PUBLIC TRANSPORT CONGESTION COSTS : THE CASE OF THE PARIS SUBWAY[1]

Rémy Prud'homme[2], Martin Koning[3], Luc Lenormand[3], Anne Fehr[3]

November 16, 2011

Abstract

The paper argues that congestion in public transportation is alike the much more studied road congestion. It sets out to produce, on the case of the Paris subway, a congestion cost curve linking the willingness to pay for non-congested travel to actual congestion levels. Congestion costs appear high. What travelers would be ready to pay to avoid congestion in the Paris subway is on average about three times the amount of their out-of-pocket payments. An 8% increase in densities experienced over the 2002-2007 period implies a welfare loss of at least 75 M€ per year. Taking into account the subway congestion costs and the resulting externalities modifies significantly the optimal car-subway mix (from 8% to 24% according to a tentative computable model).

I – Introduction

Public transport congestion is very much akin to road transport congestion. On a given road, an additional vehicle generally slows down the speed at which all other vehicles are being driven, causing a time loss to all other drivers. This time loss, often called congestion cost, is an externality leading to a suboptimal usage of the road, that can in principle be corrected by means of a congestion tax. On a given train or subway, an additional passenger will generally increase the crowding or density, causing a comfort loss to all other passengers. This comfort loss, which could equally be called a congestion cost, is also an externality leading to a suboptimal usage of the train or subway, that could in principle be corrected by means of a congestion tax. The only

[1] This piece of research did not benefit from any grant or research contract. However, it benefited from the assistance of RATP, the Agency in charge of the Paris metro ; the authors are pleased to express their gratitude to RATP. They also wish to thank Guillaume Faburel and Pierre Kopp for useful advice. We are also grateful for the very useful comments offered by two anonymous referees on a first draft. This paper has been accepted for publication in *Transport Policy*.
[2] Professor emeritus, University Paris XII
[3] PhD students, University Paris I

difference between these two types of congestion is that the time loss of the first is replaced by the comfort loss of the second.

Yet, there is a striking contrast between the abundance of studies, analysis, estimates, and policies or policy proposals, on the topic of road congestion costs and policies, on the one hand (see for instance a recent survey by Tsekeris and Vos (2009) that examines no less than 429 papers), and the dearth of similar studies on public transport congestion costs and policies on the other hand.

The topic is not new, and there is a sizable - and rapidly growing - body of literature on the causes and consequences of crowding in public transport (Litman 2008; Armelius and Hultkranz 2006; Wardman and Whelan 2010). However, the focus has been on a better understanding of individual transport choices and demand - a legitimate concern - rather than on the public policy implications of public transport congestion - an equally legitimate concern. Results are aimed at transport companies (many studies have been undertaken by consultancies for such companies, particularly rail companies, mostly in the UK, and remained unpublished) much more than at policy makers. These studies mobilize psychology or sociology or transport engineering concepts and concerns rather than economics notions. Many are dichotomic, exploring the implications of seated trips versus non-seated trips (Leurent & Liu 2009). Few talk euros (or dollars or pounds). This line of research should certainly not be abandoned, much to the contrary, but could usefully be complemented by more policy-oriented economic or public finance approaches.

To take an example, practically all cost-benefit analyses of urban tolls ignore the increased congestion in public transport caused by users shifting from car to subway. On the London congestion charge, this issue is not even mentioned in Transport for London reports, nor in Prud'homme and Bocarejo (2005). On the Stockholm trial, it is not considered as a cost in Transek (2006), and it is only discussed briefly in Kopp and Prud'homme (2010). The idea of internalizing public transport congestion externalities by means of subway tolls, to take another example, has very rarely been mentioned (Krauss 1991).

This paper tries to produce a congestion cost curve for public transport, and to utilize it to throw some light on transport policies. Public transport congestion here relates to subway and rail, not to bus transportation. This is because the supply of rail and subway services is generally rigid. That of bus services, by contrast, can more easily be increased if and when demand increases. It has even been argued, in particular by Mohring (1972) that an increase in demand, when met by more frequent buses, will reduce bus waiting times, thereby improving, not reducing, the quality of bus services.

The paper is based on the case of the Paris subway, on the basis of a passenger survey conducted in June 2009. A few words on Paris and the Paris subway may therefore be appropriate. Paris is an agglomeration of about 11 million people, consisting of a dense core of about 2 million people, the municipality of Paris, surrounded by suburbs. The Paris subway benefits mostly the Paris municipality (81% of the lines length, 86% of the stations), although several lines extend to the adjacent suburban areas (19% of the lines length, and 14% of the stations). The subway accounts for about 48% of public transportation in the Paris *agglomeration*, much more than buses (19%) or trains (33%) in terms of passenger*km. It accounts for a much larger share (around 80%) in the Paris *municipality*.

Table 1 – Recent changes in the Supply and Demand of the Paris Subway, 2002-2007

	2002	2007	% change
Supply (G seats*km offered)	25.4	26.4	+4%
Demand (G. passenger*km)	6.2	6.7	+13%
Demand/supply	0.243	0.264	+8%

Source : RATP. *Les Statistiques annuelles*, 2003 and 2007 editions.
Note : « seats offered », an indicator of supply, is calculated on the basis of seats, plus 4 people per square meter in standing areas. The demand/supply ratio suggests that only one « seat » (as defined above) out of four, is occupied on average. What matters here is the change in this ratio.

The Paris subway is increasingly congested, as shown in Table 1. The supply of subway increases very slowly, because it is extremely difficult to create new lines and even to add trains, particularly in the Paris municipality, for costs and planning reasons. The demand increases, at a rate of 2.5% per year in the past decade (as opposed to the demand for buses, which declined over the same period). As a result, congestion, or more precisely the congestion-causing demand/supply ratio, is increasing, at a rate of 1.6% per year. This is an average. The situation is better on certain lines, and worst, or much worst, on other lines, such as line 13, and complaints are on the increase.

The remainder of this paper presents the survey undertaken in the Paris subway (section II), an attempt to use the outcome of this survey to produce a subway congestion cost curve (section III), some of the policy implications of these findings (section IV), and concludes (section V).

II – Survey undertaken

Nearly 700 metro users were surveyed in June 2009. One of the most traveled subway line, line n° 1, was selected. Line 1, going from Porte de Vincennes, east of Paris, to La Défense, the new business center west of Paris, is reported to carry 213 million people per year (in 2008). Usage increased significantly in recent years (by 25% between 2000 and 2007), and serious congestion episodes are reported. The variance of congestion situations can be expected to be relatively large. Metro users were interviewed between 7h45 and 10h15 at five different subway stations. Passengers were chosen randomly: two interviewers asked the first two persons entering the subway platform after each train departure. A third member of the survey team was at the same time counting the number of passengers in each train compartment, thus producing an objective indicator of density or congestion[4].

A difficulty was that most of our observations were collected at peak hour, when the frequency of trains is very high. The time between two trains is 1 minute 45 seconds. The questionnaire had to be short. A first question was : *"Did you experience this subway line at off-peak, non-congested time ?"*. People answering "no" were no longer questioned. We wanted to be sure that all our respondents had first-hand experience of a non-congested situation, and were able to compare the present more or less congestion situation with this benchmark, non-congested one.

The remainder of the questionnaire included about ten questions on the characteristics of interviewees (age, sex, income, place of residence), on the trip undertaken, on the perception of congestion, and on the valuation of congestion-caused loss of comfort. On this key issue, we did not ask a direct question (how much do you value comfort losses?) which would not produce meaningful results in the French context, but two indirect questions.

[4] RATP, the subway company, does it as a matter of routine, and has a team of « counters », who very kindly taught us how to do it.

One was about trading time versus comfort: *"To benefit from off-peak levels of congestion on this line, would you be ready to take a subway that would spend an additional n minutes?".* In case, the answer with n = 5 minutes was "yes", the question was repeated with n = 10 minutes, then 15 minutes, then 20 minutes.

This question aims at reducing two of the biases that plague contingent analysis (Haab & McConnel 2003). When asked about the amount of money they are willing to pay to obtain a particular good, or avoid a particular bad, respondents may be tempted to manipulate their answers to make others pay for them (this is called a strategic bias). This temptation is reduced by questions not asked in money terms. Then, respondents often have difficulty understanding and visualizing the situation offered to them on which they are asked a question (this is called an hypothetical bias). In the Paris subway at peak times, it is rather frequent to see passengers that let pass a very congested train and wait for the next train in the hope it will be less congested. A question on trading time versus comfort therefore makes reference to a very concrete experience familiar to all respondents.

The other question was about trading money versus comfort: *"If there were in this subway first class carriages ensuring off-peak congestion levels and costing 10% more [then 50%; then 100% more], would you take them?"*. Such a system existed in the Paris subway until 1982. We recognize that this question does not reduce the possibility of a strategic bias as much as the previous one; and also that answers might be polluted by social attitudes and not reflect a "pure" WTP for non congestion.

Table 2 – Descriptive Statistics

	Average	σ	%
Cardinal statistics :			
Age (years)	36.6	9.9	
Income (€/month)	2000-2500	870	
Congestion[a] (index 0-5)	3.4	1.1	
Density (persons/m2)	2.1	0.7	
Length of subway trip (minutes)	27	14	
Binary statistics :			
Sex : male			49
Residence : Paris municipality			57
Trip purpose : work			86
Car availability			32
Comfort v. time:			
No with n = 5 min.			29
Yes with n = 5-10 min.			39
Yes with n = 10-15 min			27
Yes with n = 15-20 min			3
Yes with n = 20 min			1
Weighted average	7,9	4,4	
First class v. money:			
No, whatever price increase			71
Yes , with price increase of 10%			8
Yes, with price increase of 10-50%			15
Yes, with price increase of 50-100%			7

Notes : [a]Congestion is the subjective assessment of congestion by respondents, on a scale of 0 (no congestion) to 5 (very high congestion). σ = standard-error. Total number of usable questionnaires : 585.

The main findings of the survey are presented in Table 2. Respondents are about equally divided between male and female. Their average age is 37. Their average monthly income is in the 2000-2500 euros range (this refers in principle to respondent's income, not to his family income) . A majority (57%) of respondents live in the Paris municipality. The average length of trip is about half an hour. Only a third of the respondents had access to a car for their trip, which is much less than the Paris agglomeration average. The sample is heavily biased towards work trips (86% of respondents), which is explained by the choice of the line and of the survey time. Subway users are very much aware of the existence of congestion. On a scale of 0-5, with 0 = no congestion, the average ranking of congestion is 3.4, with a standard-error of 1.1. This subjective ranking is loosely correlated with the objective ranking of congestion, defined as the measured (or rather estimated) number of persons per square meter[5].

[5] Perceived congestion = 0.55*Density + 2.38 ($R2=0.101$)
(0.068) (0.137)

The question on the willingness to pay for a congestion-free first class was not well received, probably for socio-political reasons. More than 70% of respondents answered "no" for even a moderate price differential of 10%. Preferences for a class are not merely determined by the measurable characteristics of the various classes, but also by the sociological attitudes of users[6].

The question on the willingness to trade additional travel time for additional comfort met with more success. Only 29% of the respondents said they were not ready to spend even 5 additional minutes in order to travel under non-congested conditions, and many of these negative answers were recorded at times of limited actual (and perceived) congestion. In other words, a large majority (71%) of respondents exhibit a willingness to pay (in time) to travel in non-congested condition. 39% are ready to spend between 5-10 additional minutes to that effect. And 27% to spend 10-15 additional minutes for the same purpose. These answers form the basis for the econometric analysis and the determination of the congestion cost curve.

III – Producing a Congestion Cost Curve

Several econometric analysis were conducted to estimate the willingness to pay (WTP) of a passenger i for comfort (comfort defined as a non-congested situation) as a function of the characteristics X_i of passenger i and of his/her trip:

$$WTP_i = f\ (a*X_i)$$

WTP, the *explained variable*, is known in terms of minutes, and has to be translated into money terms. This can be done by means of the French official values of time for urban transport (Ministère de l'Equipement, 2004). The ministry offers two numbers, one for France outside the Paris region (7.6 €/h, in €2000), and one for the Paris region (9.3 €/h) which is 22% higher. We will retain the latter value of time, and adjust it to take account of price and income increases, to 10.8 €/h in 2009 €, or 0.18 €/minute, and conduct the core analysis with this value.

[6] Field work on Mumbai very highly congested suburban trains showed that congestion levels in first class carriages were identital to congestion levels in second class carriages : there are many people ready to pay five times more in order to be compressed with people like them, i.e. people ready to pay five times more.

We are aware that other values can be utilized, such as the value used by RATP, the subway agency, or values that would be a function of the income of travelers. A sensitivity analysis using these alternative values of time is presented at the end of this section.

The *explanatory variables* (the vector of trip characteristics X_i) consist of eight elements: (1) *density*, defined as the number of passenger per m2 at the time of the trip, which is a measure of congestion. This objective indicator was preferred to the subjective indicator, i.e. the congestion ranking by the respondents; all other things equal, one would obviously expect density (i.e. congestion) to explain WTP; (2) *trip length*, defined as the time length of the trip, in minutes; one would also expect WTP to increase with trip length; (3) *trip purpose*, which is a binary explanatory variable taking the value of 1 for work trips and 0 for non-work trips; one would expect the WTP to be greater for work trips; (4) *car availability*, also a binary variable; one would expect car availability to increase WTP; (5) *willingness to pay more for a first class;* it was hypothesized that people ready to pay more for a first class would have a greater WTP to avoid congestion than others; (6) *age;* (7)*gender;* and (8) *residence*, also a binary variable taking the value of 1 for respondents living in Paris municipality and 0 for those residing elsewhere; for these three last characteristics, it was hard to predict how the values would influence WTP.

Discrete choice models

The information collected in the survey on the WTP for comfort is of a discrete nature: respondents answer "yes" or "no" to the question about the amount of extra time they are willing to spend to enjoy non-congested travel conditions. In addition, questions – and answers - are ordered, since a "yes" triggers another question, until a "no" is obtained. Moreover, we do not know the "true" value attached to comfort, just the interval in which the "true" value lies. The probability that they answer "yes" can be explained by means of ordered logit or probit models. The models tested also include alternatives related to the samples utilized and to the treatment of trip purpose.

The results of these discrete choice models showed that three factors, namely density (i.e. actual congestion), length of trip, and acceptance of first class carriages, do increase significantly the willingness to pay for non-congestion. WTP is also positively, although loosely, influenced by car ownership. On the other hand, age, sex, locus of residence of the respondent, and - perhaps more surprisingly - trip purpose and income appear not to have a significant explanatory power.

These results, although interesting in themselves, are not reported here in detail[7] because they cannot be utilized to produce what we are primarily interested in: a congestion cost curve linking cost to density levels.

Simple Constrained Regression Models

To do this, we used the simplest type of models, a simple regression constrained by a zero constant, explaining WTP by Density. The zero constant is a logical necessity. Our WTP is the willingness to pay to avoid present congestion. When density is zero, that is when there is no congestion, travelers cannot be expected to be willing to pay anything to get what they have. In the survey, the WTP in minutes is defined as an interval (e.g. 10-15 minutes); for the regression we need one number, and retain the average (e.g. 12.5 minutes).

It is important to note that the unit of our analysis is the trip. The WTP is the WTP to undertake a non congested trip. This has limitations and advantages. It assumes that trips are homogeneous (the density is the same throughout the entire trip) and that all trips are about similar (in length, for instance), conditions which are obviously not fully met. On the other hand, the trip is a much utilized unit of transport analysis, and it is frequently utilized because analysts (fully aware of the diversity of trips) believe the notion of "average trip" is nevertheless a meaningful and useful concept.

A linear form and an exponential form were used. The exponential regression (of the $WTP=\alpha*density^{\beta}$ type), not reproduced here, turned out not to be significant. The linear form, by contrast, offers very high t-values of the coefficient for density, and is significant at the 99% threshold:

$$\text{Average WTP} = \underset{(0.018)}{0.68}*\text{Density} \qquad R^2=0.70 \qquad (1)$$

As can be seen, the logical need of a zero density, ie a non congestion situation, leading to a zero willingness to pay to go to a non congestion situation, this logical need is respected. This linearity of our

[7] For a complete presentation, see Haywood & Koning (2011)

congestion cost curve is rather convenient. It means that a x% change in density implies a welfare change of about x%*0.68 € per trip (a 10% increase leads to a 0.07 €/trip) — a very convenient finding for policy analysis.

Can this curve be considered as a congestion cost curve? Yes, with some caution. Experimental economics as well as contingent evaluation studies suggest that people give somewhat lower values to welfare gains than to welfare losses, or to put it otherwise that willingness to pay is lower then willingness to accept (Horowitz & Mc Connel 2002)[8]. The true loss of congestion for users is therefore slightly greater than revealed by our willingness to pay questions. We shall assume that the two are equal here, partly because we have no way of correcting for this asymmetry, and partly because we want to be on the safe side. We shall therefore consider our curve to be a cost curve, being it understood that it *underestimates* the true welfare losses generated by congestion.

We shall use (1) as our best estimate of the congestion cost curve on line 1 of the Paris subway. An increase in density of 1 person per square meter causes a welfare loss of 0.68 € per trip. This is an average. For a longer than average trip, the welfare loss will be greater; and it will be lower for a shorter than average trip. Similarly, the welfare loss will be higher for a traveler with a higher than average income, or for a traveler with access to a car. The figure arrived at, 0.68 € per trip, reflects the distribution of length of trips, access to car, and acceptance of a first class system actually observed in our survey.

Cost curve (1) is linear. Exponential regressions (of the $WTP=\alpha*density^{\beta}$ type), not reproduced here, turned out not to be significant. This linearity of our congestion cost curve is rather convenient. It means that an x% change in density implies a welfare change of about x%*0.68 € per trip (a 10% increase leads to a 0.07 €/trip) — a very convenient finding for policy analysis.

Impacts of different VOT

The cost curve corresponding to (1) was estimated with a single value of time of 0.18 € per minute. Other values can be used. Do they change much our cost curve?

[8] This has been known for a long time. As early as 1599, Fernando de Rojas, the illustrious author of *La Celestina* was writing : « Man has more pain at losing what he owns than joy at hoping to get it back, even if this hope is a certainty » (*La Celestina*, Act XV).

First, RATP, the subway agency, although it is 100% government-owned, does not use the government prescribed value of time for the Paris region, but a higher value of 16 €/h, or 0.27 € per minute[9]. This is 50% more. Since our cost curve is linear, increasing the WTP of all respondents by 50% would increase the coefficient by 50%. The cost curve would then become:

$$\text{Average WTP} = \underset{(0.018)}{1.02}*\text{Density} \qquad R^2=0.71 \quad (2)$$

Second, it is known that the VOT increases with income. Since we know the income of each of respondents, we can know more precisely the WTP of each respondent by multiplying the additional time he/she is willing to pay by the value of his/her time. This can be done in two ways. We can assume that the value of time is proportional to income. Because, this is known not to be case, we can assume that the value of time is semi-proportional to income. An income elasticity of 0.7, in line with the literature (Small & Verhoef 2007), is assumed: when income increases by x%, VOT increases by 0.7*x%. Table 3 presents the corresponding VOT (calculated on the basis of the average official 0.18 €/minute) for each revenue class.

Table 3 – Values of time and income classes

Income class	Number of respondents	income €/month)	VOT (in €/min) Constant	Prop.	Semi-prop.
<800	53	650	0.18	0.05	0.072
800-1100	13	950	0.18	0.07	0.095
1100-1400	35	1250	0.18	0.09	0.116
1400-1700	50	1550	0.18	0.12	0.135
1700-2000	77	1850	0.18	0.14	0.154
2000-2500	108	2250	0.18	0.17	0.176
2500-3000	76	2750	0.18	0.21	0.204
3000-3500	89	3250	0.18	0.25	0.230
>3500	84	4000	0.18	0.30	0.268
Total	585	-	-	-	-
Average	-	2372	0.18	0.18	0.18

Source: Own survey for numbers and income.
Notes: VOT = Value Of time. Prop = proportional to income. Semi-prop = semi proportional, calculated with a elasticity of VOT to income of 0.7, under the contraint of an average VOT of 0.18 /minute.

Regression (1) was recalculated using the proportional and semi-proportional VOT, or to be more precise, vectors of VOT. We obtain:

[9] This number comes from an RATP cost-benefit analysis of a tramway line under construction : projected annual time gains of 886,000 h are valued at 14.2 M€, which implies an average value time 16.03 €/h.

Table 4 – Congestion Cost curves, with Different Values of Time

VOT	Curve n°	Coefficient α	σ	R2
Constant (0.18 €/min)	(1)	0.682	0.0182	0.70
Constant (0.27 €/min)	(2)	1.083	0.0182	0.70
Proportional	(3)	0.690	0.0231	0.60
Semi-proportional	(4)	0.687	0.0211	0.64

Notes: The regressions presented are of the form : WTP = α*Density. Semi-proportional means : with VOT reflecting a 0.7 elasticity of VOT to incomee.

It appears that utilizing VOT proportional or semi proportional to income has but little impact upon the quality of the regressions, nor upon the coefficient for density. Equations (1), (2) and (3) tell very much the same story. By contrast, and not surprisingly, the 0.27 €/minute VOT used by RATP changes very much the value of α. In what follows, we will use equation (1) as our basic congestion cost curve, and keep in mind that it is based on an official value for Paris region, which may well be considered low.

IV – Policy implications

The knowledge of public transport congestion costs arrived at with this survey and its analysis remains tentative, and must be utilized with caution. It nevertheless produces orders of magnitudes that can be useful for policy analysis. We shall briefly discuss five policy implications.

Present Paris subway congestion costs

The survey clearly shows that subway passengers interviewed suffer from congestion, and that they are willing to pay (in time) to avoid it. This willingness to pay in time can be translated into a willingness to pay in money by multiplying time by the value of time, estimated to be 0.18 €/minute in Paris. The distribution of this WTP is given in table 5. The average WTP for comfort forgone appears to be around 1.4 € per trip. This is significant. This is about three times the amount passengers actually pay for the service[10].

[10] The exact amount is not known. For 2007, according to RATP (*Statistiques Annuelles 2007*, p.26), "direct traffic income" amounts to 1964 M€. For the same year, the number of trips undertaken amounted to 2873 M. This implies a payment of 0.68 € per trip. This number is an over-estimate of payments for subway usage for two reasons. First, it relates to all RATP services, including non subway trips (rail and bus trips), as well as non subway trip payments – which are often

Table 5 – Distribution of Willingness to Pay for Comfort

WTP range (€/trip)	WTP average (€/trip)	Distribution (% passengers)
Less than 0,9	0.50	29%
0.9 to 1.8	1.35	39%
1.8 to 2.7	2.25	27%
2.7 to 3.6	3.15	3%
More than 3.6	5.00	1%
All	**1.42**	100%

Source : Survey. *Notes* : WTP (willingness to pay) is what people are ready to pay to travel in an experienced non-congested situation. Answers in minutes are translated into euros by multiplying by 0.18 €/minute. WTP average is the average of the bounds of each class. The average for all (1.42 €) is a weighted average.

Unsurprisingly, the use of equation (2) produces the same result. Average measured density on line 1 in our survey is 2.1. The willingness to pay for non congested travel is therefore 2.1*0.68 €/trip, that is 1.43 € per trip.

One might be tempted to extrapolate this number in order to produce an estimate of the congestion cost associated with the Paris subway at large (by multiplying 1.4 €/trip by the total number subway trips in 2007, which was 1388 M). However, such an extrapolation would not justified, for two reasons. One is statistical, the other is theoretical.

The distribution of trips by congestion (or density) levels used is not representative of the distribution of trips by congestion levels for the Paris subway as a whole, for several reasons. One is that our sample relates to one single subway line, line 1, which may or may not be typical of Paris lines. It is known to be less congested than line 13, for instance, but also to be more congested than several other lines. A second is that our sample relates to the morning peak hours (7.45 - 10.15). We can assume that the case of the evening peak hours (16.00 to 20.30) is rather similar, and assume further that there is no congestion (ie low densities) in the non-peak hours. Traffic during these peak hours is estimated to be 47% of total traffic, ie 652 M trips per year. The average congestion cost per trip (1.4 €/trip) should, with these heroic assumptions, be multiplied by 652 M trips (not 1388 M trips) to produce an estimate of 926 M€ for Paris subway as a whole. In the absence of the detailed data on

undistinguishable from subway trips because of joint tickets or cards – which are generally longer and more costly than subway trips. Second, part of the "direct traffic income" is actually paid or reimbursed by employers. 0.50 € would be a better estimate of the actual out of pocket expenditure for an average subway trip.

congestion levels by line and hours that would be required to undertake a statistically meaningful extrapolation, this estimate is the best (the least bad) that can be produced, although admittedly not very reliable.

The theoretical reason for not paying too much attention to an overall estimate of "congestion costs" in the Paris subway is even more compelling. Defining congestion costs in fixed infrastructure transportation systems as the difference between costs at times of low usage and at times of effective usage is highly questionable, not to say meaningless. The fact that it is done routinely for automobile congestion is no excuse. The empty road or the empty subway carriage are not useful reference situations. Subways, just as roads, are not built to be empty and the zero congestion situation is not a valuable reference situation. Because the demand for infrastructure usage changes very much over the course of time, there must be moments when cars slow down each other and when subway passengers reduce the comfort of each other. A road or a subway that would never be congested would imply a shocking overinvestment.

Congestion costs of increased subway patronage

Changes in overall subway demand/supply ratio over time, lead to increases in densities at which all trips are made, and therefore to changes in congestion costs, and to welfare losses (or gains). The linearity of equation (1) facilitates the estimation of such welfare changes.

Let us assume that a given increase α in the overall subway patronage (associated with a fixed supply) leads to a similar percentage increase α in the densities of all trips at all moments. This assumes that this increased patronage does not create bottlenecks that would slow train speed and increase the length of trips – a fairly reasonable assumption. C_i, the congestion cost of a trip i, which was $0.68*d_i$ before the additional patronage will become $0.68*d_i*(1+\alpha)$ thereafter. It will increase by $0.68*d_i*\alpha$:

$$\Delta C_i = 0.68*d_i*\alpha$$

This will be true for all trips, irrespective of the initial congestion structure, or density distribution. The average density of trips (d) multiplied by the total number of trips (N) multiplied by $0.68*\alpha$ will give us the total increase (ΔC) in congestion costs:

$$\Delta C = 0.68*d*\alpha*N$$

In the case of Paris over the 2000-2007 period, the increase in total subway patronage related to constant supply was α=8%. Our survey of line 1 showed that at peak hours, average density was d=2.1. We assume that there was very little or no congestion on off-peak traffic which takes place at low densities, and that densities at peak times of line 1 are representative of densities at peak time on the entire network. Peak hours traffic, which is what we consider for congestion and congestion increases, represented about 47% of total subway traffic, therefore N=652 M trips. We then have:

$$DC = 0.68*2.1*0.08*652,000,000 = 74.5 \text{ M€/year}$$

which is an estimate of the yearly increase in congestion cost caused by the 8% increase in patronage relative to supply. Note that this is calculated with the French government estimate of the VOT for Paris. With the RATP estimate, the increase in congestion costs would be valued at 118 M€.

The cost of increased congestion brought about by increased demand/supply ratios is relevant to the cost-benefit analysis of traffic reduction policies. In Paris, for instance, it is reported that as a result of road shrinking and parking policies introduced in 2001, road traffic decreased by about 16%. The number of car trips with either/and an origin or a destination in Paris municipality was estimated to be around 690 M/year. It therefore decreased by 110 M trips per year. Had all these trips shifted to the subway, they would have increased the subway demand/supply ratio by about 8%, and a welfare loss of 75 M € per year, as estimated above. In reality, the 8% increase in the demand/supply ratio actually recorded comes only in part from this modal shift, and in part from a shift in the demand curve for urban transportation in the Paris agglomeration.

The official policy of the Paris municipality is to decrease car trips (in the Paris municipality area) by 40% over the 2000-2020 period. Assuming that the 280 M car trips per year eliminated are replaced by subway trips (an unrealistic assumption, but the one which is made by Paris officials), and no capacity increase in the subway, this would lead to a 20% increase in subway demand/supply ratio, and hence to a 187 M €/year welfare loss in terms of congestion.

Marginal congestion cost

Equation (1) depicts the individual congestion cost I(d) suffered by a traveler who takes a subway, and is a function of density d. In so doing, our traveler increases congestion, and the congestion cost borne by all other travelers. He/she inflicts upon them a congestion externality E(d) which is a function of density. This externality is equal to the derivative of I(d) multiplied by density. The total or social congestion cost S(d) is the sum of the individual cost and of the externality cost

$$I(d) = 0.68*d$$

$$E(d) = I'(d)*d = 0.68*d$$

$$S(d) = I(d) + E(d) = 1.36*d$$

S(d) varies with density. At times of high densities, such as the average density of our sample plus one standard-error for instance, i.e. about 3, S(d) equals about 4 euros. This is the congested-related social cost of a trip in the Paris subway at times of serious congestion, in addition to the individual time cost of the trip and to its economic cost.

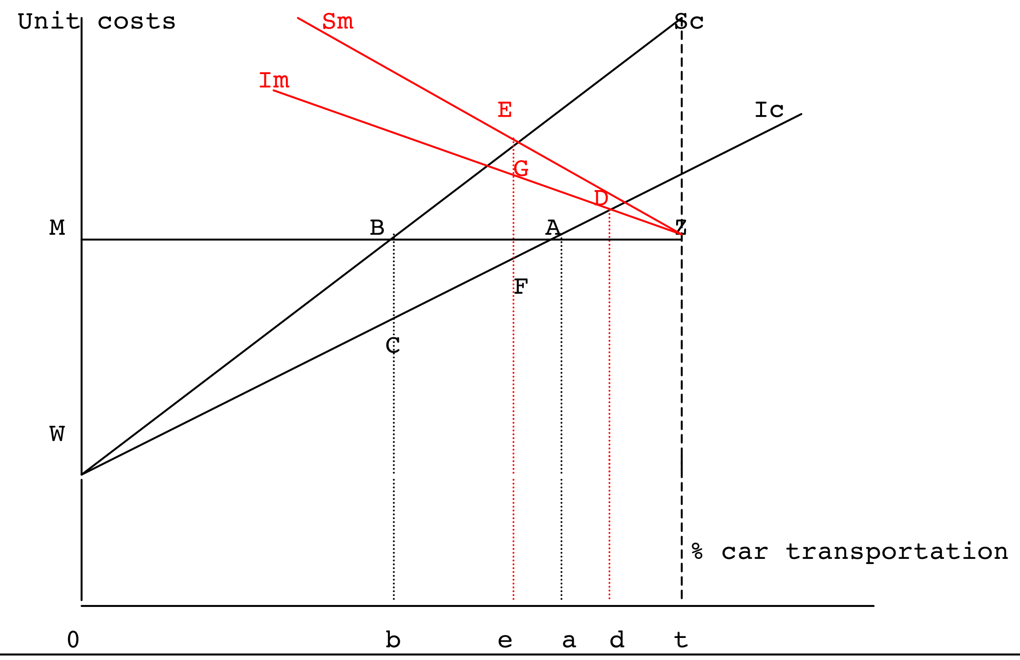
Investment appraisal

Investment in subway supply is costly, but not impossible. Since it decreases densities (which are equal to patronage divided by capacity), it produces welfare gains that must be appreciated. Hence the question: by how much should an investment of n M € reduce densities in order to be justified? We calculated the discounted cash flow (at a social interest rate of 4% over a 30 years period) of these yearly welfare gains for various densities reductions. A subway capacity increase of 6% would yield a yearly benefit of 56 M€ in lower congestion costs, and over 30 years, a total benefit of about one billion euros. In other words, an investment of up to one billion is justified if it increases subway capacity by 6%. Since all magnitudes are linear, a two billion euros investment requires a capacity increases of 12%.

A computable model of optimal modal share

An important theoretical and practical question for many cities is: what is the optimal share of car transportation and public transportation? The answer is usually based on the assumption of zero congestion in public transportation. This implies that the social cost of a public transport trip is independent of the quantity of such trips, with PT individual cost equal to PT social cost. Car transportation, by contrast, is (correctly) assumed to be congestible, with both individual and social costs of car trips increasing with the number of car trips. With these assumptions, the optimal modal share is the one for which the marginal cost of car transportation is equal to the constant cost of public transportation. This standard model can - and should - be modified by dropping the assumption of no public transport congestion.

Figure 1 – Subway and car congestion in an urban area



This is illustrated in Figure 1. It assumes that total transportation consists of car transportation and subway transportation, and that modal split is driven by cost considerations. I_c is the individual cost curve for cars or vehicles as a function of car usage; it is assumed to increase because of road congestion (this cost curve is here linear for the sake of simplicity; in reality it is quadratic). ZM is a family of cost curves of subway transportation as a function of metro usage.

If we assume that there is no subway congestion and that the metro cost curve to consider is ZM, which is constant, Ic intersects ZM in A. To the left of A, vehicle transportation is cheaper than metro transportation, and travelers use their cars. In A, car transportation costs, including the congestion cost borne by car users, are about equal to subway costs and an equilibrium is reached. The share of car transportation is equal to a.

As is well known, this natural equilibrium is suboptimal, because it neglects the road congestion externality which is equal to the derivative of I_c multiplied by the quantity of cars. Curve S_c, equal to the sum of I_c and of this externality, is the social cost curve for car transportation. It intersects ZM in B, which is the standard optimal point. This standard optimum can be reached by means of a congestion charge equal to the externality at that level of car usage, ie equal to BC. The optimal share of car transportation is, in our figure reduced to b.

Let us now introduce public transport congestion. The metro cost curve is no longer ZM, but I_m. I_m and I_c, the two individual cost curves intersect in D. This is the true equilibrium point. The resulting share of car transportation is now increased to d.

This natural equilibrium is not optimal however. The optimum is reached when the subway social cost curve S_m (defined just as the car social cost curve) intersects this car social cost curve S_c, in E. This point will be reached when and if two congestion charges are levied on users : a toll on cars, equal to FE (the marginal car congestion cost at that level of car usage), and a toll on subways, equal to GE (the marginal subway congestion cost at that level of subway usage). The resulting optimal share of car transportation will then be e.

This highly simplified model[11] shows how important it is to take public transportation congestion into consideration if we are to define truly optimal policies.

Our estimate of the subway cost curves (in addition to the quite standard motor vehicles cost curves) makes it possible to implement this model (See Annex A for a detailed presentation of the calculations).

In 2001, the number of car trips to, within and from Paris municipality was 0.69 billion trips per year, representing a 36% modal share. In the language of Figure 1, d=36%. For that number of car trips, the cost of a car trip is equal to the cost of a metro trip, taking into consideration subway congestion as well as car congestion.

Ignoring subway congestion, the equilibrium number of car trips would be much lower, at 0.15 billion trips,

[11] A more realistic model would include non congestion externalities in the social cost curve for cars (air pollution, CO2, etc) , as well as the (usually important) public transport subsidies in the social cost curve for public transportation.

representing an 8% modal share, with a=8%. For that number of car trips, the cost of a car trip (including congestion) would be equal to the time and money cost of a subway trip (excluding congestion). This shows that the system is very sensitive to the number of car trips, and that ignoring subway congestion costs can lead to very erroneous predictions.

The optimal number of car trips – for which the social cost of a car trip is equal to the social cost of a subway trip – is 0.46 billion, representing a modal share of 24%. At the optimum, e = 24%. This is less than the equilibrium modal share in the absence of policies.

This optimum requires a double toll. Car trips should be tolled 2.43 € per trip. Subway trips should be tolled 1.53 € per trip.

All these numbers must be taken with great prudence, for a number of reasons. One is that our indicators of congestion, be they for car or for subway, are global (number of trips per year divided by a given road or metro supply), when in reality congestion is in essence a changing phenomenon over time and space. What we have is a distribution of congested and less congested situations, which is poorly described by a single average number.

The other reason for prudence is that a number of parameters (trip length, maximum car speed, car speed in 2001, money costs of car and subway trips, etc.) are fragile estimates from various sources. In addition the very concept of "trip" which is utilized is questionable: it lumps together city center to city center trips and suburbs to city center trips to consider what happens in the city center part of such trips.

The model also assumes that the total number of trips is given (1.94 billion trips per year) and that the problem is merely to allocate them between the two modes considered. In reality, we know that cost increases in one mode might well lead to a decrease in the total number of trips, and not merely to modal shifts.

The model also equates public transportation with subway transportation, ignoring bus transportation. Relative to patronage increases (as mentioned above), the behavior of bus transportation is completely different from that of subway transportation. This may not matter too much in the case of Paris, where bus transportation is a small share (about 10%) of public transportation, but would be a serious problem in cities, such as London, where bus transportation is an important share of public transportation.

The numbers arrived at should therefore be seen as indicative or illustrative numbers rather than as reliable estimates. They nevertheless suggest that ignoring subway congestion can lead to serious misunderstandings and misallocations in transport policies. This is particularly relevant for the evaluation of an urban toll system. A share of the car users tolled away will use public transport systems, and cause public transport congestion costs that should be taken into consideration in appraising the car toll system.

V – Conclusions

This paper is a modest contribution to redress a curious and major imbalance in the attention given to public transport congestion as opposed to automobile congestion. The "public transport" considered here consists of subways and trains, operating with a fixed infrastructure endowment; it ignores bus transportation. The congestion considered here relates to the crowding of people in carriages; it therefore ignores the crowding of trains on infrastructure networks, which is another form of public transport congestion. The physiology of public transport congestion (thus defined) is very similar, not to say identical, to that of road congestion. When the number of people on a given train, or the number of cars on a given road, increase, the resulting crowding creates losses of comfort in one case and of time in the other, and the individual costs of transportation increase in both cases. The social costs increase even more, because of the reciprocal externalities involved. In joining a congested subway carriage or a congested road, not only will I enjoy less comfort or less speed (relative to what I would enjoy in non-congested situations), but I will also decrease the comfort or the speed of all other fellow travelers. The public policy implications of these twin types of transport congestion are massive, in terms of large public investments and delicate pricing decisions. Yet, if the knowledge base of road congestion is by now significant, the knowledge base of public transport congestion is not.

The reason is that it is much more difficult to assess and value the comfort losses associated with public transport congestion than the time losses associated with car transport congestion. We tried to conduct such an assessment and valuation on the case of the Paris subway, by using a contingent evaluation methodology. A survey of about 700 passengers was undertaken. For each of them, the effective degree of congestion, defined as density (people per square meter in the subway train) at the time and location of interview, was recorded. Only passengers who stated that they had been traveling in "non-congested" situations were considered, to make sure the respondents would have an "experienced reference". They were not asked directly how much they were ready to pay in money in order to travel in non-congested conditions, because such questions do not elicit meaningful responses. They were asked if they would prefer an hypothetical non-congested train taking more time (5, 10, 15, and 20 minutes more) rather than the actual congested train. Most of the respondents appeared ready to spend 5-10 minutes or 10-15 minutes more. This made it possible to produce for each respondent (by multiplying by a value of time) the willingness to pay in order to avoid congestion.

A significant relationship between willingness to pay for non-congested travel and actual congestion was obtained with a simple regression constrained by a zero constant, and can be seen as a subway congestion cost curve:

Congestion cost per trip = 0.68 € * congestion (density)

This 0.68 coefficient reflects the impact of the other explanatory variable. If the present average length of trips (or rather its distribution) were to increase, the value of the coefficient would also increase. If present car ownership rates were to decrease, the value of the coefficient would also decline.

Since present congestion (density) is about 2.1, it means that, on average, travelers would be ready to pay - at least - an additional 1.4 € to travel under non-congested conditions. This is the price they attach to it, or to put it otherwise, the cost actual congestion imposes upon them. This is about 3 times as much as what they actually pay in money terms to use the subway.

The knowledge of this congestion cost curve makes it possible to estimate the welfare costs (or gains) of changes in congestion levels. Over the 2002-2007 period, average congestion (again defined as number of travelers per square meter) increased by 8% : this implies a 75 M € per year increase in congestion costs.

These numbers can also be used to find out whether capacity investments in the subway (that decrease density) are justified by the congestion reduction they cause. A one billion euro investment would thus be justified if it increased capacity by about 6%.

We are well aware of the tentative character of this research. Our empirical estimates remain very fragile, and are presented as examples of what could and should be done, rather than as hard findings.

References

Armelius, Hanna & Lars Hultkrantz. 2006. « The politico-economic link between public transport and road pricing : An ex-ante study of the Stockholm road-princing trial ». *Transport Policy*, vol 13, pp. 162-172.

Cantwell, M., B. Caufield, and M. O'Mahony (2009): "Examining the Factors that Impact Public Transport Commuting Satisfaction," *Journal of Public Transportation*, 12(2), pp. 1–22.

Evans, G. and R. Wener (2007): "Crowding and personal space invasion on the train: Please don't make me sit in the middle," *Journal of Environmental Psychology*, 27, 90-94.

Haab, T.C. & McConnel, K.E. 2003. *Valuing Environmental and Natural Resources: the Econometrics of Non-Market Valuation*, Edward Elgar

Haywood L. and Koning M. 2011. « Pushy Parisian Elbows: Evidences of Taste for Comfort using Contingent Valuation Methodology", Working Paper, University Paris 1 (to be submitted for publication).

Horowitz, J.K. & McConnel, K.E. 2002. "A Review of WTP/WTA Studies", *Journal of Environmental Economics and Management,* vol. 44(3), pp. 426-447.

Kopp, Pierre & Rémy Prud'homme. 2010. "The Economics of Urban Tolls: Lessons from the Stockholm Case". *International Journal of Transport Economics*. 37:2. pp. 195-222.

Krauss, M. 1991. Discomfort externalities and marginal cost transit fares". *Journal of Urban Economics*, 29, pp. 249-259.

Leurent, Fabien & Kaisheng Liu. 2009. *On Seat Congestion, Passenger, Comfort and Route Choice in Urban Transit: a Network Equilibrium Assignment Model with Application to Paris*. 2009 TRB Meeting, Paper n° 09-1784, 16 p.

Lindsey, R. & Verhoef, E.T. 2000. "Congestion Modelling", in D.A. Hensher and K.J. Button (eds.) *Handbook in Transport 1. Transport Modelling*, Oxford, Pergamon

Litman, Todd. 2008. *Valuing Transit Service Quality Improvements*. Victoria Transport Policy Institute (www.vtpi.org)

Ministère de l'Equipement. 2004. *Instruction-cadre relative aux méthodes d'évaluation économique des grands projets d'infrastructures de transport*.

Mohring, H. 1972. "Optimization and Scale Economies in Urban Bus Transportation", *American Economic Review*, vol. 62(4), pp. 591-604.

Prud'homme, Rémy & Juan Pablo Bocarejo. 2004. "The London congestion charge: a tentative economic appraisal". *Transport Policy*. 12, n°3, pp. 279-88.

Small, K. A. and E.T Verhoef, 2007, *The Economics of Urban Transportation*, Routledge, 267 p

Transek. 2006. (for Stockholmsforsöker). "Cost-beneft analysis of the Stockholm trial" (www.stockholmsforsoket.se).

Tsekeris, T and S. Vos. 2009. "Design and Evaluation of Road Pricing : State-of-the-art and methodological advances". *Netnomics*. 10, pp. 5-52.

Wardman, Mark and Whelan, Gerard. 2010. "Twenty years of ril crowding valuation studies: Evicdence and lessons from British experience. Transport Reviews. Pp. 1-20

Wener, R., G. Evans, and P. Boately (2005): "Commuting stress: Psychological effects of a trip and spillover into the workplace," *Transportation Research Board*, 1924/2005, 112-117.

Annex A : A Computable Model of Combined Car and Subway Transportation in Paris

We have :

Q_m : Number of subway trips to, within, or from Paris municipality ;

Q_c : Number of car trips to, within, or from Paris municipality

$I_m(Q_m)$: individual cost of a subway trip

$I_c(Q_c)$: individual cost of a car trip

$S_m(Q_m)$: Social cost of a subway trip, with $S_m(Q_m)=I_m(Q_m)+I'_m(Q_m)*Q_m$

Sc(Qc) : Social cost of a car trip, with $S_c(Q_c)=I_c+I'_c(Q_c)*Q_c$

In 2002, we have, in billion trips per year : Q_m=1,25; Q_c=0,69 and Q_m+Q_c=1,94

$I_m(Q_m)$ – The individual cost of a metro trip is equal to its money cost + its time cost + its congestion cost.

The money cost of a subway trip has been estimated to be around 0.50 €/trip.

The time cost can be estimated as the time spent (about 30 minutes) by the value of time (0.18 €/minute), or 5.4 €/trip.

As explained in the text, the congestion cost is equal to 0.68*density, expressed in terms of people/square meter. We want to express it in terms of trips/year Q_m, and can write it as: $0.68*\lambda*Q_m$. We know that in 2009, density = 2.1, and that Q_m = 1.388. It follows that λ = 1.513, that $0.68*\lambda$ = 1.0288, and that congestion cost = $1.0288*Q_m$.

We therefore have:

$$I_m(Q_m) = 5.9 + 1.0288*Q_m$$

Which can also be expressed as a function of Qc, since Qm=1.94-Qc:

$$I_m(Q_c) = 7.9 - 1.0288*Q_c$$

$I_c(Q_c)$ – The individual cost of car trip is equal to its money cost + its time cost. The congestion cost is included in the definition of the time cost.

The money cost includes the fuel cost, the car amortization, the parking fees in the Paris municipality. It was estimated to be around 3.4 €/trip.

The time cost is equal to the value of time (10.8 €/year) x time spent. The time spent is equal to the trip length (6 km) / speed. It is the speed which is a function of car usage Qm: speed = maximum speed-α*Qc. Maximum speed is estimated to be 28 km/h in Paris municipality (the speed recorded between 6am and 7am). In 2001, actual average speed was 17 km/h, and Qc was 0.69 billion, which yields α=16. The time cost is therefore : 10.8€/h*6km/(28-16*Q_c) or 64.8/(28-16*Q_c)

Hence:

$$I_c(Q_c) = 3.4 + 64.8/(28-16*Q_c)$$

$S_m(Q_m)$ – The social cost of a subway trip is equal to the individual cost plus the congestion externality, ie the derivative of the individual cost multiplied by Qm:

$$S_m(Q_m) = 5.9 - 2.056*Q_m$$

Which can also be expressed as a function of Qc by replacing Qm by 1.94-Qc :

$$S_m(Q_c) = 9.9 - 2.056*Q_c$$

$S_c(Q_c)$ – The social cost of a car trip is similarly calculated as:

$$S_c(Q_c) = 3.4 + 64.8/(28-16*Q_c) + 1037/(28-16*Q_c)^2$$

Equipped with these cost curves, we can calculate the following magnitudes.

- The equilibrium number of car trips: it is the value of Qc for which Ic(Qc)=Im(Qc). Unsurprisingly, it turns out to be 0.69 billion car trips/year. As a matter of fact, we adjusted the money cost of car trips (which is difficult to estimate) to obtain that result.

- The optimal number of car trips : it is the value of Qc for which Sc(Qc)=Sm(Qc). It is 0.46 billion car trips/year.

- The car toll and the subway toll per trip required to reach this optimum.

$$\text{Car toll} = S_c(0.46) - I_c(0.46) = 2.43 \text{ € / trip}$$

$$\text{Subway toll} = S_m(0.46) - I_m(0.46) = 1.53 \text{ € / trip}$$

A more complete formulation would add in the calculation of the social cost curve of subway trips the important subsidy (about 1.5 €/trip) benefiting subway transportation. It would also include in the calculation of the social cost curve of car trips the air pollution and CO2 emissions externalities associated with car trips. It could also include the high specific taxes, mostly fuel taxes, levied on car trips.