The location and pricing of mass transit stations

André de Palma
ENS Cachan
61 avenue du Président Wilson
92230 Cachan cedex, France
Phone: +33 6 63 64 43 20
Fax: +33 1 47 40 24 60
Email: andre.depalma@ens-cachan.fr

Moez Kilani
EQUIPPE/IUT B, Université Charles de Gaulle - Lille 3
Maison de la recherche
Domaine universitaire du Pont de Bois
59653 Villeneuve d'Ascq
Phone: +33 01 20 41 68 34
Email: moez.kilani@univ-lille3.fr

Stef Proost
Catholic University of Leuven
Naamsestraat 69
BE-3000 Leuven, Belgium
Phone: +32 16-32 67 25
Fax: +32 16-32 67 96
Email: Stef.Proost@econ.kuleuven.ac.be

May 2010

Abstract

This paper presents an analytical model for the optimal location and pricing of one or more mass transit stations in a monocentric city. We compare location and pricing decisions for transit stations by a public agency, a private monopoly and a duopoly where each private operator controls the location and pricing of one station. Analytical as well as numerical results are derived for different distributions of residential density.

1 Introduction

Transport economists have analyzed extensively the optimal pricing of mass transit as mass transit is often considered as part of the solution to urban congestion problems. These studies take the demand function for transit in general as given. One of the main determinants of the demand for transit is the location of the transit stations. In this paper we consider explicitly the location and access pricing of a limited number of mass transit stations in a linear monocentric city.

We address the following research questions. First what are the main determinants of the welfare optimal location of a single or multiple mass transit stations? How do factors as relative speed of private and transit transport affect the location, and what is the role of congestion in transit and private transport? Second, if we allow private operators to choose location and access prices of mass transit stations, can we expect them to make very different decisions?

Discussion of location and pricing requires highly stylized models. We use a linear monocentric model with homogeneous and immobile residents that all want to commute to the CBD. The CBD can be reached by a private mode or by mass transit and users' decide on the basis of their generalized costs. As mass transit can only be accessed in a limited number of stations, a mass transit trip requires always first a private transport trip to the transit station. Accessing the CBD by private mode and by mass transit imply both congestion that takes the form of queuing, schedule delay and other discomforts. Once one specifies the construction and operation costs of the mass transit, one can study the optimal location of one single mass transit station.

We determine the optimal location that balances the full costs of both modes for specific residential density functions. Take the case of one station, in a monocentric station this means radial lines with one station in the CBD and one station in the residential area. In the easiest case with a uniform residential density one finds that the unique station is never located beyond 2/3 of the radius of the city. Higher mass transit infrastructure costs or lower relative speed compared to the private mode imply a location closer to the CBD. Another interesting feature of the one station solution is the location and pricing decisions of a private monopolist or of a public agency with a very demanding revenue requirement. We find that the private monopolist is not interested in serving customers that live in between his station and the CBD because this requires a fare concession that is too costly in terms of his customers living beyond the transit station.

When one considers multiple stations, we find that the discomfort of an extra stop of the train to the CBD results in optimal locations further away from the CBD. We also find that the locations chosen by the monopolist are independent of the relative speed of the train while an increase of the speed points to an optimal location of stations further away from the CBD. The main reason is again that a monopolist is not interested in serving residents located in between the station and the CBD.

Numerical illustrations show the welfare optimal and profit maximizing lo-

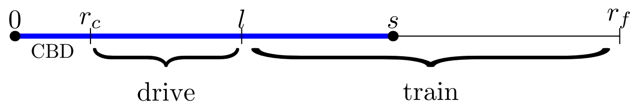

Figure 1: Spatial organization with one station.

cations for one, two, four and seven stations as well as the optimal frequencies and prices. The analysis is performed for the case of linear and non-liniear congestion.

Our paper can be compared with the work by Kraus and Yoshida (2002). They model the use of mass transit by commuter's who choose the optimal departure times in a bottleneck type of model. In the main sections of our paper, the departure time is not considered explicitly and congestion is a linear function of total use, we show that our conclusions can be extended to a bottleneck type of formulation in Appendix A. Another reference point in this literature is Kaus (1991) who models the discomfort externalities of a mass transit with several stations. Kraus and Yoshida and Kraus concentrate on optimal frequency and pricing, we also analyze the optimal location of stations.

Our paper takes the residential density and size of the city as given and concentrates on the location and operation of mass transit. Very few attempts have been made to address the two problems together (e.g. De Lara et al. 2008).Considering endogenous location and use of two competing modes as in Anas and Moses (1979), Baum-Snow (2007), preferably in a growing city would be the next step.

2 One station

2.1 Basic model

We work with a linear monocentric city where all residents want to travel to the CBD every day and this is the only type of trips considered. All residents are identical and their location is fixed. There are only two types of travel modes: mass transit and private. We refer to the private mode as driving and to mass transit as train transportation. Fig. 1 illustrates the notation used. The distance from the city center is denoted r and r_c is the radius of the CBD. We assume, w.l.o.g., that $r_c = 0$. The radius of the city is r_f. The density of the population is denoted $r(n)$.We assume that there is only one transit station located at s. l is the location of the passenger who is indifferent between transiting by train or driving directly to the CBD (the marginal user).

There is congestion in mass transit and in private transport. Congestion in mass transit is modeled as a linear function of the number of users. Congestion in private transport is represented as a linear function of the total number of drivers entering the CBD. The generalized cost of driving per unit distance is c_p (operation plus time). The corresponding generalized rail cost is c_t, and to simplify notation, all costs are normalized by c_p so $c_p = 1$. Average waiting time in the station is w and waiting cost is c_w. We assume that early and late delays have the same cost. Appendix A generalizes this feature. Residents living between l and r_f drive to the station and then commute to the CBD by train. When the user transits by train, he incurs a transfer cost, denoted K, that reflects any difference in fixed costs between the two transport alternatives (parking costs, access to train, station, etc). Users of mass transit incur each a congestion cost Γ_t. We are now able to compute the total travel cost for a passenger located at r and using rail:[1]

$$p + |r - s| + s\,c_t + w\,c_w + w\,\gamma_t \int_l^{r_f} n(r)dr \qquad r \in (l, r_f). \tag{1}$$

The first term is the users' fare for the mass transit. The second term is the private mode cost to drive to the station. The third and fourth terms represent the time cost and the waiting cost of using the train. The last term is the average congestion cost in mass transit. $n(r)$ represents the density of residents at distance r from the CBD. We work with a uniform density if not specified otherwise.

Those that do not travel by rail, drive to the CBD: residents living between 0 and l drive directly the CBD. The total travel cost for a driver is the sum of a road toll τ, a private transportation cost equal to r (remember that unitary private cost is normalized to one), and a congestion cost incurred at the entrance of the CBD that is linear function in the number of drivers (γ_p times number of drivers). Transport cost in this second group is thus:

$$\tau + r + \gamma_p \int_0^l n(r)dr, \qquad r \leq l. \tag{2}$$

The user who is indifferent between the two modes equates both travel costs. The housing location of this passenger, l, solves

$$p + s - l + s\,c_t + w\,c_w + w\,\gamma_t \int_l^{r_f} n(r)dr = \tau + l + \gamma_p \int_0^l n(r)dr. \tag{3}$$

We have the following preliminary result.

Lemma 1 (No congestion). *If $\gamma_t = \gamma_p = 0$ and if $p - \tau + w\,c_w \leq (1 - c_t)\,s$, the user indifferent between using only private and combination of private and mass transit transportation is located at l given by*

$$l^{nc} = \frac{p - \tau + (1 + c_t)s + w\,c_w}{2}. \tag{4}$$

If $p - \tau + w\,c_w \geq (1 - c_t)\,s$, then no user selects the train.

[1] The implicit assumption is that the value of travel time is not identical for sitting in a train or walking.Empirical work, e.g. Wardman (2001) and Mackie et al. (2003), recommend to use a value for walking time double that of in-vehicle time.The value of time in vehicle will also depend on the comfort conditions in the train.

Assume $p = \tau = w = \gamma_p = \gamma_t = 0$. If the relative cost of train is small, i.e. $c_t \to 0$, then all households located beyond $s/2$ will transit by train. The other users, those between 0 and $s/2$, drive directly to the CBD. As the relative cost of the train increases, the number of those who drive increases until, at $c_t = 1$, all users no longer have any gain from transiting by train. From (4), for all $s > 0$, we have $l > 0$. So, there exists always some passengers who do not transit by train, even if s is arbitrarily close to the city center.

Now consider again the case $p, \tau \geq 0$ and let us assume for simplicity that congestion does not matter, i.e. $\gamma_p = \gamma_t = 0$. We can then compute the gain from having mass transit for two groups. The first group are those located at s or beyond. The cost saving for them is:

$$\Delta C_1 = (1 - c_t)s - (p - \tau). \tag{5}$$

Notice that, in this case, all passengers have the same gain from the new transport mode (ΔC_1 does not depend on the household's location r). For the second group, the rail users living between the CBD and s (segment (l, s)), the gain depends on housing location r:

$$\Delta C_2 = -(1 + c_t)s + 2r - (p - \tau) = 2(r - l), \tag{6}$$

which is increasing in r (those who live closer to the station gain more). We see that the user's gain from having a mass transit is higher for those living beyond the station. Indeed, from (5) and (6), we have $(\Delta C_1 - \Delta C_2) = 2(s - r) \geq 0$, since $r \leq s$.

The third group are those living between the CBD and l and they drive to the CBD. In the absence of congestion they are not affected by the availability of mass transit.

Lemma 2 (linear Congestion)**.** *When a solution satisfying $l < s$ to Eq.* (3) *exists, we have*

$$\begin{aligned}
\frac{dl^c}{ds} &= \tfrac{1+c_t}{2+(\gamma_p+\gamma_t)n(l^c)} > 0 \\
\frac{dl^c}{dw} &= \tfrac{c_w+\gamma_t \int_{l^c}^{r_f} n(r)dr}{2+(\gamma_p+w\gamma_t)n(l^c)} > 0 \\
\frac{dl^c}{dp} = -\frac{dl^c}{d\tau} &= \tfrac{1}{2+(\gamma_p+w\gamma_t)n(l^c)} > 0 \\
\frac{dl^c}{d\gamma_t} = \frac{-1}{w}\frac{dl^c}{d\gamma_p} &= \tfrac{\int_{l^c}^{r_f} n(r)dr}{2+(\gamma_p+w\gamma_t)n(l^c)} > 0,
\end{aligned}$$

A solution to (3) may not exist if $p - c + w\,c_w$ is high (as for l^c) or if congestion in mass transit, i.e. $w\,\gamma_t \int_l^{r_f} n(r)dr$, is much higher than congestion in the alternative mode, i.e. $\gamma_p \int_0^{r_f} n(r)dr$. The comparative statics in Lemma 2 confirms the intuition that l^c will vary so as to increase the market share of the less congested mode. With a uniform distribution of households, i.e. $n(r) = 1$, we have an explicit solution for the equilibrium

$$l^c = \frac{p - \tau + (1 + c_t)\, s + (c_w + r_f\, \gamma_t)\, w}{2 + \gamma_p + w\, \gamma_t}.$$

2.2 Optimal location of the station

In the previous section we studied the market share of mass transit as a function of its location s. The next step is to study the optimal location of the station. For this we need to introduce the construction and operation costs of the mass transit. Let θ denote the cost to provide a unit length of train line and 2ϕ denote the operating cost of one train.Once the location of the station s and the service frequency w are fixed, the total travel cost in the city (for all households) is the sum of travel cost in the two modes (congestion included) and the construction costs.

The total travel, construction and operation cost is:

$$\begin{aligned} TC = &\int_0^l r\, n(r)dr + \int_l^{r_f} (|r - s| + s\, c_t + w\, c_w)\, n(r)dr \\ &+ \gamma_p \cdot \left(\int_0^l n(r)dr \right)^2 + w\, \gamma_t \cdot \left(\int_l^{r_f} n(r)dr \right)^2 + \theta s + \frac{\phi}{w} \end{aligned} \tag{7}$$

The first line in (7) is the travel cost, net of congestion, in the two modes. The second line in the same equation reflects congestion costs in the two modes as well as construction and operating cost of mass transit. TC in (7) is minimized when the first-order conditions with respect to decision variables location s, service frequency w, transit fare p and private road toll τ are satisfied.[2]

Proposition 1. *The location of the station and the waiting time that minimize total transport cost TC satisfy the first-order conditions:*

$$\theta + (1 + c_t) \int_l^s n(r)dr = (1 - c_t) \int_s^{r_f} n(r)dr + H, \tag{8}$$

and

$$\frac{\phi}{w^2} = \left(c_w + \gamma_t \int_l^{r_f} n(r)dr \right) \left(\int_l^{r_f} n(r)dr - H \right) \tag{9}$$

where

$$H = \frac{(1 + c_t)n(l) \left[\left(w\, \gamma_t \int_l^{r_f} n(r)dr - \gamma_p \int_0^l n(r)dr \right) - (p - \tau) \right]}{2 + (\gamma_p + w\, \gamma_t)n(l)} \tag{10}$$

[2]The optimal s is necessarily in the interior of $(0, s)$.

Moreover, if transport prices p and τ are set to minimize transport cost TC, then $H = 0$. For $\gamma_t = \gamma_p = 0$, $ds/d\theta < 0$ and Eq. (8) has a unique solution if $n(r)$ is monotone increasing in r.

Notice that H in (8) has the same sign as $(\gamma_t \int_l^{r_f} n(r)dr - \gamma_p \int_0^l n(r)dr) - (p - \tau)$. This expression reflects the difference between congestion costs and access charges for the two modes considered. A positive value of H obtains when congestion in mass transit is under-priced, and vice versa. When prices are set to reduce total transport cost, H is set to zero. If prices are not optimal then the location and frequency of service can be used as controls to minimize total cost. From (8) an increase in the congestion in one mode reduces the market share of that mode.[3] Indeed, an increase in s has a number of impacts. First, the travel cost of passengers in (s, r_f) decreases (positive impact), and this is the left hand term in (8). Second, the travel cost of passengers living in (l, s) increases (negative impact), and this is right hand term in (8). Third, locally around s, the population belonging to the first group decreases and the population in the second group increases. Finally, locally around l, the population belonging to the second group increases and the population in the third group decreases. Condition (8) means that when all passengers select the lowest transport cost mode then, the last two impacts vanish. Equation (8) has always a solution in s. Indeed, at $s = 0$, the left hand side is positive and the right hand side is zero, and at $s = r_f$, the left hand side is zero and the right hand side is positive. So, the continuity of both expressions insures the existence of one solution at least between $s = 0$ and $s = r_f$. The first-order condition with respect to s can be presented in a different way. It is interesting to compare expression (9) and Mohring's square root Mohring (1972). Indeed, it is natural to have $w = 1/2n$ where n is the number of trains serving the station per unit of time. Then (9) is the sum of two terms. The first one relates the n^2 to the number of users (thus compatible with Mohring's rule) and the second one relates n^2 to the square of the number of users. This second term, with is obtained from congestion terms in mass transit is new feature of this model (cf. Mohring 1972, Kraus 1991).

Consider the case where there is no congestion and that waiting/cost time is zero. Define $N(r) = \int_0^{r_f} n(r)dr$. Then, for $\gamma_t = \gamma_p = 0$, (8) is equivalent to

$$(1 - c_t)N(s) = (1 + c_t)N\left(\frac{1 + c_t}{2}s\right) - (1 + c_t)N(s),$$

or

$$N(s) = \frac{1 + c_t}{2}N\left(\frac{1 + c_t}{2}s\right).$$

The same argument for the existence of a solution applies. The second-order condition, which ensures the existence of a unique solution that minimizes the total travel cost, simplifies to

$$n(s) \geq n(l)\left(\frac{1 + c_t}{2}\right)^2. \tag{11}$$

[3] Notice that the first term in the right hand side of (8) is decreasing in s.

Condition (11) is satisfied provided that the distribution of households $n(r)$ is not "too decreasing". Indeed, notice that the left hand side in (8) is monotone decreasing in s, while variation in the right hand side depends on the distribution $n(r)$. If the population density increases as we move further away from the city center than the latter is increasing and Equation (8) has necessarily a unique solution. This structure, is not compatible with the monocentric city model, however. Notice that by construction there are no external costs in this model and there is no need to charge the access to the train and the optimal toll is zero.

Totally differentiating (8) with respect to c_t and s and simplifying, yields (s^o is optimal location)

$$\frac{\mathbf{d}\, s^o}{\mathbf{d}\, c_t} = -\frac{1}{2} \frac{\int_{l^o}^{r_f} n(r)dr - l^o n(l^o)}{n(s^o) - \left(\frac{1+c_t}{2}\right)^2 n(l^o)}. \tag{12}$$

At s^o, where the second order condition (11) is satisfied, the denominator in (12) is positive. So, the derivative in (12) is negative if $\int_{l(s^*)}^{r_f} n(r)dr - l(s^*)n(l(s^*)) > 0$, which is the case, for example, for the uniform distribution.

Proposition 2. *The train service is welfare improving (decreases the social cost $T(s,\theta)$) if, and only if,*

$$(1 - c_t)\, N > \theta. \tag{13}$$

Notice that proposition 2 does not depend on the exact distribution of households, but only on their total number. It is independent from congestion terms γ_t and γ_p, too, since H in (10) vanishes as s gets close to zero. Condition (13) can be used as a first approximation to see whether it is interesting to invest in a new mass-transit line. As an illustration assume that the construction cost of an underground line is 38M€/km[4], that the relative cost between private transport and public transport is 1/4 (based on INSEE survey, 2005, and RATP prices). Suppose also that we focus on a horizon of 20 years with 200 days in a year. Then, according to (13) more than 13 000 traveler per day are required before investment is worthwhile.

When the distribution of households is uniform over $(0, r_f)$ we have an optimal location, denoted $s^o(\theta)$, given by

$$s^o(\theta) = s^o(0) - \frac{2\theta}{(1-c_t)(3+c_t)}, \tag{14}$$

so, as expected, $s^o(\theta) \leq s^o(0)$. If $\theta = 0$ then the optimal location of the station is given by

$$s^o(0) = r_f \frac{2}{3+c_t}. \tag{15}$$

[4]This is based on Madrid underground enlargement (1999-2008), source: Urban Public transport integration and funding.

So, the optimal location is a weighted average of 0 and r_f. If the relative cost of the train is not small ($c_t = 1$) it is optimal to locate the station halfway between border and CBD. As the relative cost of the train decreases the relative importance of r_f's weight increases and at the limit it moves to point $s^o = (2/3)r_f$. Notice that s^o is monotonic in c_t, so for all possible values of the relative speed of the train, the optimal location of the station varies in 16.7% of the segment $(0, r_f)$.

Under other distributions of households we still obtain comparable results. To be consistent with empirical observations let us consider a distribution that decreases in the distance from the city center. The negative exponential density, (cf. Stewart 1947, Clark 1951), has been proposed for this purpose. In our model, if we consider a distribution $n(r) = A\exp(-\lambda r)$, where A and λ are both positive constants, then we obtain (for $A = 5$ and $\lambda = 0.2$) an optimal location of the station at $0.361\,r_f$ for $c_t = 1$, and that increases to $0.492\,r_f$ when c_t decreases to 0. However, the exponential distribution does not provide a closed form solution,

For this purpose, on may consider a simpler linear distribution of the form $n(r) = A(1-r/r_f)$ where A is a positive constant. Similar calculations as above yield the optimal location at

$$s^o = \frac{2\left(3 + c_t - \sqrt{2(1+c_t)}\right)}{7 + c_t(4 + c_t)} r_f.$$

In this case (and for $A = 5$), s^o increases from $0.333\,r_f$ to $0.454\,r_f$ as c_t decreases from 1 to 0. So, in both cases the optimal location varies in a relatively small ranges as the relative travel speed covers all possible values.

2.3 Private operator

Most mass transit corporations are publicly controlled but with a budget constraint and in that case their optimal behavior is a mixture of welfare optimum and profit maximizing behavior (cf. Boiteux 1956). For this reason we are interested in the extreme case where the station is managed by a private operator who maximizes profit and imposes a charge p on each passenger transiting by the train.

The operator's profit is equal to total revenues minus the construction cost of the line and the operating cost of train,

$$p\int_l^{r_f} n(r)dr - \theta\, s - \frac{\phi}{w}. \tag{16}$$

The operator maximizes the profit in (16) over non-negative variables s, w and p. The charge should be set such that $l \leq s$, or

$$(1 - \Gamma(s) - c_t)s \geq p, \tag{17}$$

where $\Gamma(s) = \gamma_t \int_l^{r_f} n(r)dr - \gamma_p \int_0^l n(r)dr$. At the same time, the station should be located inside the city, i.e.

$$r_f \geq s. \tag{18}$$

Let (P) denote the program that consists in maximizing (16) subject to constraints (17) and (18) (and the non-negativity constraints on s, w and p). The Lagrangian corresponding to (P) is

$$L(s, w, p, \lambda_1, \lambda_2) = p \int_l^{r_f} n(r)dr - \theta\, s - \frac{\phi}{w} + \lambda_1[s - l] + \lambda_2(r_f - s). \tag{19}$$

An internal solution of this problem requires s, w, $p > 0$. Consider the first-order conditions. We have respectively $\partial L/\partial p = -pn(l)\,\partial l/\partial p + \int_l^{r_f} n(r)dr - \lambda_1 \partial l/\partial p$, $\partial L/\partial s = -p\,n(l)\partial l/\partial s + \lambda_1(1 - \partial l/\partial s) - \lambda_2$ and $\partial L/\partial w = -p\,n(l)\partial l/\partial w + \phi/w^2 - \lambda_1 \partial l/\partial w$. With these conditions, $\lambda_1 = \lambda_2 = 0$ implies that $\partial L/\partial s < 0$ and thus $s = 0$. Then $s > l$ and $s > 0$ imply that $\lambda_1 = 0$ and thus $\lambda < 0$, a contradiction. So $s = l$. Finally, $\lambda_2 > 0$ implies that $s = r_f$ but with $s = l$ this implies zero profit for the operator ($l = r_f$). The following proposition summarizes the behavior of the private operator.

Proposition 3. *A private operator who maximizes profit given in* (16) *subject to constraints* (17) *and* (18) *charges a price p^* so that*

$$l = s^*, \tag{20}$$

locates the station at s^*, which is a solution to

$$\frac{\theta}{1 - c_t} + \int_{s^*}^{r_f} n(r)dr = s^*\, n(s^*), \tag{21}$$

and chooses a service frequency w^* which is a solution to

$$\frac{\phi}{(w^*)^2} = \left(c_w + \gamma_t \int_{s^*}^{r_f} n(r)dr\right)\left(\int_{s^*}^{r_f} n(r)dr\right). \tag{22}$$

When households are uniformly distributed over $(0, r_f)$, we obtain $s^* = r_f/2$ and $p^* = (1 + \Gamma(r_f/2) - c_t)r_f/2$. There are two important consequences of Proposition 3. First, Equation (20) implies that the private operator does not serve the passengers located closer to the CBD than s. To see this point remember that users beyond s have all the same benefit from using mass transit and that this gain is higher than the gain for households living closer than s. Attracting passengers from the latter group requires a relatively important decrease in the mass transit fare. The net impact on the operator's revenue of lowering the fare and attracting these customers turns out to be negative, and the private operator prefers not to serve this group. In this case, the decrease in the transport costs generated by mass transit is fully transformed into profit for the private operator.

From (21), we see that the infrastructure cost of mass transit influences the location of the station only through the first term of the left hand side. If $\theta = 0$, then the station location is independent of c_t, and the private operator takes another decision than the one described above. Notice that the location decision of the private operator is the same as the socially optimal decision when the generalized cost of the train is relatively high. Indeed, a Taylor series development of optimality condition (8), around $c_t = 1$, yields

$$\left(\int_s^{r_f} n(r)dr - sn(s) \right) + O(c_t - 1)^2 = 0, \tag{23}$$

which is equivalent to (21) for c_t close to 1. So, the private operator's problem (for all c_t) is equivalent to the optimal welfare problem when the travel cost of the train is comparable to the driving travel cost.[5] For a higher speed of the train, the private operator locates the station too close to the CBD, i.e. $s^* < s^o$ for $0 < c_t < 1$.

It is interesting to see the extent to which the private operator adopts a pricing policy that gives a tight condition (17). In particular, if the location of the station is exogenously fixed as it may be the case in reality, and if it is located close enough to r_f he gets a small market share with that pricing practice. If s is given, the marginal user depends only on p. Let us write $r_l(p)$. We have the following result.

Proposition 4. *When the location of the station s is given exogenously, the optimal charge for the private operator is the solution to*

$$\int_l^{r_f} n(r)dr = \frac{1}{2} p\, n(l) \tag{24}$$

when the condition

$$\int_s^{r_f} n(r)dr < \frac{1}{2}(1 + \Gamma(s) - c_t + wc_w)s\, n(s) \tag{25}$$

is satisfied. Otherwise, we have $l = s$.

It follows that the private operator chooses a charge such that $l < s$ when s is sufficiently close to r_f, and (s, r_f) corresponds to a small part of the market.

A private monopolist running a mass transit system may appear as an extreme, unrealistic solution. Often a public monopolist is given a binding budget constraint and this solution will be a combination of of the private monopolist and the social welfare optimum solution.

3 Multiple stations

We extend the analysis to the case of n stations. This step allows us to introduce new features in the model. In particular, a new variable that reflects

[5] Thus, to compute the optimum location of the private operator it is easier to solve for the optimum and then fix c_t to 1.

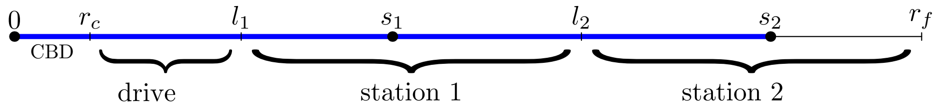

Figure 2: Two stations.

delays incurred by passengers in stops in intermediate stations before arrival is considered. This variable will affect the equilibrium and optimal quantities. A further issue to be considered with multiple stations is the construction cost of the stations. With a single station, this cost is included in the cost of the infrastructure (line construction), but with multiple stations we need to distinguish the two types of costs. We ignore congestion and waiting time in this section. The two stations are located at respective distances $s_1,\ldots,\ s_n$ from the city center. l_i is the location of the passenger who is indifferent between transiting through station i and station $i+1$.[6] Let g reflect a penalty (delay) incurred by a passenger from the stop in an intermediate station. Our focus will be limited, from here on, to the location of stations, and we ignore congestion terms as well as waiting time.

The user who is indifferent between stations i and $i-1$ is located at

$$l_i(s_{i-1}, s_i) = \frac{s_{i-1}}{2}(1-c_t) + \frac{s_i}{2}(1+c_t) + \frac{p_i - p_{i-1}}{2} + \frac{g}{2} \tag{26}$$

for $i = 2, \ldots, n$. A public manager chooses the location of the stations along the line $(0, r_f)$ in order to minimize total transport cost in the city. Without congestion, as in the case of one station, it is optimal to set the mass transit fare $p_i = 0$. The following result characterizes the optimal location solution.

Proposition 5. *For $\theta = g = 0$, the first-order condition for the location of station i, is*

$$(1-c_t)\int_{s_i}^{l_{i+1}} n(r)dr = (1+c_t)\int_{l_i}^{s_i} n(r)dr, \tag{27}$$

for $i = 1, \ldots, n-1$. For $i = n$, replace l_{i+1} in (27) by r_f.

Let us illustrate this result in the case of two stations. Figure 2 illustrates the notation adopted.

The conditions that determine the user indifferent at l_1 (using private transport or station 1) and l_2 (using station 1 or 2), respectively, are

$$(s_1 - l_1) + s_1 c_t + p_1 = l_1, \tag{28}$$

and

$$(s_2 - l_2) + s_2 c_t + g + p_2 = (l_2 - s_1) + s_1 c_t + p_1, \tag{29}$$

[6]Except l_1 who is indifferent between transiting by train or driving directly to the CBD.

From (28) and (29), we obtain

$$l_1(s_1) = \frac{s_1}{2}(1+c_t) + \frac{p_1}{2}, \quad \text{and} \tag{30}$$

$$l_2(s_1, s_2) = \frac{s_1}{2}(1-c_t) + \frac{s_2}{2}(1+c_t) + \frac{g+p_2-p_1}{2}. \tag{31}$$

Because we did set the cost per unit distance of driving equal to 1, the delay g is a cost relative to the driving cost. Similarly, p_1 and p_2 are also charges divided by the unit cost of private transport.

With a uniform distribution of households over line segment $(0, r_f)$, we can solve for the optimal locations of the stations[7][8]

$$s_1^o = \frac{2r_f}{5+c_t} + \frac{4g}{(1-c_t)(1+c_t)(5+c_t)} \tag{32}$$

$$s_2^o = \frac{4r_f}{5+c_t} + \frac{(3+c_t)g}{(1-c_t)(5+c_t)}, \tag{33}$$

After rearranging terms we see that the optimal locations have two components. The first one reflects the geometry of the city and depends on r_f. The second component is introduced by the discomfort imposed on users of station 2 at the stop in station 1. The discomfort penalty g moves the locations of the stations away from the city center. We can check that we have always $s_1 \geq l_1$, but that $s_2 \geq l_2$ only when

$$g \leq \frac{r_f(1-c_t^2)}{3+c_t},$$

which we assume to hold. The distance between the two stations is

$$s_2^o - s_1^o = \frac{2r_f}{5+c_t} + \frac{(c_t^2+4c_t-1)g}{(1-c_t)(1+c_t)(5+c_t)}. \tag{34}$$

So, an increase in g increases the distance between the two stations as long as $c_t^2 + 4c_t - 1 \geq 0$, or $c_t \leq \sqrt{5} - 2$. Otherwise an increase in g induces stations to be located closer to each other. Let us inspect the impact of the relative travel speed c_t on the optimal location of the stations. We have,

$$\frac{\partial s_1^o}{\partial c_t} = -\frac{2r_f}{(5+c_t)^2} + \frac{4(3c_t^2+10c_t-1)g}{[(1-c_t)(1+c_t)(5+c_t)]^2}, \quad \text{and} \tag{35}$$

$$\frac{\partial s_2^o}{\partial c_t} = -\frac{4r_f}{(5+c_t)^2} + \frac{(17+6c_t+c_t^2)g}{[(1-c_t)(5+c_t)]^2}. \tag{36}$$

Assume first that $g = 0$. An increase in c_t generates locations closer to the city center for the two stations. The impact is twice as high on the location

[7] With $s_1 < s_2$ we check that we have $l_1 < s_1$ and $l_2 < s_2$. When a mass transit fare is charged these constraints should be taken explicitly into account.

[8] Investment costs are not considered here

of station 2 than on the location of station 1. If $g > 0$, then a second impact appears. For station 2, this new impact goes in the opposite direction: to move the station away from the CBD. The reason is that moving station 2 further away reduces the number of passengers that will experience the extra delay g. For station 1, the new impact is in the same direction as the initial one when c_t is small (moves the station closer to the city center), and has the opposite impact for $c_t > 2\sqrt{7} - 5$.

Assume that $\theta > 0$ and let us denote optimal locations by $s_1^o(\theta)$ and $s_2^o(\theta)$, respectively for station 1 and station 2. Similar calculations as above yield[9]

$$s_1^o(\theta) = s_1^o(0) - \frac{r_f}{N}\,\theta\,\frac{2}{(1-c_t)(5+c_t)} \tag{37}$$

$$s_2^o(\theta) = s_2^o(0) - \frac{r_f}{N}\,\theta\,\frac{4}{(1-c_t)(5+c_t)}. \tag{38}$$

The distance between the two stations does not depend on g and decreases with c_t. From (37) and (38), we have $s_1^o(0) - s_1^o(\theta) = (s_2^o(0) - s_2^o(\theta))/2$. The optimal charge here is zero, because there is no externality. Although, the more passengers at station 1 the more stops are necessary for trains coming from station two.

3.1 Private operators

A private operator collects revenues from charges p_1 and p_2 in station 1 and in station 2. We discuss below, respectively, the case of monopoly and the case of duopoly.

3.1.1 Multi-stations Monopoly

A private operator is responsible for both location and pricing at the two stations. The profit function is

$$p_1 \int_{l_1}^{l_2} n(r)dr + p_2 \int_{l_2}^{r_f} n(r)dr. \tag{39}$$

Any solution must satisfy the ordering of the stations, i.e. $l_1 \le s_1$, $l_2 \ge s_1$, $l_2 \le s_2$ and $s_2 \le r_f$. In explicit form these constraints are respectively

$$(1-c_t)s_1 - p_1 \ge 0 \qquad (\text{or, } l_1 \le s_1) \tag{40}$$

$$(s_2 - s_1)(1-c_t) - (g + p_2 - p_1) \ge 0 \qquad (\text{or, } s_1 \le l_2) \tag{41}$$

$$(s_1 - s_2)(1+c_t) - (g + p_2 - p_1) \ge 0 \qquad (\text{or, } l_2 \le s_2) \tag{42}$$

$$s_2 - r_f \ge 0 \qquad (\text{or, } s_2 \le r_f). \tag{43}$$

Furthermore, all decision variables s_1, s_2, p_1 and p_2 are non-negative. The optimal solution for a monopolist is given in the following proposition.

[9] In the case of more than one station, it is more convenient to distinguish the cost of providing the lane and the cost of providing a station. This is useful to find the optimal number of the stations (to develop).

Proposition 6. *Assume that $g \leq (1-c_t)r_f/2$. Then, the locations chosen by the monopoly are a solution to the system of equations*

$$\int_{s_1^M}^{s_2^M} n(r)dr = s_1^M n(s_1^M) \tag{44}$$

$$\int_{s_2^M}^{r_f} n(r)dr = \left(s_2^M - s_1^M - \frac{g}{1-c_t}\right), \tag{45}$$

where the superscript M refers to monopoly values. The optimal charges are given by,

$$p_1^M = (1-c_t)s_1^M \tag{46}$$

$$p_2^M = (1-c_t)s_2^M - g. \tag{47}$$

Thus, the solution of the Monopoly with two stations is a direct extension of the case of one station managed by a private operator. Notice that with $g = 0$, the location of the two stations does not depend on the relative speed of the train. From (45) we notice the substitution effect between g and c_t. Indeed, an increase in g may be compensated for by a decrease in c_t, i.e. a decrease in the relative cost of the train.

If users are uniformly distributed over $(0, r_f)$, and if $g \leq (1-c_t)r_f/2$, the two stations are located at

$$s_1^M = \frac{1}{3}\left(r_f + \frac{g}{1-c_t}\right) \tag{48}$$

$$s_2^M = \frac{2}{3}\left(r_f + \frac{g}{1-c_t}\right), \tag{49}$$

and the charges are given by

$$p_1^M = (1-c_t)\, s_1^M \tag{50}$$

$$p_2^M = (1-c_t)\, s_2^M - g. \tag{51}$$

In order to compare with alternative distributions of households we consider the output of the linear distribution, $n(r) = A(1 - r/r_f)$, and the negative exponential distribution $n(r) = A\exp(-\lambda r)$, where in both cases, A and λ are positive numbers. The optimal locations for each case are given in Table 1. To find the optimal location of the monopoly, recall that the problem of the private operator is equivalent to the optimal problem when driving speed is similar to the train speed. So, $c_t = 1$ in Table 1 is also the optimal value for a public monopolist.

3.1.2 Duopoly

Assume that each station is managed by a private operator who charges the users of his station. Assume that the operation cost is zero and that passengers

	$n(r) = A\left(1 - \frac{r}{r_f}\right)$		$n(r) = Ae^{-\lambda r}$	
	s_1	s_2	s_1	s_2
$c_t = 0$	2.68	5.98	2.81	6.73
$c_t = 1$	1.18	4.79	2.27	5.32

$A = 5$; $\lambda = 0.2$; $r_f = 10$

Table 1: Numerical examples.

are distributed uniformly over $(0, r_f)$. It is natural to consider a two-stage game where, at stage one each firm chooses the location of its station, and at stage two both firms observe the choice made at stage one and compete in prices.

Consider the decision of firm 1 at stage 2. The firm observes s_1 and s_2 and has to choose the best reply to the charge of firm 2, p_2. When firm 2 chooses the charge such that $l_2 = s_2$, firm 1 may be seen as the private operator in the model of one station, but in a city of radius s_2 instead of r_f. Thus, firm 1 sets its charge so that $l_1 = s_1$. Similar reasoning applies to firm 2, and it is likely that at the equilibrium both firms choose their charges so that $l_1 = s_1$ and $l_2 = s_2$.

Then consider the location decisions at stage 1. Assume that the city has a unit length and that $g = 0$. We have seen above that a monopoly will locate the two stations at $s_1 = 1/3$ and $s_2 = 2/3$, respectively. Let us check if this is an equilibrium for the duopoly. When firm 2 chooses the location at 2/3, firm 1 is faced with a city that extends from 0 to 2/3, and we know from the one station model that it is optimal for a private operator to locate in the middle of the city, i.e. at 1/3 in this case. Similarly, when firm 1 locates at 1/3, firm 2 is facing a location problem in a city that extends from 1/3 to 2/3. From the same argument, its optimal decision implies a location at 2/3. Thus, $s_1 = 1/3$ and $s_2 = 2/3$ is an equilibrium choice for the duopoly in this case. Pricing and location decisions are the same under duopoly and monopoly in this model. That a duopoly behaves like a monopoly is the result of the separation of the two markets. The result extends to the case where the delay penalty is positive, i.e. $g > 0$.

If entry is sequential, then the first firm will locate at the middle of the city (for simplicity, we assume the uniform distribution), i.e. $s_1 = 1/2$. If the entrant firm has to choose any location, the solution would be to locate at $s_2 = s_1 + \varepsilon$, where ε is an arbitrarily small positive number. This allows firm 2 to collect (almost) all the gain of firm 1, which gets a zero profit now. This may induce firm 1 to reconsider its initial decision. An existence problem arises in this case. It is better to restrict the strategy space of the entrant in order to avoid such situation. For example, if firm 2 is restricted to choose between 0 and s_1, then would locate at 1/4. An entrant restricted to choose location between $\underline{r}$ and r_f, with $s_1 < \underline{r}$ will choose $s_1 = \underline{r}$. We summarize our conclusion in this

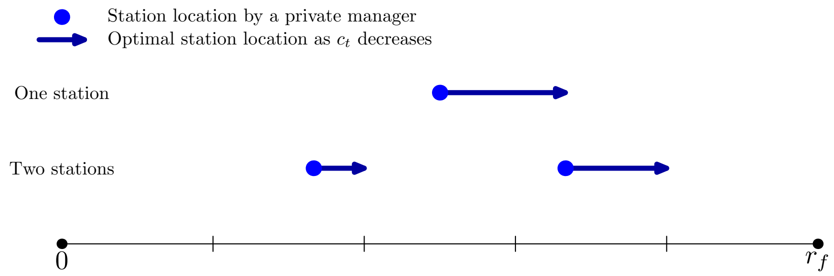

Figure 3: Comparison between one and two stations ($g = 0$).

proposition.

Proposition 7. *Under the framework discussed above, the duopoly problem is identical to monopoly problem.*

3.2 Comparison between one and two stations

We compare the case of one station with the case of two stations, assuming a uniform distribution of households. From (15) and (33), we have (for $g = 0$)

$$s_2^o - s^o \geq 0 \quad \text{and} \quad s^o - s_1^o \geq 0, \tag{52}$$

and

$$\frac{s_2^o - s^o}{s^o - s_1^o} = \frac{1}{2}(1 + c_t). \tag{53}$$

Let us call "initial station" the optimal location of the station in the first model (one station). Condition (52) says that station 1 will be located between the initial station and the city center, and station 2 will be located beyond the initial station. Condition (53) says that station 2 will be closer to the initial station than station 1.

Figure 3 illustrates the difference between one and two stations for $g = 0$. The dot indicates the location of the station chosen by a private operator, while the white lines correspond to the optimal location (but would never build a rail line). The latter depends on the relative cost of the train. For $c_t = 1$ the private operator chooses the optimal locations. As c_t decreases (the relative cost of the train is lower), the optimal location moves away from the city center, while the private operator's remain unchanged. Under both management regimes, the impact of g is a translation further away from the city center. The general case is not particularly different and the same result hold.

3.3 Discomfort in mass transit

We now give the formulation of the problem when discomfort is taken into account. This step makes the model less tractable and is mainly intended to introduce the numerical experiments we undertake in the next section. Total cost for each mass transit passenger has four parts.

1. Home-to-station trip : $|r - s_i|$. Taking into account the distribution of the population and summing for all stations $i = 1, \ldots, n$ we have

$$\sum_{i=1}^{n} \int_{l_i}^{l_{i+1}} |r - s_i| n(r) dr,$$

   where $l_{n+1} = r_f$.

2. The trip from station s_i to station s_{i-1} in the train implies two costs: time-value c_t and discomfort reflected by a function $f(\cdot)$. Summing on all trips, we get

$$\sum_{i=1}^{n} N(l_i)(s_i - s_{i-1})[c_t + f(N(l_i))],$$

   where $N(x) = \int_x^{r_f} n(r)dr$, i.e. the population living beyond x. The function $f(\cdot)$ may reflect a linear discomfort, as we have done in the case of one station, but we keep it in this more general and flexible form for instance. Below we propose a new and more convenient formula to mass transit.

3. All those who use mass transit incur a waiting time in the station,

$$N(l_1)\, w\, c_w.$$

4. Under each stop those who are in the train incur a delay cost equal to g. The total cost in the city is:

$$\sum_{i=1}^{n} [N(l_{i+1}) - N(l_i)]\, (i-1)\, g.$$

Total cost in mass transit in the city is the sum of these four parts. The computation of this quantity requires, as a first step, the identification of market segmentation. This provides for each station the location of all passengers which use it. Thus, we have to find all those locations where a passenger is indifferent between two adjacent stations. This leads to the following equation

$$p_{i-1} + l_i - s_{i-1} + s_{i-1}\, c_t = p_i + g + s_i - l_i + N(l_i) \cdot f(N(l_i))(s_i - s_{i-1}), \quad (54)$$

which have to be solved for $i = 2, \ldots, n$ (with $s_0 = 0$). For $i = 1$ the user who is indifferent between using his car or mass transit is located at l_1 that solves

$$\tau + l_1 + \gamma_p(N(l_1) - \overline{N}) = p_1 + s_1\, c_t + w\, cw + N(l_1)\, f(N(l_1))\, s_1. \quad (55)$$

4 Numerical illustration

With more than two stations it becomes difficult to derive analytical results and it becomes interesting to use a numerical approach. When a numerical approach is used it is also interesting to use a more general formulation of the discomfort in mass transit. We start this section with a discussion of this feature and next discuss the numerical results.

4.1 Non-linear discomfort formula

The mass transit vehicle has n_1 seats and its loading capacity allows $n_2 - n_1$ other passengers to be standing under non-crowded conditions. Let c_0 denote the time cost for a passenger having a seat and c_1 denote the time cost for a standing passenger (but non crowded situation). When there are more than n_2 passengers in the vehicle, the discomfort increases strongly with every new passenger.

The following congestion function reflects this situation :

$$C(n) = c_0 + \frac{c_1 - c_0}{1 + e^{\alpha(n_1 - n)}} + \beta e^{\gamma(n - n_2)}, \tag{56}$$

where α, β and γ are parameters used to calibrate the congestion or discomfort function. A large value of α(more than ten) provides satisfactory results, but for β and γ convenient values have to be discussed on the basis of empirical evidence.

Equation (56) is illustrated on Figure 4. As long as $n \leq n_1$ the second and third terms are small, and as soon as the number of passengers becomes larger the denominator of the second term gets close to 1 and (since the third term remains small) the cost (for new passengers) is almost equal to c_1. As more passengers arrive and as their number gets higher than n_2 the vehicle gets more and more crowded and time cost increases substantially (the third in (56) term becomes important).

Notice that the above formula gives the travel cost for single user and for all passengers. This is because passengers who have a seat will not have an increase in their cost as new passengers arrive. Total travel cost in the vehicle is:

$$TC(n) = \begin{cases} n\, c_0 & \text{if } n \leq n_1 \\ n_1\, c_0 + (n - n_1)C(n) & \text{if } n > n1. \end{cases} \tag{57}$$

Consequently the marginal cost of a new passenger is:

$$TC(n) = \begin{cases} c_0 & \text{if } n \leq n_1 \\ C(n) + (n - n_1)C'(n) & \text{if } n > n1. \end{cases} \tag{58}$$

4.2 Summary table

Except for N and r_f all other parameter values are expressed as relative values with respect to c_p. Base-case parameter values are given in Table 2. Total

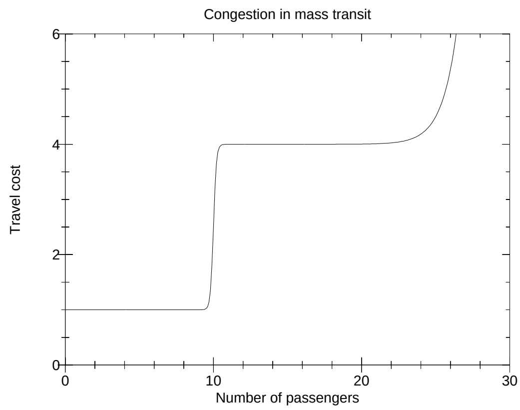

Figure 4: Discomfort in mass transit.

population N and the radius of the city are set to the same value so that the average distribution of households per unit of distance is equal to one. The travel cost in mass transit, net of discomfort, is set to one third travel cost in private vehicle, net of congestion cost. The waiting cost in the station is about one half of the travel cost. The implicit assumption here is that travelers are better off being in the station than being in the vehicle. This is not crucial but seems convenient as passengers have more flexibility in the station (the possibilities of doing some shopping, or changing the travel decision, ...). Construction cost θ (normalized to a daily cost) are smaller than travel cost. This is also the case for the service cost ϕ, but the sum of these two costs is higher than one. The delay cost g resulting from intermediate stops of the train is equal one tenth the travel cost. Concerning congestion and discomfort cost γ_p, γ_t and the parameter in the non-linear discomfort function, there aren't obvious values to propose initially. We have tried many values and considered those which provide the most coherent results. Empirical investigation of this question remains important. n_1 and n_2 are set so that only a small part of the population finds a seat but most of the population travels under non-crowded conditions. Of course, alternative analysis focusing on peaked hours characterized by crowding in mass transit have to be considered in the future.

Summary output is given in Table 3. There are three situation considered : non congestion, linear congestion and non-linear congestion. To obtain the first case simply set $\gamma_p = \gamma_t = 0$ in the analytical model above. The linear case corresponds to positive values of these two parameters and the non-linear congestion corresponds to the formulation introduced in Section 3.3.

For each case we provide output indicators for the optimum decision, which minimizes total transport cost in the city, and values related to the monopoly decision. There are three set of indicators provided. First the optimum number of stations, then transport costs in the city and monopoly profit then the locations of the stations, the service frequency, the pricing at each station and

Variable	Value
r_f	10.0
N	10.0
c_t	0.33
c_w	0.15
θ	0.4
ϕ	0.7
g_t	0.001
g_p	0.08
g	0.1

Table 2: Parameter values in the base case.

road toll level. Market segmentation is given only for the optimum since for monopoly the corresponding constraint is always binding and the stations' location is sufficient.

There are some visible facts one can observe from Table 3. In all cases the monopoly chooses more stations than in the optimum case. As a consequence, the transport cost induced by a monopoly is always substantially higher than the optimal transport cost. Continuing the comparison between the monopoly and the social optimum, notice that the private operator always chooses more stations (about twice). The intuition for this outcome is that doing so the monopoly is able to discriminate more between mass transit users and thus collects higher profits.

Service frequency is better under the optimal regime unless for the case of non-linear congestion. In this case, it is optimal to locate the first station far away from the city center. The last part of the train trip, between the first station and the CBD, is characterized by an important congestion. To reduce this term, the monopoly increases the train service instead of moving the stations away from the CBD. The function played by the pricing is not the same under the two regimes. The private operator chooses its prices in order to increase the profit, while at the optimum pricing of mass transit induces correct segmentation of the market, i.e. each station is used by a convenient number of passengers. Despite this divergence we notice that the pricing of mass transit increases with the distance between the station and the city center. The increase is much higher under monopoly, however.

Acknowledgment

Moez Kilani acknowledges financial support from Université Charles de Gaulle - Lille 3 (Project BQR) and from Predit (Project: Tarification des transports individuels et collectifs à Paris - dynamique de l'acceptabilité-).

No congestion					
n^o	4	TC	25.89	s	$(1.81, 3.63, 5.45, 7.30)$
				l	$(1.51, 2.79, 4.60, 6.50)$
				w	0.16
				p	$(0.77, 0.83, 0.90, 0.97)$
				τ	0.70
n^m	7	PRF	17.72	s	$(0.71, 1.59, 2.49, 3.59, 4.91, 5.71, 6.74, 7.71)$
		TC	31.21	w	0.5
				p	$(0.48, 1.07, 1.67, 2.41, 3.29, 3.82, 4.52, 5.17)$
Linear congestion					
n^o	4	TC	26.13	s	$(1.52, 3.40, 5.30, 7.23)$
				l	
				w	0.15
				p	$(0.71, 0.78, 0.85, 0.91)$
				τ	0.75
n^m	7	PRF	18.78	s	$(1.51, 2.38, 3.28, 4.23, 5.23, 6.28, 7.39)$
		TC	31.84	w	0.70
				p	$(1.18, 1.76, 2.36, 3.00, 3.66, 4.37, 5.10)$
Non-linear congestion					
n^o	4	TC	23.05	s	$(4.20, 5.77, 7.02, 8.00)$
				l	$(2.23, 5.33, 6.69, 7.72)$
				w	0.087
				p	$(0.94, 1.86, 2.84, 3.90)$
				τ	2.36
n^m	8	PRF	13.16	s	$(1.07, 2.00, 2.92, 3.86, 4.84, 5.86, 7.08, 7.79)$
		TC	33.43	w	0.05
				p	$(0.61, 1.15, 1.69, 2.27, 2.85, 3.48, 4.26, 4.72)$

Table 3: Optimal and equilibrium values for transport sector.

5 Conclusion

This paper discussed the endogenous location of mass transit stations. The analysis has been conducted under both public and private management regimes, and with respect to various formulations of discomfort in mass transit. In this context we have proposed a new formulation of a non linear discomfort function that takes into account the number of seats available and the possible crowding that may occur in the vehicles.

The literature on mass transit is relatively poor by comparison to road transport, in particular with respect to congestion and discomfort. Our analysis extends earlier models by Mohring (1972) and Kraus (1991).

The first part of the paper has focused on an analytical model of station's location, which we have used to derive a number of results with respect to decisions of a private manager by comparison to optimal decisions (which minimizes total transport cost in the city). The relative cost of the two transport modes plays an import role in the solution of the public manager. Indeed, the optimal location in this case depends on the relative cost. As mass transit becomes more attractive, the public manager locates the station further away from the city center. The private manager, by contrast, sets the price and locates the station independently of the relative cost. In this case all the surplus, generated by mass transit, is collected by the private manager and all users are set to their willingness to pay.

The model incorporates congestion in the two transport modes. Congestion in mass transit has been considered by a limited number of authors (cf. Kraus 1991). With an arbitrarily distribution of households, users split between the two modes cannot be derived explicitly, but it remains possible to conduct sensitivity analysis with respect to key parameters of the model. We find that, as a transport mode becomes more congested, it is optimal to reduce its market share by relocating the mass transit station.

With multiple stations the new feature introduced in the model is the delay incurred at the intermediate stops of the train before it reaches the destination. In both regimes, this delay induces a location of the station away from the city center. With many stations the model is less tractable and we have omitted the congestion impacts, unless in the numerical part, where a non-linear discomfort function has been considered. The adequate calibration of this function with respect to empirical data has to be undertaken in the future, but the first analysis conducted here shows a good explanatory potential of this formulation.

References

Boiteux, M. (1956), 'Sur la gestion des monopoles publics astreints à l'équilibre budgétaire', *Econometrica* **24**, 22–40.

Clark, C. (1951), 'Urban population densities', *Journal of the Royal Statistical Society* **114**(Series A), 490–496.

De Lara, M., de Palma, A., Kilani, M. & Piperno, S. (2008), Congestion pricing and long term urban form: application to île-de-france, Cahiers du Laboratoire d'Econometrie 2008-14, Ecole Polytechnique.

Kraus, M. (1991), 'Discomfort externalities and marginal cost transit fares', *Journal of Urban Economics* **29**(2), 249–259.
URL: *http://ideas.repec.org/a/eee/juecon/v29y1991i2p249-259.html*

Mackie, P. J., Wardman, M., Fowles, A. S., Whelan, G., Nellthorp, J. & Bates, J. (2003), Value of travel time savings in uk, Report to the transport department, Institute of Transport Studies.

Mohring, H. (1972), 'Optimization and scale economies in urban bus transportation', *American Economic Review* **62**(4), 591–604.
URL: *http://ideas.repec.org/a/aea/aecrev/v62y1972i4p591-604.html*

Stewart, J. (1947), 'Suggested principles of social physics', *Science* pp. 179–180.

Wardman, M. (2001), Public transport values of time, Working paper 564, Institute of Transport Studies.

A Dynamic departure time

This section extends the analysis with respect to waiting time by distinguishing between early and late arrival costs. We consider early delay and arrival time. To make the analysis at this stage simple, we assume that all the households have a uniform arrival time preferences over a closed interval between date $t = 0$ and date $t = T$, with $T > 0$. We follow the related literature by denoting Early delay cost by β and late delay cost by γ, with $\gamma > \beta$. Let s denote the time when service is given $s \in (0, T)$, i.e. the time when the train leaves the station. When there are two trains, the user who is indifferent between both has a preferred arrival time that we denote by $\hat{t}$. When needed, s and $\hat{t}$ will have indices.

If two services are provided at dates s_1 and s_2 respectively, the indifferent user will have a preferred arrival time $\hat{t}$ that satisfies

$$(\hat{t} - s_1)\gamma = (s_2 - \hat{t}\,)\beta \tag{59}$$

or,

$$\hat{t} = \frac{\gamma}{\gamma + \beta}\, s_1 + \frac{\beta}{\gamma + \beta}\, s_2. \tag{60}$$

Equation (59) reflects a situation when users are not aware of congestion. If the latter is taken into account, and if it take a linear form, then[10]

$$(\hat{t}^c - s_1) + \delta\hat{t}^c\frac{N}{T} = (s_2 - \hat{t}^c)\beta + \delta(T - \hat{t}^c)\frac{N}{T} \tag{61}$$

[10] As it becomes clear from the analysis below, the exact formulation of congestion not crucial as long as the preferred arrival time of households is uniformly distributed.

or,

$$\hat{t}^c = \frac{\gamma s_1 + \beta s_2 + \delta N}{\gamma + \beta + 2\delta N/T}. \tag{62}$$

Thus $\hat{t}^c - \hat{t}$ have the same sign as $\gamma + \beta - 2(\gamma s_1 + \beta s_2)$, or in other words $\hat{t}^c > t$ if and only if

$$\frac{T}{2} > \frac{\gamma s_1 + \beta s_2}{\gamma + \beta}. \tag{63}$$

For example, if $s_1 = 0$, $s_2 = T$ and $\gamma > \beta$ then $\hat{t}^c > \hat{t}$. The intuition for this result is straightforward: Higher congestion cost induces a smaller importance of delay costs. The following result describes the behavior of a public administrator who wishes to minimize the transport cost in mass transit.

Proposition 8. *Whether congestion is taken into account or not, the optimal timing of the service when there are n trains is*

$$s_i^o = \left(i - \frac{\gamma}{\gamma + \beta}\right)\frac{T}{n} \tag{64}$$

for $i = 1, \ldots, n$.

A consequence of this proposition is that the problem with n services in the time interval $(0, T)$ is equivalent to the problem of one service in the time interval $(0, T/n)$ repeated n times. Indeed, we have the following conclusion.

Corollary 1. *If train services are provided according to* (64)*, then the marginal users have respective preferred arrival times at*

$$\hat{t}_i = i\frac{T}{N} \tag{65}$$

for $i = 1, \ldots, n$.

Consequently, $\hat{t}_i^c = \hat{t}_i$, i.e. the allocation that takes into account congestion leads to the same allocation that does not take into account congestion.

With an alternative transportation mode available only a part of the population uses the train. To simplify, assume that users have a fixed reservation cost, denoted η. When P users are uniformly distributed, according to their preferred arrival time, over $(0, T)$, and if there are n trains during the same time interval, then the number of households using the train is given by the following demand function

$$D(n) = \begin{cases} \eta P \frac{n}{T} \frac{\beta+\gamma}{\beta\gamma} & \text{if } \eta \leq \frac{\beta\gamma}{\beta+\gamma}\frac{T}{n} \\ P & \text{if not.} \end{cases} \tag{66}$$

To determine the optimal number of services, assume that the cost of one service is ξ. The total cost of n services can be written in the form

$$\varphi(n) = \frac{T}{n}\Gamma + \xi\, n, \tag{67}$$

where Γ is a constant that depends on β and γ. It is clear that total travel cost in the mass transit is minimized with n^f services where, n^f is given by

$$n^f = \sqrt{T \frac{\Gamma}{\xi}}. \tag{68}$$

The expression of the optimum number of services corresponds to the well known Mohring's square root rule (cf. Mohring 1972, Kraus 1991) adapted to our model. If congestion is taken into account, then (68) remains true but with a higher value of Γ, inducing a larger number of services at the optimum.