R. (1998), Recent developments in the bottleneck model, in K. Button and E. Verhoef, eds, 'Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility', Edward Elgar, Aldershot, pp. 79–110.
- Asensio, J. and Matas, A. (2008), 'Commuters valuation of travel time variability', *Transportation Research Part E: Logistics and Transportation Review* **44**(6), 1074–1085.
- Basso, L. J. and Zhang, A. (2007), 'Congestible facility rivalry in vertical structures', *Journal of Urban Economics* **61**(2), 218–237.

- Ben-Elia, E. and Ettema, D. (2009), ‘Carrots versus sticks: Rewarding commuters for avoiding the rush-hour—a study of willingness to participate’, *Transport Policy* **16**(2), 68–76.
- Ben-Elia, E. and Ettema, D. (2011), ‘Rewarding rush-hour avoidance: A study of commuters travel behavior’, *Transportation Research Part A: Policy and Practice* **45**(7), 567–582.
- Brueckner, J. K. (2002), ‘Airport congestion when carriers have market power’, *The American Economic Review* **92**(5), 1357–1375.
- Brueckner, J. K. (2005), ‘Internalization of airport congestion: A network analysis’, *International Journal of Industrial Organization* **23**(7–8), 599–614.
- Brueckner, J. K. and Van Dender, K. (2008), ‘Atomistic congestion tolls at concentrated airports? seeking a unified view in the internalization debate’, *Journal of Urban Economics* **64**(2), 288–295.
- Catoni, S. and Pallottino, S. (1991), ‘Technical note—traffic equilibrium paradoxes’, *Transportation Science* **25**(3), 240–244.
- Cominetti, R., Correa, J. R. and Stier-Moses, N. E. (2009), ‘The impact of oligopolistic competition in networks’, *Operations Research* **57**(6), 1421–1437.
- Daniel, J. I. (1995), ‘Congestion pricing and capacity of large hub airports: A bottleneck model with stochastic queues’, *Econometrica: Journal of the Econometric Society* **63**(2), 327–370.
- Daniel, J. I. (2009), The deterministic bottleneck model with non-atomistic traffic, Working Papers 09–08, University of Delaware, Department of Economics.
- Daniel, J. I. and Harback, K. T. (2008), ‘(When) do hub airlines internalize their self-imposed congestion delays?’, *Journal of Urban Economics* **63**(2), 583–612.
- de Palma, A., Ben-Akiva, M., Lefevre, C. and Litinas, N. (1983), ‘Stochastic equilibrium model of peak period traffic congestion’, *Transportation Science* **17**(4), 430–453.
- de Palma, A. and Engelson, L. (2012), Network market conduct with atomic and non-atomic players, Technical report, May 14, paper presented at Kuhmo Nectar – ITEA Conference on Transportation Economics, Berlin, June 21-22.
- de Palma, A. and Fontan, C. (2001), ‘Éléments d’analyse de la composante horaire des déplacements: le cas de la région ile-de-france’, *Les Cahiers Scientifiques du Transport* **39**, 55–86.

- de Palma, A. and Fosgerau, M. (2011), Dynamic traffic modeling, *in* A. de Palma, R. Lindsey, E. Quinet and R. Vickerman, eds, ‘Handbook in Transport Economics’, Edward Elgar, Cheltenham, UK and Northampton, Mass, USA, pp. 188–212.
- Devarajan, S. (1981), ‘A note of network equilibrium and noncooperative games’, *Transportation Research Part B: Methodological* **15**(6), 421–426.
- Fotakis, D., Kontogiannis, S. and Spirakis, P. (2008), ‘Atomic congestion games among coalitions’, *ACM Transactions on Algorithms (TALG)* **4**(4), 52.
- Gutiérrez-i Puigarnau, E. and Van Ommeren, J. N. (2012), ‘Start time and worker compensation implications for staggered-hours programmes’, *Journal of Transport Economics and Policy (JTEP)* **46**(2), 205–220.
- Harker, P. T. (1988), ‘Multiple equilibrium behaviors on networks’, *Transportation Science* **22**(1), 39–46.
- Hofer, M. and Skopalik, A. (2009), Altruism in atomic congestion games, *in* A. Fiat and P. Sanders, eds, ‘Lecture Notes in Computer Science’, Vol. 5757, Springer, Berlin, pp. 179–189.
- Levinson, D. (2005), ‘Micro-foundations of congestion and pricing: A game theory perspective’, *Transportation Research Part A: Policy and Practice* **39**(7), 691–704.
- Lindsey, R. (2004), ‘Existence, uniqueness, and trip cost function properties of user equilibrium in the bottleneck model with multiple user classes’, *Transportation Science* **38**(3), 293–314.
- Lindsey, R., van den Berg, V. A. C. and Verhoef, E. T. (2012), ‘Step tolling with bottleneck queuing congestion’, *Journal of Urban Economics* **72**(1), 46–59.
- Marcotte, P. (1987), ‘Algorithms for the network oligopoly problem’, *Journal of the Operational Research Society* **38**(11), 1051–1065.
- Miller, T. C., Tobin, R. L. and Friesz, T. L. (1991), ‘Stackelberg games on a network with cournot-nash oligopolistic competitors’, *Journal of Regional Science* **31**(4), 435–454.
- Mun, S.-i. and Yonekawa, M. (2006), ‘Flexitime, traffic congestion and urban productivity’, *Journal of Transport Economics and Policy (JTEP)* **40**(3), 329–358.
- Newell, G. F. (1987), ‘The morning commute for nonidentical travelers’, *Transportation Science* **21**(2), 74–88.
- Otsubo, H. and Rapoport, A. (2008), ‘Vickrey’s model of traffic congestion discretized’, *Transportation Research Part B: Methodological* **42**(10), 873–889.

- Pels, E. and Verhoef, E. T. (2004), ‘The economics of airport congestion pricing’, *Journal of Urban Economics* **55**(2), 257–277.
- Silva, H. E. and Verhoef, E. T. (2013), ‘Optimal pricing of flights and passengers at congested airports and the efficiency of atomistic charges’, *Journal of Public Economics* **106**, 1–13.
- Silva, H. E., Verhoef, E. T. and van den Berg, V. A. C. (2014), ‘Airlines’ strategic interactions and airport pricing in a dynamic bottleneck model of congestion’, *Journal of Urban Economics* **80**, 13–27.
- Small, K. A. (1982), ‘The scheduling of consumer activities: work trips’, *The American Economic Review* **72**(3), 467–479.
- The Economist (2013), Air cargo: Cabin fever, Technical report, November 2, p.72 (<http://www.economist.com/news/business/21588923-fedex-and-ups-have-turned-memphis-and-louisville-aerotropolises-cabin-fever>).
- Vickrey, W. S. (1969), ‘Congestion theory and transport investment’, *The American Economic Review* **59**(2), 251–260.
- Wardman, M. (2001), ‘A review of british evidence on time and service quality valuations’, *Transportation Research Part E: Logistics and Transportation Review* **37**(2), 107–128.
- Zhang, A. and Zhang, Y. (2006), ‘Airport capacity and congestion when carriers have market power’, *Journal of Urban Economics* **60**(2), 229–247.
- Zou, X. and Levinson, D. (2006), ‘A multi-agent congestion and pricing model’, *Transportmetrica* **2**(3), 237–249.