Discomfort in mass transit and its implication for scheduling and pricing

André DE PALMA, Moez KILANI and Stef PROOST

Energy, Transport & Environment

Faculty of Economics
And Business

Discomfort in mass transit and its implication for scheduling and pricing

André de Palma
ENS Cachan - France
Email: andre.depalma@ens-cachan.fr

Moez Kilani
Université Charles de Gaulle - Lille 3
Email: moez.kilani@univ-lille3.fr

Stef Proost
KULeuven
Email: Stef.Proost@econ.kuleuven.ac.be

February 1, 2013

Abstract

This paper proposes an analytical formulation of discomfort in mass transit and discusses its micro-economic properties. The formula we introduce reflects real situations faced by the passengers, it has nice mathematical properties and it is easy to compute. The discomfort formulation is used to analyze optimal scheduling and pricing of transit in a dynamic model.

1   Introduction

The quality of mass transit and in particular crowding has become a severe problem in several metropolitan areas in developing countries, but also in European countries. In Paris and London it is hard to get into the metro or local commuter trains in urban areas at peak times. As a consequence, passengers

may change their departure time, their route or mode in order to avoid excessive congestion. These changes are costly, the transit authority has difficulties to associate a money value to those costs.

Many metropolitan areas have adopted a second best low price policy to attract car users into public transportation. This policy can be justified (cf. Parry and Small, 2009; Proost and Van Dender, 2008), but the necessary capacity decisions calling upon public revenues have not always been adopted and this resulted in severe congestion in mass transit. In this paper, we focus on the characterization of congestion in mass transit.

Note that there are very few studies of congestion in mass transit while there are many studies of congestion in private transportation. This may be understood in the US where the use of public transport is typically low except in some cities like New York and Boston. But this is more difficult to understand for European cities where the fraction of commuters can exceed 50% in the peak period.

There is a long tradition in dealing with road congestion. The so-called BPR formula - Bureau of Public Roads - is now the standard in the literature. But the micro-economic theory of mass transit has not dealt frequently with crowding and riding comfort. There are only a few exceptions in the literature. Among those let us mention the pioneering contribution of Kraus (1991), where a distinction is made between the value of time for standing and for seated passengers. Kraus and Yoshida (2002) focus on the congestion on the rail platform where passengers may have to wait for several trains before they can enter. Several papers have integrated congestion in public transport in a multi-modal model but treat this important issue in a simplified way. Huang (2000) has comfort costs linearly increasing in the number of users and Rouwendal and Verhoef (2004) have crowding as an increasing function of occupancy ratio. Many empirical studies confirmed the importance of discomfort conditions in rail (see Wardman and Whelan (2011)). Vuchic (2005), provides a major contribution to the recent theory of transit but does not explicitly deal with crowding in urban transit. Jara-Díaz and Gschwender (2005) offers an informal discussion of crowding. Initial papers dealing with rail (and bus) transport were mainly concerned by the optimal service frequency and vehicle capacity (cf. Mohring, 1972; Jansson, 1980; Rietveld et al., 2001). An approach of congestion in public transportation based on multi-prize contests, has been proposed recently by de Palma and Soumyanetra (2012).

The value of time (VOT) in mass transit depends on riding conditions. There are mainly three different situations a passenger may face:

- A seat is available, and in that case the value of time can be assumed to be independent of the number of passengers in the vehicle.

- The passenger has to stand but the vehicle is not crowded. In that case the VOT will be higher but again constant.

- The passenger has to stand and there are too many passengers in the vehicle. In that case the VOT, or the riding conditions, would depend on

the number of passengers.

Similar situations were considered in Lam et al. (1999), where they distinguish between three discrete situations in their empirical analysis. The development of tractable micro-economic theory that includes riding comfort and crowding in mass transit could benefit from a mathematical expression reflecting these alternative comfort situations. Our paper deals only with the stylized case of a homogeneous population that wants to make a fixed number of trips between an origin and a destination but can be generalized to the case of variable demand for trips, to the case of a network and to the case of a heterogeneous population. It offers also new options to plan and better optimize the capacity of mass transit. To the best of our knowledge, public transportation congestion is not modeled in commercial software which describes public and private transportation. The functional form we propose may help to remedy this situation. It is simple, tractable and can be explained intuitively.

In the next section we propose the discomfort function and analyze its properties. In Section 3, we use our formulation to analyze the properties of user equilibrium and optimal scheduling and pricing in a dynamic mass transit model for the case of a uniform desired arrival time. In Section 4, the discomfort formulation is used to analyze the scheduling in the case of a randomly distributed desired arrival time. In Section 5, we discuss possible generalizations.

2   Defining discomfort functions in mass transit

Time cost in public transit depends on riding conditions. The most comfortable situation is when a passenger has a seat. Not having a seat is not enjoyable but acceptable when there is no crowding and the trip is not too long. Discomfort becomes particularly important when too many passengers have to stand.

Most important notation and assumed numerical values are provided in Table 1. Let n^s denote the number of seats in the vehicle and let n^x denote the standing capacity. The standing capacity is sometimes defined by the manufacturer of the bus or by the regulating authority. But it is often exceeded at peak times. Vuchic (2005) distinguishes between five situations that range from "independent standing, easy circulation" when passenger density is less than one by one meter square to "crashes loads, possible injuries forced movements" when there are 6.7 passengers per meter square (cf. Vuchic, 2005, Table 1.2, page 12). Of course these standards vary when some passengers are carrying some luggages or strollers. We assume that the standing capacity still allows passengers to travel in comfortable conditions. What matters here is not how it is defined but that the discomfort of standing is increasing with the number of passengers. So, the total seating and comfortable standing capacity of the vehicle is $n^s + n^x$. We define user costs for a given standard trip that takes a given time. The user cost of the nth passenger, is given by

$$C(n) = \begin{cases} \alpha_0 & \text{if } n \leq n^s \\ \alpha_1 + b \, e^{c(n - n^s - n^x)} & \text{if } n > n^s, \end{cases} \tag{1}$$
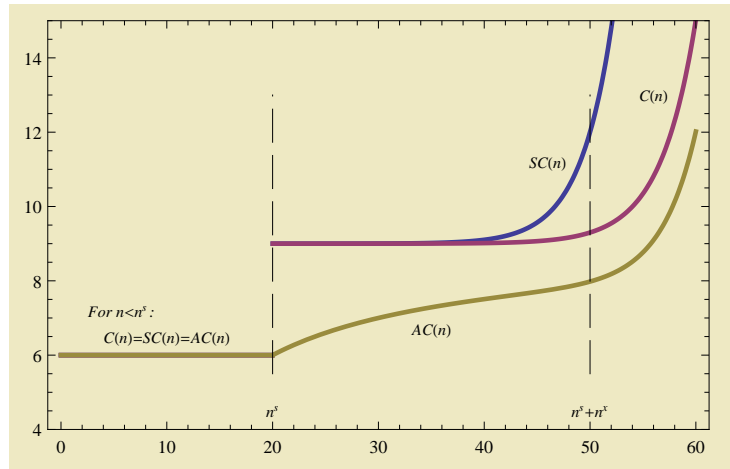
3

Figure 1: In-vehicle average cost for the passengers (based on values in Table 1).

where b and c are positive parameters that reflect how crowding impacts time cost. The user cost for those who can sit is α_0. The user cost for the passenger that have to stand is given by the second line in (1). It is equal to α_1, where $\alpha_0 < \alpha_1$, unless $n > n^x + n^s$ (the vehicle is crowded) and then it increases strongly. The total cost corresponding to (1) is

$$TC(n) = \begin{cases} n\,\alpha_0 & \text{if } n \leq n^s \\ n^s\,\alpha_0 + \left(\alpha_1 + b\,e^{c(n-n^s-n^x)}\right)(n-n^s) & \text{if } n > n^s. \end{cases} \quad (2)$$

The average cost is $AC(n) = TC(n)/n$, and the marginal social cost is obtained by differentiation of (2) with respect to n, i.e. $SC(n) = dTC(n)/dn$, except at point $n = n^s$ where it is not defined. These functions are illustrated on Figure 1. Notice that both $C(n)$ and $SC(n)$ are discontinuous at point n^s, and defined using a conditional statement. In practice a continuous formulation is generally preferred, whenever possible. In Appendix A, we discuss possible issues on how to approximate Eq. 1 with a smooth function.[1]

For $n < n^s + n^x$ and a positive value of parameter c the term with exponential is very small and can be neglected. As the number of passengers increases and becomes higher than the "capacity" of the vehicle ($n^s + n^x$), crowding increases and this is captured in the exponential term.

If there are less than n^s passengers then they all have a seat and the marginal passenger contributes to total travel cost by α_0. If the number of passengers is between n^s and n^x, the marginal social cost is almost constant and equal to α_1. For more than n^x passengers the marginal social cost increases reflecting crowding and the difficulty to get into the vehicle. Notice that user cost is not the same for all passengers: those who have a seat have lower travel cost.

---

[1] Instead of the term $\exp(c\,(n-n^s-n^x))$ in Eq. 1 one could use $\exp(c\,\max(n-n^s-n^x, 0))$. Appendix B shows that both formulations lead to similar impacts.

Parameter	Comment	Illustrative value
α_0	VOT with seat	6 ($/hour)
α_1	VOT without seat	9 ($/hour)
β	Early arrival penalty	5 ($/hour)
γ	Late arrival penalty	12 ($/hour)
n^s	Number of seats	20 (seat)
n^x	Standing capacity (legal)	30 (passenger)
$n^s + n^x$	Vehicle capacity (legal)	50 (passenger)
b, c	Discomfort parameters	(0.3, 0.3)
t_i	Departure time of train i	–
T	Travel time	0.25 (hour)

Table 1: Parameter values.

3 User equilibrium, scheduling and optimal pricing with identical desired arrival time

In this section we explore optimal scheduling and pricing in a simple dynamic model where some users want to travel via mass transit from one origin to a given destination. We study first the simpler case of identical desired arrival times.

The objective is to illustrate how crowding in rail or bus system forms in peak hours. On the side of passengers, we assume a group of N individuals who have the same desired arrival time window $(\underline{t}, \overline{t})$, and who incur a schedule delay cost whenever they arrive too early (before time $\underline{t}$) or too late (after time $\overline{t}$) at their destination. Let t denote the actual arrival time of a given passenger. We consider the following penalty function:

$$f(t) = \begin{cases} (\underline{t} - t)\,\beta & \text{if } t < \underline{t} \\ 0 & \text{if } \underline{t} < t < \overline{t} \\ (t - \overline{t})\,\gamma & \text{if } t > \overline{t}, \end{cases} \tag{3}$$

where β and γ are the schedule delay parameters, respectively, for early and late arrival. It is usually assumed that $\beta < \gamma$, i.e. the penalty of an early arrival is lower than the penalty of a late arrival. When the travel speed in the rail system is constant, we can assume without loss of generality that the arrival times are also the departure times of the train. As in Kraus and Yoshida (2002), we assume successive departures of the train at times $t_i = \underline{t} + \delta\,i$, where

$i = \ldots, -2, -1, 0, 1, 2, \ldots$, and δ is the technical time interval between two departures. Denote by n_i the number of passengers that select the train at time t_i. We take a simple configuration where all passengers take the train in the same (single) station and have the same destination. We discuss the user equilibrium, the system optimum and the optimal pricing case.

3.1 User equilibrium

At a user equilibrium, each passenger's objective is to minimize his own travel cost plus schedule delay cost, called generalized user cost. If there is no crowding (unlimited capacity) then the best solution for all passengers is to take a train that departs in time interval $(\underline{t}, \overline{t})$. With crowding, each passenger will trade off the schedule delay cost with discomfort cost to select the best departure time. At equilibrium no passenger will have an incentive to change his departure time. The individuals will not take into account their external costs, so that the user equilibrium may differ from the system optimum. The total in-vehicle cost in train i is $TC(n_i)$, and the average cost is $TC(n_i)/n_i$, where n_i the loading of train i. Taking into account the schedule delay cost, the generalized user cost in train i is $C(n_i) + f(t_i)$. At equilibrium no passenger has an incentive to switch to another train and this obtains when generalized user costs are equalized among all trains. Formally, we have

$$C(n_i) + f(t_i) = \overline{c},$$

where $\overline{c}$ is the generalized user cost in all the trains.[2]

3.2 System optimum

The optimum distribution of users over trains minimizes total transport cost, i.e. $\sum_i (TC(n_i) + f(t_i) \, n_i)$, where n_i represents the number of users of the train leaving at time t_i. Travel cost considered here is based on the MAS cost function given in (1) that represents the discomfort costs as a function of arrival order. Total number of users is fixed and we have $\sum_{i=1}^{k} n_i = N$. The optimal occupancy rate in each train is determined so that total costs are minimized. We can formulate this problem as an unconstrained minimization program if we substitute for $n_k = \sum_{i=1}^{k-1} n_i$. If the optimum solution is an interior solution where $n_i > n^s$, we can use first-order condition for a given train i (where $i = 1, \ldots, k-1$), that is

$$SC(n_i) + f(t_i) = SC(n_k) + f(t_k).$$

---

[2] If there are k trains, $i = 1 \ldots k$, we have a system of $k+1$ nonlinear equations: k equations, one for each train stating $C(n_i) + f(t_i) = \overline{c}$, plus the condition that each user chooses only one train $n_1 + \cdots + n_k = N$. The unknowns are the train loadings n_i and the generalized travel cost $\overline{c}$. To find a solution to this problem, it suffices to solve the first k equations by considering $\overline{c}$ as a parameter, and then find the value of $\overline{c}$ that yields the condition on the total number of users. The solution is clearly unique.

This condition implies that, at the optimum, the marginal cost of a new passenger in-vehicle i is equal to the marginal social gain obtained from removing the same user from vehicle k. The optimality conditions form a set of $n - 1$ non-linear equations that can be solved numerically. We illustrate the user equilibrium and the system optimum in Table 2 using again the parameters of Table 1. Remember that in-vehicle travel cost is $C(n) \cdot T$, where T is the travel time, here equal to 15 minutes. In the same table we also illustrate the solution where the users are distributed uniformly over the trains (one third of the total population in each train).

The difference between the user equilibrium and the system optimum is the suboptimal allocation of users over vehicles. The system optimum is to load the trains arriving too early or on time more or less like the trains arriving too late and the occupancy rate is not too different from uniform (compare first and third lines in Table 2). There are two forces in play: discomfort in the vehicle and schedule delay cost. The optimal solution reaches the best trade off between these two forces. The reduction of schedule delay cost made possible by loading more passengers on the first trains is limited by the crowding it will induce. Since early delay cost is smaller than late delay cost in the example of Table 3 (first line) there are more passengers in the first train than in the third one.

Consider now the user equilibrium. Users will make efforts to improve their comfort and will try to minimize the user cost also by picking the wrong train. They disregard the extra in-vehicle comfort costs they generate for the other users. This leads to too full trains that arrive too early and just in time. In the example of Table 2 (line 2) there is more crowding in the first and second train and much less passengers in the third one. This suboptimal distribution of passengers over the trains leads to a much higher total travel cost.

Comparing the uniform distribution of passengers over trains (third line) with the system optimum tells us that, in our example, the inefficient allocation over trains is less important in terms of efficiency than the efforts of the users to improve their in-vehicle comfort. In this case, the travel cost in equilibrium is higher by more than 35% than the cost obtained with the optimal loadings of the trains. The travel cost in the uniform distribution is higher only by less than 3%.

Notice however that the uniform distribution will not be as good as in this illustration if the total population were smaller and capacity not fully used. In that case it is natural to lower cost by putting more passengers in the middle train since schedule delay cost is reduced but crowding is not increased. A similar point is discussed below in Section 4.2 below.[3]

3.3   Optimal pricing

The source of inefficiency is the inefficient allocation of passengers over the different trains. This requires a differentiation of the charges for the different

---

[3] The interested reader may check the Mathematica notebook accompanying this paper that we make available from `http://perso.univ-lille3.fr/~mkilani/codes/`

	Train loadings			Average costs		
	1	2	3	C	SD	Total
	−30 mins	on time	+30 mins	($/day/passenger)		
System Optimum	53.5	55.1	41.4	2.08	1.70	3.78
User Equilibrium	61.5	63.3	25.3	3.75	1.36	5.11
Uniform	50.0	50.0	50.0	1.99	1.89	3.88

Table 2: Train loading and user costs under three alternatives (parameters from Table 1). Abbreviations: "C" stands for crowding, and "SD" for schedule delay. Travel time is 15 minutes.

trains so that the system optimum condition is satisfied

$$C(n_i) + f(t_i) = C(n_k) + f(t_k), \qquad i = 1\ldots, k-1.$$

In order to internalize the crowding externalities we need to charge more the trains with more crowding or alternatively those trains with the best arrival times.

4 User equilibrium, scheduling and optimal pricing with randomly distributed desired arrival times

We turn now to the more general case where passengers differ in their desired arrival time. We consider a uniform distribution as it allows to derive analytical expressions. We start by studying the passenger choice facing two departure times t_A and t_B both located in time interval $(0,1)$ and $t_A < t_B$. Next we optimize the departure times for multiple trains, and discuss pricing alternatives. In order to simplify mathematical expressions, we assume that the travel time is set equal to one.

4.1 User equilibrium and system optimum with 2 trains

Consider the choice between two trains. Train A leaves at $t_A \in (0,1)$, while train B leaves at t_B where $t_A < t_B < 1$. Assume that travel times are normalized to zero (constant travel speed), and that the desired arrival times are continuously distributed in $(0,1)$, with density ρ. We first compute the equilibrium average user costs for any t_A, t_B. In a second step we optimize the departure times. For

an equilibrium it is necessary that the last entrant is indifferent between the two trains. Denote the departure time of the user indifferent between the two trains A and B by t^*. The most interesting case is where $t^* \in (t_A, t_B)$. The number of users of train A is ρt^*. The generalized cost of user t^* is:

$$
\begin{aligned}
C_A^G &= C(n_A) + \beta(t^* - t_A) \\
&= C(\rho t^*) + \beta(t^* - t_A).
\end{aligned}
$$

Similarly, the generalized cost for using train B, is

$$
\begin{aligned}
C_B^G &= C(n_B) + \gamma(t_B - t^*) \\
&= C(\rho(1 - t^*)) + \gamma(t_B - t^*).
\end{aligned}
$$

The indifference condition for a user equilibrium reads:

$$
\begin{aligned}
C(\rho t^*) + \beta(t^* - t_A) &= C(\rho(1 - t^*)) + \gamma(t_B - t^*) & (4) \\
\text{or,} \quad (\beta + \gamma)t^* &= \gamma t_B + \beta t_A + C(\rho(1 - t^*)) - C(\rho t^*).
\end{aligned}
$$

The form of $C(n)$ precludes a general analytical solution but we can inspect its properties numerically, as illustrated in Figure 2(a). The case of a small small enough ρ (when all passengers have a seat and incur the same travel cost) can be computed explicitly, since then $C(\rho t^*) \to \alpha_0$ and $C(\rho(1 - t^*)) \to \alpha_0$. Therefore, in this case, equation (4) reduces to:

$$
\begin{aligned}
\beta(t^* - t_A) &= \gamma(t_B - t^*) \\
t^* &= \frac{\gamma t_B + \beta t_A}{(\beta + \gamma)}.
\end{aligned}
$$

For $\gamma > \beta$, we have $t^* < 1/2$. The reason is that transit users prefer to be one minute too early rather than one minute too late. $\gamma > \beta$ is the usual assumption and is empirically sound. This is the solution when congestion (crowding) does not matter ($\rho = 0$). As the density parameter ρ increases there will be more and more passengers in train B. We check from (4) that

$$
\frac{dn_A}{d\rho} < 0, \text{ if } \frac{\gamma t_B + \beta t_A}{(\beta + \gamma)} > 0.5,
$$

and at the limit we have $\lim_{\rho \to 1} \widehat{t}^* = 0.5$, where $\widehat{t}^*$ denotes desired arrival time with crowding for the passenger indifferent between train A and train B. As expected, the effect of crowding is to equalize the number of users in each train. This statement also means that an increase of total demand for mass transit will decrease the number of users of train A if train A transports already more than 50% of total demand.

In Figure 2(a) the two curves correspond to the number of users in train A, under equilibrium and optimal regimes. When, the number of users is small (by comparison to the vehicle capacity) external costs are small and the equilibrium outcome is optimal. As the number of users increases, the comfort decreases,

9
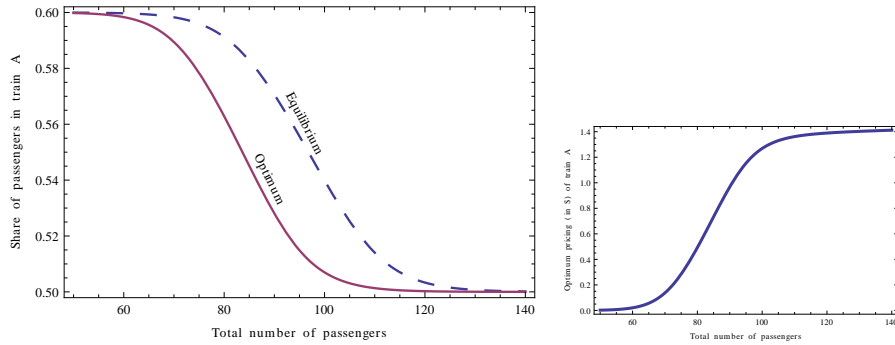
(a) Equilibrium and optimum train loadings in Train A. (b) Optimum pricing of train A.

Figure 2: Comparison of optimum and equilibrium when $t_A = 0.2$ and $t_B = 0.8$ and other parameters from Table 1.

still, train A remains overused since each user considers his own cost, not the external crowding cost imposed on the other passengers. As the number of users continues to increase, crowding becomes the main concern and passengers are almost equally split between the two trains, both at equilibrium and at the optimum. This figure, however, hides a particular detail. From the fact that equilibrium loadings converge to optimum, one may conclude that pricing is no longer required in this case, or more precisely that optimum pricing of train A is $p_A \to 0$, as ρ gets larger. This is a false conclusion. Indeed, optimum pricing is increasing in ρ as shown in Figure 2(b). The reason is that as the number of passengers increases, crowding increases strongly. Even if the difference in the two quantities is small, it still induces an increasing differential in cost. At the same time, one must be careful in the practical implication of this "theoretical result". Indeed, in practice the number of passengers is integer, so when two quantities are close it means that they are equal, and indeed no pricing is needed to adjust quantities. In the next section we study optimal departure times and show that with optimized departure times and uniformly distributed arrival times, price differentiation over time is again not needed.

4.2   Optimal departure times for k trains

For an early study of optimal time tables see de Palma and Lindsey (2001). We treat the case of a single train. Total number of users is fixed and so is the crowding discomfort and it plays no role in the optimization computation. The total schedule delay cost is: $\int_0^{t_A} \gamma(t_A - t)g(t)dt + \int_{t_A}^1 \beta(t - t_A)g(t)dt$. The minimum cost is obtained at departure time t_A that satisfies

$$\beta \cdot N^{\text{early}} = \gamma \cdot N^{\text{late}}, \tag{5}$$

10

where $N^{\text{early}} = \int_{t_A}^{1} g(t)dt$, where $N^{\text{late}} = \int_{0}^{t_A} g(t)dt$, denotes the number of passengers that arrive after their desired arrival time. When $g(t)$ is uniform over $(0,1)$, we get a simple solution: $\gamma t_A = \beta(1 - t_A)$. Solving this equation for t_A we find that the train should depart at time $\beta/(\beta + \gamma)$. With k trains, the same optimal solution applies in each subinterval. The first order condition for a maximum gives the optimal departure time t_i (for $i = 1 \ldots k$), that is

$$t_i^{\text{unif}} = \frac{1}{k} \left[\frac{\beta}{\beta + \gamma} + (i - 1) \right]. \tag{6}$$

The main result is that *when users only differ in their preferred arrival time and the distribution of preferred arrival times is uniform, crowding does not modify the optimal departure times*. Notice however that when the distribution of desired arrival times is not uniform, this result no longer holds.

For the sake of comparison, consider this alternative nonuniform distribution of desired arrival times.

$$h(t) = \begin{cases} 4\rho t & \text{if} \quad 0 \leq t \leq 1/2, \\ 4\rho(1-t) & \text{if} \quad 1/2 < t \leq 1, \\ 0 & \text{elsewhere.} \end{cases} \tag{7}$$

Notice that $\int_{0}^{1} g(t)dt = \rho$, so we have the same number of users as in the case of a uniform distribution. With distribution $h(t)$, most users have desired arrival time near $1/2$. With only one train, Eq. (5) still applies, since there will be ρ passengers in the train independently of its departure time. Using this condition and the first order condition leads to

$$t_1^{\text{h}} = \frac{\beta}{\beta + \sqrt{\beta(2\gamma - \beta)}},$$

and we check that for $\beta < \gamma$, $T^{\text{g}} \in (0, 1/2)$.

With two trains and more, we can no longer use condition (5). Instead, total user cost should be minimized with respect to departure times of the two trains. Let $SD(t_A, t_B)$ denote total schedule delay cost for all users when the departure times of the two trains are at t_A and t_B, respectively. The optimal departure times minimize total cost $TC(n_A) + TC(n_B) + SD(t_A, t_B)$, where n_A and n_B denote the number of passengers in the first and second train, respectively. Notice that these train loadings depend on departure times t_A and t_B. Indeed, for any given t_A and t_B one has to find the user who is indifferent between the two trains in order to compute n_A and n_B. A solution to this problem cannot be derived analytically, given the nonlinear expression of $C(n)$, but may be solved using a numerical procedure. For a numerical illustration,[4] computing the solution with distribution $h(t)$ defined above and parameter values in Table 1, we find the solution values given in the second column in Table 3. Users are almost split equally between the two trains ($\bar{t} = .499$) and we check, as expected, that $t_1^{\text{unif}} < t_1^{\text{h}} < t_2^{\text{h}} < t_2^{\text{unif}}$.

---

[4] The computational details are given in the Mathematica notebook. See footnote [3].

	Optimizing over t_A and t_B with	
	same fares $(p_A = p_B)$	optimized p_A and $p_B = 0$
Average cost	8.68	7.61
t_A	0.276	0.256
t_B	0.587	0.558
$\bar{t}$	0.499	0.472
% arriving late	31.38	29
p_A	−	1.439

Table 3: Optimizing over departure times and fares.

4.3 Optimal pricing for k trains

We know that optimal pricing depends in principle on the departure times as they determine the levels of congestion and the external congestion costs. In a few cases, prices do not matter. When there is only one train, congestion level is fixed and prices do not matter as long as total demand is fixed. More generally, for any set of departure times, optimal prices are always equal to the marginal external congestion costs. When total demand is fixed only the differences in marginal external congestion costs of the different trains matters. Also, with a uniform distribution of desired arrival times there is no need for pricing to decentralize the optimum. Indeed, in this case, departure times are given by Eq. 6 and there are ρ/n passengers in each train. One can check that users with desired arrival time at i/n are indifferent between train $i - 1$ and train i. So, the private decision leads to the optimal choice.

Then, to discuss the case of nonuniform distribution of arrival times, we consider the case of two trains. Desired departure times distribution is $g(t)$ over $(0, 1)$. Total passenger cost has two parts: schedule delay cost and in-vehicle cost which depend on the loadings. Let N_A and N_B denote passengers in trains A and B, respectively. Let the fares[5] in the two trains be p_A and p_B, respectively, and let the train loadings be determined by $N_A = \rho \int_0^{\tilde{t}} g(t)dt$ and $N_B = \rho \int_{\tilde{t}}^1 g(t)dt$, where $t \in (0, 1)$. The social planner can choose $\tilde{t}$ by setting fares p_A and p_B conveniently. Given $\tilde{t}$, the departure times of the two trains can be determined on the basis of condition 5, respectively applied on $(0, \tilde{t})$ and

---

[5]When demand is not elastic, only one train need to be priced, with a value that may be positive or negative. So in this discussion one may assume that $p_A = 0$ and only consider fares on train B.

$(\tilde{t}, 1)$. Now let

$$\tilde{t}^* = \arg\max_{\tilde{t}} \widetilde{SD}(\tilde{t}),$$

where $\widetilde{SD}(\tilde{t})$ is total schedule delay cost for all passengers (in the two trains) where fares are set to make passenger $\tilde{t}$ indifferent between the two trains and where the departure times are based on condition 5 as explained above. So, $\tilde{t}^*$ indicates the train loadings that yield the minimum total schedule delay cost. Fares p_A and p_B obtained under this solution would be optimal fares when crowding in the vehicles is not considered.

Crowding is minimized when all the passengers are equally split between the two trains. If we start from the situation where only schedule delay cost matters and increase progressively the importance of crowding, the marginal passenger moves from $\tilde{t}$ to the median passenger $\overline{t}$ (we have $\int_0^{\overline{t}} g(t)dt = \int_{\overline{t}}^1 g(t)dt = 1/2$). A numerical illustration is given in the last column of Table 3. Comparing with the situation where the optimization of prices is not possible we see that departures times are slightly advanced. There are more passengers in the second train, and this allows more passengers to arrive earlier. Average user cost is (of course) smaller when prices are optimized.[6]

5   Conclusion

This paper has developed an analytical expression for the discomfort in mass transit. Our expression distinguishes between passengers with a seat and those who have to stand. For those who have to stand, the discomfort will depend on the number of standing passengers compared to the capacity of the vehicle. This formulation helps to derive optimal timetables and optimal user charges. The model presented in this paper is very simple and many improvements can be envisaged. In particular, we have omitted the waiting time and the fact that when a train arrives, heavily congestion, only a fraction of passengers are able to enter. The remaining passengers have to wait for the next train, incurring an extra waiting time. For empirical applications, the parameters of the model should be estimated in order to derive a congestion function for public transportation, comparable to the BPR function widely used for private transportation.

Acknowledgments

---

[6] Notice that since demand is inelastic, one only need to consider fares for one train (train A here).

References

de Palma, A. and Lindsey, R. (2001), Optimal Time Tables for public transportation, *Transportation Research B* **35**, 789–813.

de Palma, A. and Soumyanetra, M. (2012), Multi-player, Multi-prize, Imperfectly Discriminating Contests, Working paper hal-00683688, HAL.

Glazer, A. and Proost, S. (2010), Reducing rent seeking by providing wide public service, Technical report, Discussion paper series DSP 10.31, Leuven (Belgium): KULeuven CES.

Huang, H. J., (2000), 'Fares and tolls in a competitive system with transit and highway: the case with two groups of commuters, *Transportation Research Part E* **36**(4), 267–284.

Jansson, J. O. (1980), 'A simple bus line model for optimisation of service frequency and bus size', *Journal of Transport Economics and Policy* **14**(1), 53–80.

Jara-Díaz, S. R. and Gschwender, A. (2005), Making pricing work in public transport provision. In D. A. Hensher and K. J. Button, eds. *Handbook of Transport Strategy, Policy and Institutions*. Elsevier, Amsterdam, 447–459.

Judd, K. L. (1998), *Numerical methods in economics, MIT Press.*

Kraus, M. (1991), 'Discomfort externalities and marginal cost transit fares', *Journal of Urban Economics* **29**(2), 249–259.

Kraus, M. and Yoshida, Y. (2002), 'The commuter's time-of-use decision and optimal pricing and service in urban mass transit', *Journal of Urban Economics* **51**(1), 170–195.

Lam, W. H. K., Cheung, C.-Y. and Lam, C. F. (1999), 'A study of crowding effects at the Hong Hong light rail transit stations', *Transportation Research A* **33A**, 401–415.

Mohring, H. (1972), 'Optimization and scale economies in urban bus transportation', *American Economic Review* **62**(4), 591–604.

Parry, I. W. H. and Small, K. A. (2009), 'Should urban transit subsidies be reduced?', *American Economic Review* **99**(3), 700–724.

Proost, S. and Van Dender, K. V. (2008), 'Optimal urban transport pricing in the presence of congestion, economies of density and costly public funds', *Transportation Research Part A: Policy and Practice* **42**(9), 1220–1230.

Rietveld, P., Dekkers, J. and van Woudenberg, S. (2001), Choice of frequency and vehicle size in rail transport: Implications for marginal external costs, Tinbergen Institute Discussion Papers 01-109/3, Tinbergen Institute.

Rouwendal, J. and Verhoef, E. T. (2004), '2. second-best pricing for imperfect substitutes in urban networks', *Research in Transportation Economics* **9**(1), 27–60.

Vuchic, V. R. (2005), *Urban transit operations, Planning and Economics*, Wiley.

Wardman, M. and Whelan, G. (2011), '20 years of rail crowding valuation studies: Evidence and lesssons from british experience', *Transport Reviews* .

A Approximation of the MAS formula

A drawback of the MAS formula discussed in this paper is that it is not continuous at point n^s. Also, the conditional statement in definition (1) may lead to a complications in the practical implementation of the MAS formula. Whenever possible a smooth function is preferred. A continuous alternative may be obtained by replacing the original function by a good approximation. We discuss here how one can construct a function ψ that approximates the user cost given in Eq. (1).

There are several approximation procedures and techniques available. These are generally simple to apply to unidimensional functions. We use two standard techniques, a simple Chebyshev interpolating polynomials and a sophisticated implementation in Mathematica.[7] Both of these solutions are illustrated on Figure 3.

For Chebyshev polynomials, we have used the standard procedure as given in Algorithm 6.2 in Judd (1998). Let us denote this approximation function by ψ^c. The result shown in Fig. 3(a) with $m = 150$ (number of interpolation points) and $n = 50$ (polynomial degrees). The approximation quality remains poor, due mainly to the discontinuity at n^s. Increasing further the number of interpolation points does not improve the quality of the approximation. Increasing the degree of the interpolating algorithm leads to an instability in the output function and makes the computation much more complicate. The quality of this approximation is not good, and there are two problems. For some values of n, particularly around n^s, the values it generates aren't close to those of $C(n)$. This can be confirmed by measuring the error approximation $\int_0^{60} |C(n) - \psi^c(n)| dn$. The second problem is that an equation of the form $\psi^c(n) = A$, where A is positive number may have more than one solution (depending on the values of A). This occurs because the approximation here does not preserve the monotonicity of $C(n)$. Definitely ψ^c is not a good choice for the approximation of $C(n)$.

The sophisticated approximation is denoted ψ^l. Fig. 3(b) shows both C and ψ^l. In this case we obtain a better approximation. In particular, The error between the approximation and the original function is small. It is clear that $\psi^s(n)$ fits the original function $C(n)$ much better than ψ^c. This observation may be confirmed by computing the error measure $\int_0^{60} |C(n) - \psi^s(n)| dn$.

---

[7]The latter uses divided differences to construct Lagrange or Hermite interpolating polynomials.

(a) Chebyshev polynomials.

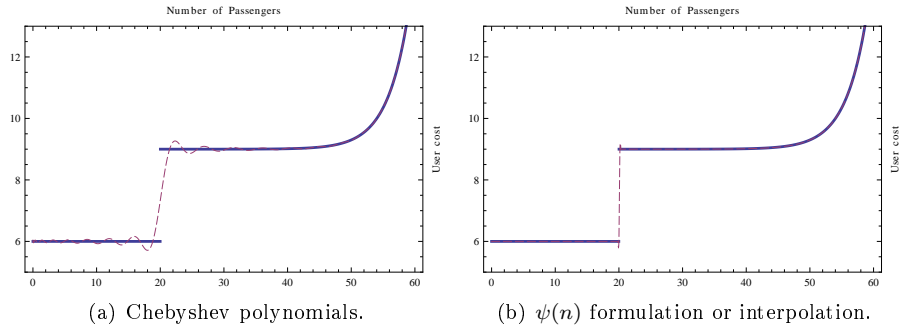(b) $\psi(n)$ formulation or interpolation.

Figure 3: Approximation of user cost function.

The explicit formula is actually a quiet long expression that can be handled by computers but not really useful for direct analytical usage. We can, however, provide a relatively simple formula that is comparable with respect to the error generated to the second approximation provided above. Indeed using the same notation as above, the function given by

$$\psi^{\mathrm{MAS}}(n) = \alpha_0 + \frac{\alpha_1 - \alpha_0}{1 + e^{a(n^s - n)}} + b\, e^{c\,(n - n^s - n^x)} \tag{8}$$

is a good approximation for the user cost as defined by Eq. 1. It is comparable to the function $\psi^s(n)$ defined above but has the merit of being very simple and avoids all conditional expressions. Apart, the problem of multiple solutions (as discussed for the case of ψ^s) it could be used for practical purposes.

B   Comparison with $\max$ function

The formulation in Eq. 1 uses the exponential form to take into account the fact that loadings below $n^s + n^x$ does not generate crowding. One may wonder whether we can replace the term $b \exp(c\,(n - n^s - n^x))$ by the simpler $b \exp(c \max(n - n^s - n^x), 0)$. Figure 4 shows that both expressions give similar curves for user cost and average cost (the dashed curve is the one using the max function).
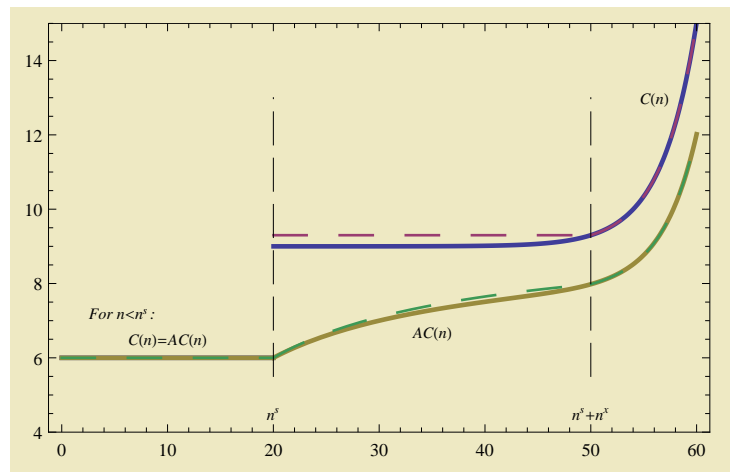
16

Figure 4: Comparison with initial formulation (bold curves) and max formulation (dashed curves) .