

OPTIMAL FARES AND CAPACITY DECISIONS FOR CROWDED PUBLIC TRANSPORT SYSTEMS¹

André de Palma, Ecole Normale Supérieure de Cachan
Robin Lindsey, University of British Columbia
Guillaume Monchambert, Ecole Normale Supérieure de Cachan

Introduction

Crowding on urban mass transit systems is common in both developed and developing countries. A roundtable report by the International Transport Forum identifies crowding as a major source of inconvenience that increases the generalized cost of travel (OECD, 2014). Crowding occurs not only while riding buses and trains, but also when boarding and alighting from them, when purchasing tickets, while waiting on platforms or at stops, and while accessing stations by escalator, elevator or on foot (King et al., 2014).

Several recent studies have documented the cost of crowding on transit networks. For example, Prud'homme et al. (2012) estimate that the 8% increase in densities in the Paris subway between 2002 and 2007 imposed a welfare loss of at least €75 million per year. Veitch et al. (2013) estimate the annual total cost of crowding in Melbourne metropolitan trains in 2011 at \$280 million.

Tirachini et al. (2013) provide a detailed review of the extensive literature on public transport crowding. Studies find that crowding increases in-vehicle time and waiting time, reduces travel time reliability, and causes stress and feelings of exhaustion. A number of studies document how disutility from in-vehicle time increases with the number of users.

There is a large operations research literature on public transit system design. An extensive economic literature has also developed on public transit capacity investments, service frequency, and optimal pricing and subsidy policy. These two branches of literature have made significant advances in understanding public transit systems. However, in contrast to the literature on automobile traffic congestion, most of the studies have employed static models that cannot account for travelers' time-of-use decisions and the dynamics of transit congestion and crowding.

The time profile of ridership is driven by the trade-off that users face between traveling at peak times and suffering crowding, and avoiding the peak by traveling earlier or later than they would like. A few studies² have explored this trade-off using simple microeconomic models that combine trip-scheduling preferences as introduced by Vickrey (1969) with a crowding cost function that describes how utility from travel decreases with passenger loads. In this paper we use this modeling framework to analyze usage of a rail transit line, and assess the potential benefits from internalizing crowding externalities by setting differential train fares. We also present results on optimal train capacity and the number of trains put into service.

The model

Consider a transit line that connects two stations without intermediate stops. The line operates on a timetable to which the operator adheres precisely. There are m trains, indexed in order of departure. Train k leaves the origin station at time t_k . The time headway between successive trains is a constant, h , so that $t_{k+1} = t_k + h$, $k = 1, \dots, m-1$. The headway is set at the minimum feasible value consistent with safe operation. Travel time aboard a train is independent of both departure time and train occupancy, and without loss of generality it is normalized to zero. Each train therefore arrives at the destination as soon as it leaves the origin station.

Each morning, a fixed number, N , of identical individuals use the line to get to work. Users know the timetable, and have to choose which train to take. They prefer to reach the destination at a common time,

t^* . If they arrive before t^* , they incur a disutility or “schedule delay cost” of β per minute they are early. For late arrival the penalty is γ per minute. Hence, the schedule delay cost of taking train k , $\delta(t_k)$, is

$$(1) \quad \delta(t_k) = \beta[t^* - t_k]^+ + \gamma[t_k - t^*]^+, \quad k = 1, \dots, m,$$

where $x^+ = \max\{0, x\}$. Users are assumed to board a train in random order, and (as in de Palma et al., 2015) thus cannot increase their chance of securing a good seat by arriving at the origin station early. The expected cost of crowding is $g(n)$, where n is the number of users taking the same train. Function $g(n)$ is an average over possible states: securing a good seat, getting a bad seat, having to stand in the middle of the corridor, standing close to the door, etc.. In this paper we assume that $g(n) = \lambda n/s$, where $\lambda > 0$ and $s > 0$ is a measure of train capacity. The linear form is supported by the meta-analysis of Wardman and Whelan (2011).

Let n_k denote the number of users on train k . A user taking train k incurs a combined schedule delay and crowding cost of

$$c_k = \beta[t^* - t_k]^+ + \gamma[t_k - t^*]^+ + \lambda \frac{n_k}{s}, \quad k = 1, \dots, m.$$

User equilibrium with a uniform fare

In this subsection we derive and characterize user equilibrium (UE) when N is fixed, and the fare is uniform; i.e. the same for all trains. A uniform fare does not affect either the division of users between trains or crowding costs. Let c^e denote the equilibrium trip cost net of fare, and n_k^e equilibrium ridership on train k . In UE, users distribute themselves between the m trains so that the user cost net of fare on every train is c^e . Given eq. (1) this implies

$$(2) \quad \delta(t_k) + \lambda \frac{n_k^e}{s} = c^e, \quad k = 1, \dots, m.$$

Since every user has to take some train,

$$(3) \quad \sum_{k=1}^m n_k^e = N.$$

Equilibrium ridership is solved by substituting eq. (2) into (3):

$$(4) \quad n_k^e = \frac{N}{m} + \frac{s}{\lambda} [\bar{\delta} - \delta(t_k)], \quad k = 1, \dots, m,$$

where $\bar{\delta} \equiv \sum_{k=1}^m \delta(t_k)/m$ is the unweighted average scheduling cost per train. Equation (4) reveals that timely trains (i.e., train that arrive closer to t^* and have lower $\delta(t_k)$) carry more users than other trains. The difference in passenger loads between two successive trains is proportional to train capacity, s , and the headway between trains, h , and inversely proportional to users' sensitivity to crowding, λ . The equilibrium user cost works out to

$$(5) \quad c^e = \bar{\delta} + \frac{\lambda N}{ms}.$$

Aggregate travel costs are easily derived. Let SDC denote total schedule delay costs, TCC total crowding costs, and TC total travel costs net of the fare. It is straightforward to show that

$$(6) \quad SDC^e = \bar{\delta} N - \frac{s}{\lambda} (\Delta - m\bar{\delta}^2),$$

$$(7) \quad TCC^e = \frac{\lambda N^2}{ms} + \frac{s}{\lambda} (\Delta - m\bar{\delta}^2),$$

$$(8) \quad TC^e = SDC^e + TCC^e = \bar{\delta} N + \frac{\lambda N^2}{ms},$$

where $\Delta \equiv \sum_{k=1}^m [\delta(t_k)]^2$ and $\Delta - m\bar{\delta}^2 > 0$ by the Cauchy-Schwarz inequality.

In equilibrium, total schedule delay costs are lower than if users were equally distributed across trains (in which case $SDC^e = \bar{\delta} N$) because users crowd onto timely trains that arrive closer to t^* . Total crowding costs are higher by the same amount, so that total costs are the same as if users were equally distributed.

We now derive the optimal uniform fare. The equilibrium private cost of a trip, p^e , equals the user cost plus the fare, τ :

$$(9) \quad p^e = \bar{\delta} + \frac{\lambda N}{ms} + \tau.$$

From eq. (8), the marginal social cost of a trip is:

$$(10) \quad MSC^e = \frac{\partial TC^e}{\partial N} = \bar{\delta} + \frac{2\lambda N}{ms}.$$

The first term in (10) is the same as in (5). The second term is proportional to N , and it is twice the corresponding term in (5) because the marginal social cost of crowding is twice the average cost. The average external cost is

$$MEC^e = MSC^e - c^e = \frac{\lambda N}{ms}.$$

With elastic demand, considered later, it is optimal to charge a uniform fare equal to the average external cost:

$$(11) \quad \tau^u = \frac{\lambda N}{ms},$$

where superscript “ u ” denotes the uniform-fare optimum. Total revenue from the optimal uniform fare is

$$(12) \quad R^u = \tau^u N = \frac{\lambda N^2}{ms}.$$

Given eqs. (11) and (9), the equilibrium private cost of a trip is

$$(13) \quad p^u = \bar{\delta} + \frac{2\lambda N}{ms}.$$

The uniform fare does not support the social optimum because the marginal external cost of crowding varies with train occupancy, and it is larger on more heavily used trains. The social optimum is examined in the following subsection.

Social Optimum

In the social optimum (SO), the marginal social cost of trips is the same on all trains. The marginal social cost of using train k is

$$MSC_k = \frac{\partial(c_k n_k)}{\partial n_k} = \delta(t_k) + \frac{2\lambda n_k}{s}.$$

Let superscript “ o ” denote the SO, and MSC^o the marginal social cost of a trip. At the optimum,

$$(14) \quad MSC_k = \delta(t_k) + \frac{2\lambda n_k}{s} = MSC^o, \quad k=1, \dots, m.$$

Since every user takes some train, the counterpart to eq. (3) holds:

$$(15) \quad \sum_{k=1}^m n_k^o = N.$$

Solving eqs. (14) and (15) one obtains

$$(16) \quad n_k^o = \frac{N}{m} + \frac{s}{2\lambda} [\bar{\delta} - \delta(t_k)], \quad k=1, \dots, m,$$

$$(17) \quad MSC^o = \bar{\delta} + \frac{2\lambda N}{ms}.$$

It is evident from eqs. (16) and (4) that train loads in the SO are more even than in the uniform-fare equilibrium. Spreading users is desirable because the cost of crowding is a quadratic function of load, and the benefits from reducing loads on timely trains exceed the costs of greater crowding on other trains.

The SO usage pattern can be decentralized by charging a fare on train k equal to the marginal external cost of usage. We will call this fare pattern the *SO-fare*. The marginal external cost of usage on train k is

$$MEC_k = MSC_k - c_k = \frac{\lambda n_k}{s}, \quad k=1, \dots, m.$$

The SO-fare on train k is therefore:

$$(18) \quad \tau_k^o = \frac{\lambda n_k^o}{s} = \frac{\lambda N}{ms} + \frac{\bar{\delta} - \delta(t_k)}{2}, \quad k=1, \dots, m.$$

Compared to the uniform fare in eq. (11), the fare in (18) is higher on timely trains and lower on the earliest and latest trains. Users of all trains incur a private cost equal to the social cost of a trip:

$$p_k^o = c_k^o + \tau_k^o = MSC^o = \bar{\delta} + \frac{2\lambda N}{ms}, \quad k=1, \dots, m.$$

The private cost is the same as the cost with the uniform fare in eq. (13). It is higher than the cost if no fares are charged. To see this, note that at least one train is more crowded in the SO than the UE. Compared to the UE, in the SO a rider of this train incurs the same schedule delay cost, but a higher crowding cost and a positive fare. Since all users incur the same private cost in the UE, and all users incur the same private cost in the SO, private costs are higher in the

SO. Unless fare revenues are used to improve service in some way, charging fares to price crowding costs leaves users worse off.

Total revenue from the SO-fare is $R^o = \sum_{k=1}^m \tau_k^o n_k^o$. Using eqs. (16) and (18) one obtains

$$(19) \quad R^o = \frac{\lambda N^2}{ms} + \frac{s}{4\lambda} (\Delta - m\bar{\delta}^2).$$

The first term in (19) matches revenue from the optimal uniform fare in (12). The second term is extra revenue due to variation of the fare. This will be called *variable revenue*, RV^o , where³

$$(20) \quad RV^o = \frac{s}{4\lambda} (\Delta - m\bar{\delta}^2) \cong \frac{s(\beta\gamma)^2}{48\lambda(\beta+\gamma)^2} h^2 m^3.$$

Other aggregate costs in the SO work out to:

$$(21) \quad SDC^o = SDC^e + 2RV^o,$$

$$(22) \quad TCC^o = TCC^e - 3RV^o,$$

$$(23) \quad TC^o = TC^e - RV^o.$$

Total schedule delay costs are higher in the SO than the uniform-fare or no-fare equilibrium. However, crowding costs are smaller by 1.5 as much, and total costs are lower by variable revenue.

As indicated in eq. (20), variable revenue is independent of N . The welfare gain from setting optimal fares is therefore independent of total usage. This may seem surprising since intuition would suggest that the welfare gain increases with N : first because crowding becomes more onerous for users on average, and second because more users suffer the higher cost. To understand the result, recall that the welfare gain arises from distributing users more evenly between trains. Since the difference in crowding costs between two successive trains equals the difference in schedule delay costs, the benefit from reallocating users between trains is independent of N . As N increases, marginal crowding costs on each train rise at the same rate. The uniform-fare component of the optimal toll in eq. (18) increases linearly with N , but the variable component does not change.

Another way to view the result is in terms of total costs. The marginal social costs of usage in the UE and SO are given in eqs. (11) and (17) respectively. Since the equations are the same, the marginal social costs are the same. In effect, the benefits of internalizing the crowding cost externality are exhausted once total usage is large enough for all trains to be used — as assumed throughout the paper.

In a longer version of this paper (de Palma, Lindsey and Monchambert, 2015) we have shown that the welfare gains from congestion pricing depend on the shape of the crowding cost function, $g(\cdot)$. If $g(\cdot)$ is convex, the welfare gain decreases with N , and if $g(\cdot)$ is concave, the welfare gain increases with N . Most empirical studies find that $g(\cdot)$ is either linear or convex. This suggests that the benefits of differentiating fares by train (or, equivalently, time of day) may be limited on heavily used transit systems. The model thus offers one explanation for why peak-load pricing on transit systems is not very common.

Optimal transit service with elastic demand

We now turn attention to the long run when the transit authority can choose the number of trains, m , train capacity, s , and the train timetable. We continue to assume that all trains have the same capacity, and that the headway is uniform and exogenous. When the schedule delay and crowding cost functions are linear, as assumed, the optimal timetable is straightforward to derive and, due to space constraints, it is not described here.

To investigate the optimal values of m and s we assume that the capital, operations and maintenance costs of providing service are described by a function $K(m, s)$ which is strictly increasing in m and s . To facilitate analysis, m will be treated as a continuous variable.

We now allow total usage to depend on the price, defined in eq. (9), and let $p(N)$ denote the inverse demand curve. With a uniform fare, social surplus net of capacity costs is

$$(24) \quad SS^e = \int_{n=0}^N p(n) dn - \left(\bar{\delta} N + \frac{\lambda N^2}{ms} + K(m, s) \right).$$

The transit authority chooses m and s to maximize SS^e . With a zero fare, the first-order conditions work out to:

$$(25) \quad \text{For } s \text{ with zero fare: } \frac{\lambda N^2}{ms^2} \frac{p_N N}{p_N N - \lambda N / ms} = K_s,$$

$$(26) \quad \text{For } m \text{ with zero fare: } \left(\frac{\lambda N}{m^2 s} - \frac{\partial \bar{\delta}}{\partial m} \right) N \frac{p_N N}{p_N N - \lambda N / ms} = K_m,$$

where K_s and K_m are derivatives of $K(m, s)$ with respect to s and m respectively. The first term of the product on the left-hand side of (25) is the marginal benefit from expanding train capacity if usage remained fixed. The average cost of crowding would decrease by $\lambda N / (ms^2)$ for the N users. The actual reduction in crowding is smaller than this because the improved service quality attracts new users. Because usage is underpriced, the increase in usage is welfare-reducing which shrinks the benefit from greater capacity. This latent demand effect accounts for the second term of the product on the left-hand side of (25) which is less than 1. In the limit of perfectly elastic demand (i.e., $p_N \rightarrow 0$), the potential benefit is completely dissipated. In the opposite limit of fixed demand (i.e., $p_N \rightarrow -\infty$), the second term converges to 1, and there is no dilution of benefit.

Equation (26) for m is interpreted similarly. The first term in brackets on the left-hand side is the marginal benefit from less crowding. The second term in brackets is the marginal disbenefit due to greater schedule delay costs. This net benefit is diluted by the same factor as in eq. (25). With the optimal uniform fare, the first-order conditions work out to:

$$(27) \quad \text{For } s \text{ with optimal uniform fare: } \frac{\lambda N^2}{ms^2} = K_s,$$

$$(28) \quad \text{For } m \text{ with optimal uniform fare: } \left(\frac{\lambda N}{m^2 s} - \frac{\partial \bar{\delta}}{\partial m} \right) N = K_m.$$

In contrast to eqs. (25) and (26), the marginal benefits from expanding service in eqs. (27) and (28) are not diluted by additional usage because usage is priced efficiently. This might suggest that the optimal values of s and m , s_*^o and m_*^o , are larger than their counterparts with a zero fare, s_*^n and m_*^n . However, at least for given values of s and m , usage is higher in the no-fare regime because the private cost of usage is lower. This leaves the rankings of s_*^o and s_*^n , and m_*^o and m_*^n , theoretically ambiguous in general.

With the SO-fare, social surplus net of capacity costs is given by

$$(29) \quad SS^o = \int_{n=0}^N p(n) dn - \left(\bar{\delta} N + \frac{\lambda N^2}{ms} + K(m, s) \right) + RV^o(m, s).$$

Equation (29) is the same as eq. (24) except for the last term, RV^o , which is a function of m and s , but does not depend on N . Since usage is priced efficiently in both the SO-fare and optimal uniform-fare regimes, the first-order conditions for s and m are the same as for the optimal uniform toll, (27) and (28), with the derivatives of $K(m, s) - RV^o(m, s)$ in place of the derivatives of $K(m, s)$. Hence,

$$(30) \quad \text{For } s \text{ in the SO: } \frac{\lambda N^2}{ms^2} = K_s - RV_s^o,$$

$$(31) \quad \text{For } m \text{ in the SO: } \left(\frac{\lambda N}{m^2 s} - \frac{\partial \bar{\delta}}{\partial m} \right) N = K_m - RV_m^o,$$

where RV_s^o and RV_m^o denote the derivatives of $RV^o(m, s)$ with respect to s and m respectively. The right-hand sides of eqs. (30) and (31) are smaller than their counterparts for the optimal uniform toll, (27) and (28). The generation of variable revenue from the SO-fare effectively reduces the marginal financial cost of expanding capacity. In the case of (30) this implies that optimal train capacity conditional on the values of m and N is larger in the SO: $s_*^o(m, N) > s_*^u(m, N)$. Similarly, eq. (31) implies that the optimal number of trains conditional on the values of s and N is also larger in the social optimum: $m_*^o(s, N) > m_*^u(s, N)$.

These rankings may seem surprising given that total system costs are lower in the SO than the uniform-fare equilibrium. Inequality $m_*^o(s, N) > m_*^u(s, N)$ is explained by the fact that ridership is distributed more evenly across trains in the SO. More users take the earliest and latest trains in the SO which makes adding extra trains more beneficial. The reason for the inequality $s_*^o(m, N) > s_*^u(m, N)$ is more subtle. It can be shown that while total user costs in the uniform-fare regime decrease with s , the deadweight loss from imbalanced ridership between trains increases with s . Expanding capacity is therefore more valuable in the social optimum. By contrast, in the Vickrey (1969) bottleneck model of road traffic congestion, optimal capacity is higher with a uniform fare.

A numerical example

Further properties of the model are difficult to derive analytically, and to proceed further we now adopt a specific capacity function that is based on Kraus and Yoshida (2002):

$$(32) \quad K(m, s) = (\nu_0 + \nu_1 s)m + \nu_2 s,$$

where ν_0 , ν_1 and ν_2 are all nonnegative parameters. The term $\nu_0 + \nu_1 s$ in (32) is the incremental cost of running an additional train, which includes both capital and operating costs. It is a linear increasing function of train capacity. If $\nu_0 > 0$, there are scale economies with respect to train size. The second term in (32), $\nu_2 s$, accounts for costs that depend on train capacity, but not the number of trains, such as terminal capital costs.

The numerical example draws on recent empirical estimates of crowding costs, and is loosely calibrated to describe service on the Paris RER A line during the morning peak.⁴ Base-case parameter values are: $\beta = 5$ [€/hr·user], $\gamma = 20$ [€/hr·user], $\lambda = 2.72$ [€/user], and $h = 2$ [min/train]. The demand function is assumed to have a constant-elasticity form $N = N_0 p^\eta$ with $\eta = -1/3$.⁵ Parameters N_0 , ν_0 , ν_1 and ν_2 are chosen to yield equilibrium values for the optimal

uniform-fare equilibrium of $N^u = 18,700$, $m_*^u = 24$, $s_*^u = 800$ and fare revenue equal to 5/6 of capacity costs. The resulting values are: $N_0 = 35,598$, $v_0 = 412.8$, $v_1 = 0.5058$ and $v_2 = 49.78$. Results for the three fare regimes are reported in Table 1.

With no fare, the equilibrium private cost (which equals the equilibrium user cost) is €4.54. There are $N^n = 21,500$ users who are accommodated in $m_*^n = 25.27$ trains with (nominal) capacities of $s_*^n = 811$. Total crowding costs are more than double total schedule delay costs. Capital costs (K^n) amount to a little over 60% of total user costs (TC^n). Given no fare, cost recovery, ρ , is zero.

The optimal uniform fare works out to $\tau^u = €2.65$. It boosts the equilibrium private cost to $p^u = €6.90$ which is €2.36 above the no-fare equilibrium price. Ridership drops to $N^u = 18,700$ which is 13% below the no-fare level. Both the number of trains and train capacity are also lower than with no fare although capacity costs are reduced by only 2.8%. Total crowding costs and total schedule delay costs are also lower than with no fare. Consumers' surplus is lower than with no fare, but social surplus is higher by €4,135 or about €0.22 per rider in the uniform-fare equilibrium. The relative efficiency of the optimal uniform fare can be measured as a fraction of the difference in social surplus between the SO and the no-fare equilibrium. It works out to about 0.82 so that the optimal uniform fare yields most of the efficiency gains from the SO-fare.

The SO features more trains than either the no-fare or the uniform-fare regime. However, train capacity is slightly lower than in the other two regimes. Ridership and consumers' surplus are slightly higher than with a uniform fare, whereas price, revenue per user and cost recovery are slightly lower. Crowding costs are significantly lower than in the other two regimes, but schedule delay costs are higher than with a uniform fare because the SO-fare spreads usage more evenly over trains. Capacity costs are intermediate between the

other two regimes. Social surplus is higher than with no fare by about €0.27 per rider.

	Fare regime		
	No-fare (n)	Opt. uniform (u)	Social opt. (o)
m	25.27	24	25.69
s	811	800	787
N	21,500	18,700	18,814
p	4.54	6.90	6.77
Rev/user	0.00	2.65	2.63
TCC	68,493	55,563	49,424
SDC	29,092	23,896	28,588
TC	97,584	79,459	78,012
K	61,164	59,447	60,029
R	0	49,539	49,424
ρ	0	0.833	0.823324901
CS	1,004,019	956,898	959,244
SS	942,856	946,990	948,638
Gain/ N^u		0.22	0.27
Rel.eff	0.00	0.822	1

Table 1: Comparison of no-fare, optimal uniform fare, and SO-fare (i.e., social optimum) regimes: base-case parameter values

The number of trains has been treated as a continuous variable although it is discrete in reality. The model was resolved by restricting m to integer values to obtain $m_*^n=25$ and $m_*^o=26$. Results were hardly affected, and social surplus was virtually unchanged.

To test the sensitivity of the results to the demand function, solutions were derived with different values of the elasticity. With $\eta = 0$, the uniform fare yields no welfare gain at all and merely transfers money from users to the transit authority. The SO-fare yielded a welfare gain of only €0.097 compared to €0.27 in the base case. With $\eta = -2/3$,

the optimal uniform-fare yields a much higher welfare gain of €0.42. The welfare gain from the SO increases too to €0.51, and the relative efficiency of the uniform-fare regime is little changed.

Conclusion

In this paper we have studied the time profile of ridership on a crowded rail transit line. We solve for the user equilibrium and social optimum when supply is fixed, as well as the long run when service can be optimized. Some of the results parallel those obtained with road traffic congestion models. Passenger loads are distributed more evenly across trains in the social optimum than the user equilibrium. The social optimum can be decentralized by charging higher fares on more popular trains to internalize the crowding cost externality on each train. Imposing differentiated fares makes users worse off --- at least before accounting for how the revenues are used. Other results are less obvious. The welfare gains from tolling are independent of total ridership. Expanding the number of trains can also be more valuable in the social optimum than the user equilibrium even though total system costs are lower in the social optimum.

Bibliography

- de Palma, A., Kilani, M., and Proost, S. (2015), Discomfort in mass transit and its implication for scheduling and pricing, *Transportation Research Part B*, 71(1), 1-18.
- de Palma, A., Lindsey, R., and Monchambert, G. (2015), The economics of crowding in public transport. Working paper.
- Huang, H.J., Tian, Q., and Gao, Z.Y. (2005), An equilibrium model in urban transit riding and fare policies, *Lecture Notes in Computer Science*, 3521, 112–121.
- King, D., Srikukenthiran, S., and Shalaby, A. (2014), Using simulation to analyze crowd congestion and mitigation at Canadian subway interchanges: case of Bloor-Yonge Station, Toronto, Ontario, *Transportation Research Record*, 2417, 27-36.
- 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.

- OECD (2014), Valuing convenience in public transport. Technical report, ITF Round Tables.
- Oum, T. H., Waters II, W. G., and Fu, X. (2008), Transport demand elasticities. In: Hensher, D., Button, K. (Eds.), Handbook of Transport Modelling, 2nd edition. Oxford: Elsevier.
- Prud'homme, R., Koning, M., Lenormand, L., and Fehr, A. (2012), Public transport congestion costs: the case of the Paris subway, *Transport Policy*, 21, 101-109.
- Tian, Q., Huang, H. J., and Yang, H. (2007), Equilibrium properties of the morning rush-hour commuting in a many-to-one mass transit system, *Transportation Research Part B*, 41(6), 616-631.
- Tirachini, A., Hensher, D. A., and Rose, J. M. (2013), Crowding in public transport systems: effects on users, operation and implications for the estimation of demand, *Transportation Research Part A*, 53, 36-52.
- Veitch, T., Partridge, J., and Walker, J. (2013), Estimating the costs of over-crowding on Melbourne's rail system. 36th Australasian Transport Research Forum. Brisbane, Queensland, Australia.
- Vickrey, W. S. (1969), Congestion theory and transport investment, *American Economic Review*, 59(2), 251-260.
- Wardman, M., and Whelan, G. A. (2011), Twenty years of rail crowding valuation studies: evidence and lessons from British experience, *Transport Reviews*, 31(3), 379–398.

¹ This is an abridged version of de Palma, Lindsey and Monchambert (2015) which provides a more general and formal treatment. For helpful comments we would like to thank participants at the Annual Conference of the International Transportation Economics Association (ITEA) in Toulouse, June 2014, seminar participants at the Department of Civil and Environmental Engineering, the Hong Kong University of Science and Technology, July 2014, and seminar participants at the Department of Spatial Economics, Free University of Amsterdam, February 2015.

² Huang et al. (2005), Tian et al. (2007), de Palma et al. (2015).

³ The last expression in (20) is obtained when the timetable is optimized.

⁴ See http://en.wikipedia.org/wiki/Paris_M%C3%A9tro_Line_1 and http://www.debatpublic-reseau-grandparis.org/site/DEBATPUBLIC_GRANDPARIS_ORG/_SCRIPT/NTSP_DOCUMENT_FILE_DOWNLOADDCB59.PDF

⁵ An elasticity of $-1/3$ is in the mid-range of empirical estimates (Oum et al., 2008). Consumers' surplus is infinite with $\eta > -1$. To enable comparisons of consumers' surplus between regimes, the area under the demand curve is computed for $p \leq \text{€}100$.