

# A model of modal choice with heterogenous agents

Quentin David\*, CREA, Université du Luxembourg.

Renaud Foucart†, ECARES, Université Libre de Bruxelles.

March 17, 2011

## Abstract

We build up a model where a mass of heterogenous commuters has to choose between public transportation and the car. Our main assumption comes from cross modal externalities: depending on the policy implemented, the congestion generated by car affects the quality of public transportation. We show that such a model can lead to multiple equilibria, and that the equilibrium involving the largest use of public transportation is always Pareto-enhancing. We discuss two policy tools: tax and separation of public transportation from car congestion. We show that, when one of the two is not strictly more efficient than the other, then separation should be preferred for large-scale policies, while taxation should be preferred for marginal modifications of commuting patterns.

## Contents

<b>1</b>	<b>Literature</b>	<b>4</b>
1.1	Congestion costs . . . . .	4
1.2	Network externalities in public transportation . . . . .	5
1.3	Heterogeneous agents . . . . .	5
<b>2</b>	<b>The model</b>	<b>6</b>
2.1	Setup . . . . .	6
2.1.1	Heterogeneous preferences for the car use . . . . .	6
2.1.2	The network externality of PT . . . . .	7
2.1.3	Car use, congestion and commuting time . . . . .	7
2.1.4	The game . . . . .	10
2.2	Equilibria . . . . .	11
2.2.1	Existence . . . . .	11
2.2.2	Uniqueness and multiplicity . . . . .	11
2.2.3	Efficiency . . . . .	13
<b>3</b>	<b>Governement's intervention and welfare</b>	<b>15</b>
3.1	Policy tools . . . . .	15
3.2	Efficiency of policy tools . . . . .	16
3.3	Graphical illustration . . . . .	20

---

\*Contact: quentin.max.david@gmail.com

†Contact: rfoucart@ulb.ac.be

<b>4 Discussion</b>	<b>21</b>
<b>A Technical appendices</b>	<b>28</b>
A.1 Proof of proposition 2: multiple equilibria . . . . .	28
A.2 Proof of proposition 3: Efficiency . . . . .	29
A.3 Proof of proposition 4 . . . . .	30
A.4 Proof of Lemma (2) . . . . .	30
A.5 Proof of Lemma (3) . . . . .	31
<b>B Figures</b>	<b>34</b>

In many cities, the congestion caused by cars is an increasing issue, as much for the ease of the transportation within the city as for environmental concerns. We build up a theoretical model in which heterogenous commuters have to decide simultaneously to use a car or public transportation. We assume that the car users (CU) generate congestion on all the commuters and that users of public transportation (PT) enjoy a positive network externality (the efficiency of PT increases with the number of users). We show that it is also possible that ex-ante similar cities end up with very different modal shifts, due to the presence of multiple equilibria (ME). In the presence of ME, we demonstrate that the equilibrium involving the lowest share of CU Pareto dominates all the other equilibria. We also study two policies: taxation and traffic separation. Both can be used to imperfectly enforce coordination. We show that the main drawback of taxation happens when the number of car users is small. A shrinking tax base can be detrimental to the last car user and poorly efficient. The main drawback of separation is observed when the number of car users remain high, as it increases congestion for cars.

The scientific literature addressing this issue can be conveniently divided in two families: the physical planning and the car dependant cities. The purpose of this paper is not to reject neither to support any of those schools. As will be made clear below, we show that both approaches can be pertinent, but correspond to different scope of policies.

We denote the first one by ‘physical planning’. They emphasize the fact that the structure of the city is the most important driver of commuting patterns. The main idea to improve transportation policies is to have a shift towards ‘transit oriented development’. Belzer and Aultier

(2002, p.7) define it as follows: ‘transit-oriented development must be mixed-use, walkable, location-efficient development that balances the need for sufficient density to support convenient transit service with the scale of the adjacent community’<sup>1</sup>.

We denote the second by ‘car dependant cities’. The idea is that, given both, the structure of our cities and, the intrinsic preferences that many consumers have for the car, one should focus on the best way to accomodate traffic flows, while the question of public transportation is of minor interest. It is rather consensual among economists that road pricing should be the privileged way to deal with congestion problems (Beesley and Kemp, 1987, Calfee and Winston, 1998) and the ‘games of congestion’ have been widely studied in economic theory (e.g. Rosenthal, 1973). Many applied papers also deal with congestion costs and car taxation. One of the most famous results is due to Vickrey (1963). He argues that pricing should vary at different times of the day as to make commuters pay for the marginal cost of congestion.

In the late eighties, Pucher (1988) observed that ‘*Urban transportation and traveler behavior vary widely, even among countries with similar per capita income, technology and urbanization*’. Kenworthy and Laube (1999) documented this puzzle and show that the fraction of workers using transit is 6 times higher in wealthy Asian cities compared to the US. They also show that the commuting time to work is higher (and cheaper) where the use of public transport is higher and that the cost recovery of transit increases with the share of passengers using it. In other words, cities where transit is intensively used appear to need a smaller share of subsidies for operating it. This is confirmed by Pucher et al. (1983) who state that ‘*[D]irect benefits to transit riders have been small relative to the increase in subsidy [and] alleged environmental and secondary economic benefits are negligible or inexistant*’. While Pucher and Renne (2003) computed that, in the US, public transport accounts for less than 2% in the urban travel in 2001.

---

<sup>1</sup>Cervero et al. (2002, p.2), emphasize that it does ‘involve some combination of intensifying commercial development around stations, inter-mixing land uses, layering in public amenities (e.g., civic spaces, landscaping), and improving the quality of walking and bicycling’. One should also consider the book by Dittmar and Ohland (2003) that summarizes literature and ‘good practices’ in transit oriented development.

In our model, the existence of ME come from anticipations on the modal shift among heterogeneous commuters. The fraction of CU affects the costs and benefits of each mode of transportation. An increase in the share of CU increases not only the costs of transportation by car (congestion) but also the cost public transportation by reducing the positive network externality associated with having many commuters using PT.

The paper is organized as follow. In the next section, we review the literature that we associate to our research. We use this review to provide support for our hypothesis. In section 3, we present the model, and show under which conditions ME emerge and discuss their relative efficiency. In section 3, we discuss two policy tools that can be used to coordinate agents to the best equilibrium (especially in the ME case). We consider two policy tools: taxation of cars and traffic separation between cars and public transportation. We provide theoretical results showing that, if no policy is strictly more efficient than the other, then separation should be preferred for large-scale policies. Minor changes in a world where the share of CU is high are better dealt with taxation. We also provide graphical illustrations of those results, giving some intuition for policy makers. We discuss some policies (such as building an underground) and conclude in section 4.

## **1 Literature**

### **1.1 Congestion costs**

The effect of the number of cars on the efficiency of transport has been widely studied in the economic literature. Anas and Small (1998, p.1456) describe this negative externality as follows: ‘The congestion externality arises because the user of a motor vehicle does not pay for its marginal contribution to congestion’.

Mirabel (1999) considers similar externalities as we do. This is the so called ‘crossed modal externalities’, i.e. the fact that the congestion generated by cars also affects the efficiency of public transport. Those ‘crossed modal externalities’ have been empirically measured in Brussels by Dobruzkas and Fourneau (2007).

De Vlieger et al. (2000) empirically show that these congestion costs are concave in terms of pollution. In terms of perceived cost, Wardman (2001) shows in a meta-analysis that time spent in congested traffic is valued 50% higher. This is again an argument for the concavity of the cost of congestion, as congestion (i) increases travel time and (ii) increase the marginal cost of travel time. This principle is applied by Santos and Bhakar (2006) to assess the benefits of the congestion toll in London.

## 1.2 Network externalities in public transportation

A relatively large literature also exists on the network effect of the number of transit users on the efficiency of public transport. For instance, Mohring (1972, p.591) explains that *‘Transportation differs from the typical commodity price theory texts in that travelers and shippers play a producing, not just a consuming role’*.

The idea is the existence of a so-called ‘dynamic network externality’. If the demand for bus service doubles, a company is expected to double the number of busses serving the route, at the same *per capita* price. Thus, the waiting time for an individual commuting by bus decreases, which improves the efficiency of public transportation.

In a quite similar spirit, Tabuchi (1993) considers a railway parallel to a road. The commuters can use public transportation (the railway) or car (the road). The CU are expected to face congestion in a bottleneck. The users of public transport face a positive network externality, by sharing the fixed costs of the railway. This differs from our model in two ways: (i) there is full separation between public and private transportation and (ii) commuters are homogeneous in type.

## 1.3 Heterogeneous agents

A paper by Berhoef and Small (2004) considers heterogeneous agents in a model of pricing for car use only, but the question of heterogeneity of preferences among agents has, to the best of our knowledge, not been taken into account in economics models of modal choice.

However, several empirical papers confirm the heterogeneity of preferences among commuters. One could refer to Beirão et al. (2007), Handya et al. (2005) Jensen (1999), Steg (2005) Hiscock (2002) or Van Vught et al. (1996).

## 2 The model

### 2.1 Setup

We consider a closed city with a mass 1 of commuters. The space is finite and it is not possible to increase neither the number of road nor the number of the traffic lines. Our assumption concerning the structure of the city are not specific. We assume that people are exogenously located within the city and have to commute everyday to another exogenous place (within the city) either using a car (C) or public transport (PT). One can consider that they commute either to get to work, to search for a job, etc. We also assume that commuters have intrinsic preference for the car (relative to PT) drawn from a given distribution. We consider that these preferences do not only correspond to psychological costs, but also to more physical costs, as for instance the ease of access to the public transportation network.

In comparizon to the use of public transportation, we suppose that the use of a car implies an additional (fixed) cost,  $f_c$  and produces congestion for the other CU and for PT as long as the traffic lines are not fully separated. We also assume congestion costs to be concave. On the other side, we consider that PT users produce a (positive) network externality on public transportation that can be interpreted alternatively as a lower price for a given level of quality, or a higher quality for a given price.

For the sake of clarity, we present and discuss one by one the main hypotheses described above.

#### 2.1.1 Heterogeneous preferences for the car use

Commuters are heterogeneous with respect to their preference  $\varepsilon_i$  for the use of a car compared to the use of public transport. This parameter  $\varepsilon_i$  encapsulates all the heterogeneity among

commuters preferences and is supposed to embody many dimensions: the intrinsic degree of love for cars, the access to public transportation, the daily distance to cover, etc. We assume that  $\varepsilon_i$  comes from a cumulated distribution function  $F$  whose support is  $(-\infty, +\infty)$ , that is some love public transport so much that they would never accept not to use it ( $\varepsilon_i \rightarrow -\infty$ ) while others will never use public transports ( $\varepsilon_i \rightarrow +\infty$ ). The associated density is  $f$ .

$$\varepsilon_i \sim F(\varepsilon)$$

### 2.1.2 The network externality of PT

As stated before, we model the network externality through the waiting time for PT users. The idea is that if there are more user, the frequency of public transport increases and the waiting time decreases. For simplicity, we assume this network externality to be linear. If there are  $n$  users of public transport, the waiting time of each of them is given by  $W(n) > 0$  ( $W(n) \in \mathbb{R}_0^+$ ), and

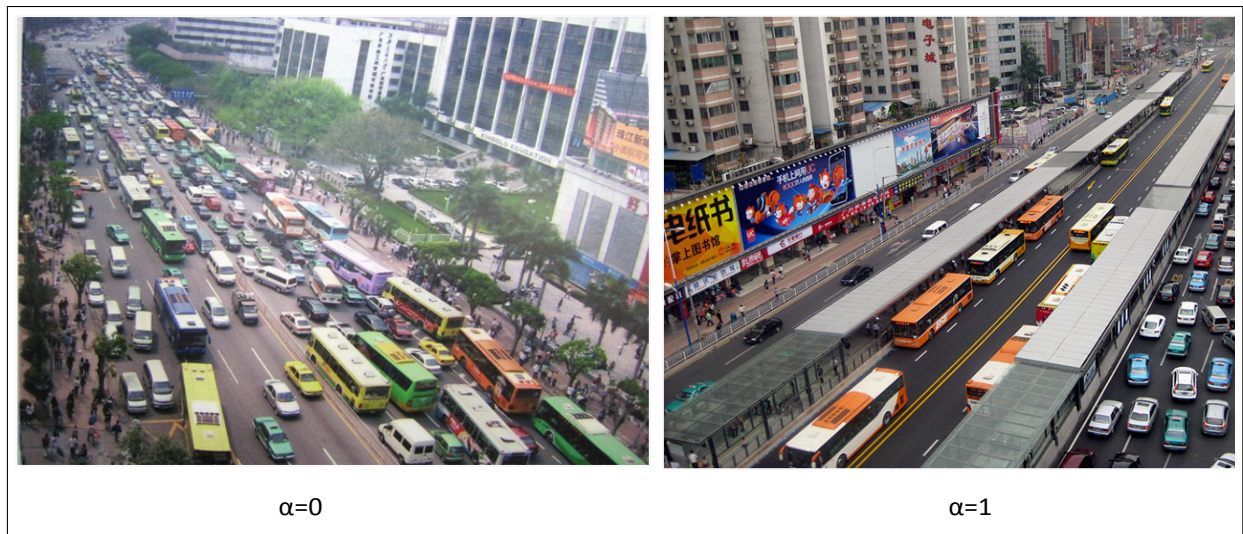
$$W'(n) < 0 \text{ and } W''(n) = 0$$

### 2.1.3 Car use, congestion and commuting time

The CU generate externalities on both the other CU and on PT. We assume that the effect of the congestion on public transport depends on the share of CU and the share of road exclusively dedicated to PT.

Therefore, if  $n$  is the number of users of PT,  $z = (1 - n)$  workers commute by car. We assume that travel time is an increasing function of  $z$ , whatever the modal choice. Furthermore, a share of the road can be exclusively dedicated to public transport. We denote  $\alpha$  ( $0 \leq \alpha < 1$ ), the share of roads dedicated to public transport, that is, exempt from congestion. As we assume a finite amount of space in the city, there is a share  $(1 - \alpha)$  of the roads that are accessible to both, cars and public transports. Therefore, more roads devoted to public transportation reduces, by the same amount, the space available for cars (we discuss as an extension the possibility of building an underground that allows to improve the efficiency of public transport without reduction of

Figure 1: Illustration from Ghuanzou (China): two extreme values of  $\alpha$



space for cars). We start by assuming that  $\alpha$  is exogenously determined and discuss in Section 3.1 the possibility for a central planner to use  $\alpha$  as a political tool and study its efficiency. An illustration of two different values of  $\alpha$  is provided in Figure 1.

The commuting time of CU is denoted by  $t^c(\alpha, z)$  and by  $t^{pt}(\alpha, z)$  for PT. As discussed in subsection 1.1, we assume that congestion is increasing and concave in  $z$ . Moreover, a higher degree of separation of cars and public transport (higher  $\alpha$ ) is likely to generate more congestion for cars (because of less space for them) and less congestion for PT. This effect on CU increases with the fraction of CU (the idea is the following: separation has an impact only if there is actually a problem of congestion).

Therefore,  $t^c(\alpha, z)$ , the commuting time of CU has the following properties

$$\begin{aligned}
 \frac{\partial t^c(\alpha, z)}{\partial z} &> 0 \text{ and } \frac{\partial^2 t^c(\alpha, z)}{\partial z^2} \geq 0, \\
 \frac{\partial t^c(\alpha, z)}{\partial \alpha} &> 0 \text{ and } \frac{\partial^2 t^c(\alpha, z)}{\partial \alpha^2} = 0, \\
 \frac{\partial^2 t^c(\alpha, z)}{\partial z \partial \alpha} &> 0.
 \end{aligned} \tag{1}$$



and  $t^c(\alpha, z)$ , the time spend commuting by public transport is such that

$$\begin{aligned} \frac{\partial t^{pt}(\alpha, z)}{\partial z} &> 0 \text{ and } \frac{\partial^2 t^{pt}(\alpha, z)}{\partial z^2} \geq 0, \\ \frac{\partial t^{pt}(\alpha, z)}{\partial \alpha} &< 0 \text{ and } \frac{\partial t^{pt}(\alpha, z)}{\partial \alpha^2} = 0, \\ \frac{\partial^2 t^{pt}(\alpha, z)}{\partial z \partial \alpha} &< 0. \end{aligned} \tag{2}$$

On the one hand, the efficiency of PT decreases with the fraction of CU, but this impact can be mitigated by offering some separation from car traffic. On the other hand, the efficiency of separation is more important when the fraction of CU is high.

We also assume that (i)  $t^{pt}(0, z) = t^c(0, z)$  i.e. if there is no separation, commuting times are equal; (ii)  $t^{pt}(\alpha, 0) = t^c(\alpha, 0) \forall \alpha \in [0, 1]$  i.e. if there is no congestion and some place available for CU, commuting times are equal (iii)  $t^{pt}(\alpha, z)$  and  $t^c(\alpha, z) \in \mathbb{R} \forall \alpha \in [0, 1], z \in [0, 1]$ , i.e. commuting times are finite (iv)  $\frac{\partial t^c(\alpha, z)}{\partial z} > \frac{\partial t^{pt}(\alpha, z)}{\partial z} \forall \alpha \in (0, 1)$ , i.e. the marginal effect of congestion is higher for cars than for public transport as long as there is some separation (v)  $\frac{\frac{\partial t^{pt}(\alpha, z)}{\partial \alpha}}{\frac{\partial t^c(\alpha, z)}{\partial \alpha}}$  is constant. This measures the ‘efficiency’ of the policy - what share of the difference in time generated by separation corresponds to an increase in transportation time for CU, and what share corresponds to a decrease in transportation time for PT.

**Definition 1** *The additional time required to commute by car in comparison to the time using PT is given by*

$$\Delta_1(\alpha, z) = t^c(\alpha, z) - t^{pt}(\alpha, z) \tag{3}$$

Using the properties of  $t^c(\alpha, z)$  and  $t^{pt}(\alpha, z)$ , we have

**Lemma 1** *Properties of  $\Delta_1(\alpha, z)$ .*

- (i)  $\Delta_1(\alpha, z) \geq 0$ ;
- (ii)  $\frac{\partial \Delta_1(\alpha, z)}{\partial \alpha} > 0$ ;
- (iii)  $\frac{\partial \Delta_1(\alpha, z)}{\partial z} > 0, \forall \alpha > 0$ ;
- (iv)  $\Delta_1(0, z) = \Delta_1(\alpha, 0) = 0$ ;

(v) **Supermodularity of  $\Delta_1(\alpha, z)$** : the effect of separation on the differential of commuting time increases with congestion (with the number of CU), i.e.  $\frac{\partial^2 \Delta_1(\alpha, z)}{\partial z \partial \alpha} > 0$ .

**Proof.** The results (i), (ii) (iii) and (iv) come straightforward from the assumptions in (1) and (2). Here is the proof for result (v). From the properties described in (1), we have  $\frac{\partial^2 t^c(\alpha, z)}{\partial z \partial \alpha} > 0$  and from the properties in (2), we have  $\frac{\partial^2 t^{pt}(\alpha, z)}{\partial z \partial \alpha} < 0$ . The definition of  $\Delta_t(\alpha, z) = t^c(\alpha, z) - t^{pt}(\alpha, z)$  leads to:

$$\frac{\partial^2 \Delta_1(\alpha, z)}{\partial z \partial \alpha} = \frac{\partial^2 t^c(\alpha, z)}{\partial z \partial \alpha} - \frac{\partial^2 t^{pt}(\alpha, z)}{\partial z \partial \alpha} > 0$$

An illustration of the shape of  $t^c(\alpha, z)$ ,  $t^{pt}(\alpha, z)$  and  $\Delta_1(\alpha, z)$  is presented in Appendix B. ■

#### 2.1.4 The game

**Definition 2** The *modal choice* for a commuter consists in choosing the mode of transportation (CU or PT) that maximizes its utility. The modal choice is a simultaneous game among a mass 1 of commuters.

The utility of a commuter  $i$  commuting by car is given by

$$U_i^c(\alpha, z) = -f_c - t^c(\alpha, z) + \frac{\varepsilon_i}{2} \quad (4)$$

where  $f_c$  the fixed cost associated to the use of a car,  $t^c(\alpha, z)$  is the commuting time associated to the use of a car and  $\varepsilon_i$  is the individual preference for the use of a car.

The utility of a commuter  $i$  using PT is given by

$$U_i^{pt}(\alpha, z) = -W(n) - t^{pt}(\alpha, z) - \frac{\varepsilon_i}{2} \quad (5)$$

where  $W(n)$  is the waiting time,  $t^{pt}(\alpha, z)$  is the commuting time associated to the use of PT and  $\varepsilon_i$  is the individual desutility associated with the use of PT. Without loss of generality, we normalize to  $\frac{\varepsilon_i}{2}$  in both utility function for the ease of exposition, as the differential in utilities becomes  $\varepsilon_i$ .

Commuter  $i$  uses commutes by car if  $U_i^c(\alpha, z) > U_i^{pt}(\alpha, z)$ , which is the case when

$$\varepsilon_i > f_c - W(n) + [t^c(\alpha, z) - t^{pt}(\alpha, z)]$$

Using (3), this last equation rewrites

$$\varepsilon_i > f_c - W(n) + \Delta_1(\alpha, z) \tag{6}$$

In the illustrations provided in the next section, the curves depicted correspond to both sides of equation 6. Note that, when  $\alpha = 0$ , the right-hand side is linear.

## 2.2 Equilibria

### 2.2.1 Existence

**Proposition 1** *There exist, at least, one Nash Equilibrium in pure strategy*

**Proof.** First, we sort individuals according to their relative preference for the car in an ascending order ( $\varepsilon_i < \varepsilon_j \forall i < j$ ).

Assume that individual  $i$  is such that  $\varepsilon_i > f_c - W(n) + \Delta_1(\alpha, z)$ , ( $i$  commutes by car). Therefore, for any individual  $j > i$ , we have  $\varepsilon_j > f_c - W(n) + \Delta_1(\alpha, z)$  i.e.  $j$  also commutes by car. The same reasoning applies if  $\varepsilon_i < f_c - W(n) + \Delta_1(\alpha, z)$ ,  $\forall j < i$ . In a Nash Equilibrium (NE),  $n = n_k$  is a solution to  $\varepsilon_i = f_c - W(n) + \Delta_1(\alpha, n)$  where  $F(i) = n$ . Therefore, as the support of  $F$  is  $(-\infty, +\infty)$ , there exist at least one NE such that:

$$n_k = F [f_c - W(n_k) + \Delta_1(\alpha, z)] \tag{7}$$

■

### 2.2.2 Uniqueness and multiplicity

**Proposition 2** *There exist multiple equilibria if and only if there exist a solution  $n_k$  satisfying the following condition:*

$$\frac{\partial [f_c - W(n_k) + \Delta_1(\alpha, z)]}{\partial n} > \frac{\partial F^{-1}(n_k)}{\partial n} \tag{8}$$

**Proof.** We provide here the intuition of the proof, the full development is presented in Appendix A.1.

1. As the support for  $\varepsilon_i$  is  $(-\infty, \infty)$ , we know that  $\lim_{x \rightarrow 1} F^{-1}(x) = -\lim_{x \rightarrow 0} F^{-1}(x) = +\infty$ , i.e. there is a left hand side asymptote in 0 toward  $-\infty$  and a right hand side asymptote in 1 toward  $+\infty$  (some people will always take their car, some will always use PT);
2. From the domain of the parameters, we know that the slope of  $f_c - W(n) + \Delta_1(\alpha, z)$  with respect to  $n$  is strictly lower than infinity;
3. If  $f_c - W(n) + \Delta_1(\alpha, z)$  crosses  $F^{-1}(n)$  (definition of an equilibrium) in a point  $n_k$  where the slope of  $f_c - W(n) + \Delta_1(\alpha, z)$  with respect to  $n$  is higher than the slope of  $F^{-1}(n)$ ; then the function  $f_c - W(n_k) + \Delta_1(\alpha, z)$  will be above (resp. below)  $F^{-1}(n_k)$  on the right (resp. left) neighborhood of the equilibrium;
4. From point 2, this means we have at least 3 equilibria.

■

Intuitively, this need for a zone of high density corresponds to unimodality of preferences. For multiplicity of equilibria to happen, one needs to have few people with extremely polarized preferences, and a large fraction of people with close preferences. For instance, figure 2 represents a city with 3 equilibria when  $\alpha = 0$ . One can find two ‘stable’<sup>2</sup> equilibria, one with really few users of PT ( $n_1$ ) and one with a large fraction ( $n_3$ ). There is also one so-called ‘non-stable’ equilibrium. We show in the next subsection that these first equilibria are inefficient. The fraction of commuters that does not make the same choice in both equilibria are all better off by taking public transportation in equilibrium  $n_3$ .

---

<sup>2</sup>The concept of stability we use throughout the paper implies that if there is a small change in the beliefs on  $n$  the best responses still converge to this equilibrium.

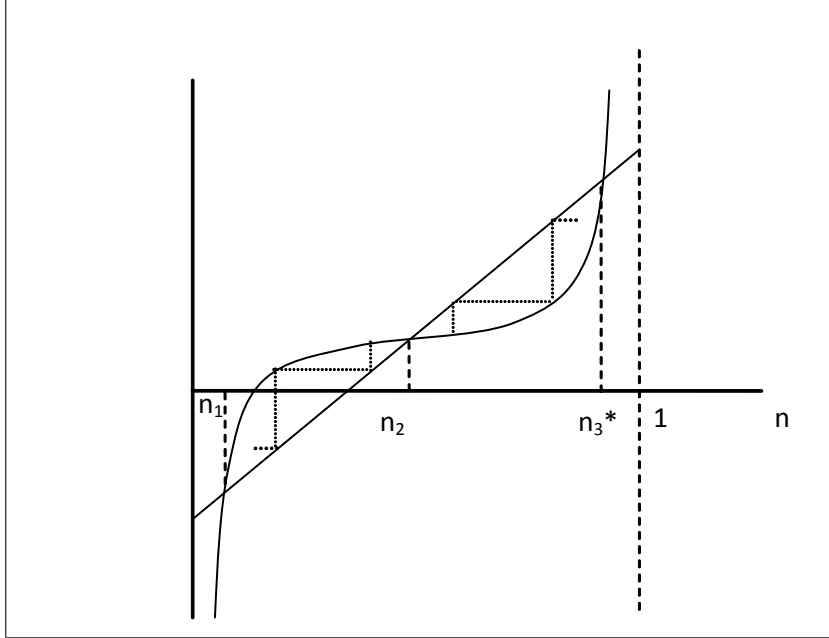


Figure 2: An illustration of multiple (3) equilibria

### 2.2.3 Efficiency

**Proposition 3** *If there are multiple equilibria, the equilibrium involving the higher use of PT Pareto dominates all the other equilibria. We denote this equilibrium by  $n^*$ .*

**Proof.** The formal proof is provided in Appendix A.2. The intuition is the following, using figure 2 as an illustration. People commuting by car in both situations (those located on the right of  $n_3$ , i.e. with a very strong preference for using a car) are better off because they face less congestion if the equilibrium is located in  $n_3$  rather than in  $n_1$ . People using PT in both equilibria (those located on the left of  $n_1$ , i.e. which have lower preference for the car) are also better off because public transportation is now more effective thanks to the network externality. Finally, one can show that people changing their modal choice (those located between  $n_1$  and  $n_3$ ) from using car to PT are necessarily better off. Indeed,  $n_3$  is now associated with a higher utility than the one they had in  $n_1$  if they use a car (since there is less congestion), but they choose to use public transportation instead. This implies, by revealed preferences, that they

necessarily improve their utility by doing it and  $n_3 = n^*$ . ■

It follows from this proposition that, if we have multiple equilibria and the prevailing one is not the one involving the highest use of PT, any costless coordination device leading to another equilibrium associated with a higher use of PT would increase efficiency.

### 3 Government's intervention and welfare

In this section, we assume that a government can set up two different policies likely to affect the commuters' modal choice. We assume that the government's objective function is to improve welfare by, if possible, set Pareto improving policies.

Note that in our setting, a government could set a very high tax and to remove it almost instantaneously in order to force commuter to coordinate on the efficient equilibrium. Nevertheless, we believe that it is not realistic as the switching dynamic, to switch from one equilibrium to the other is a long and progressive process that we cannot implement in our model without considering a dynamic model. To solve this issue, we assume that setting a policy implies, for the government to keep it in place forever. This assumption can be seen as being ad hoc, but it is realistic for an intervention to take effect to be credible and assuming that the government has to commit on the permanence of its application seems to be a nice and convenient hypothesis.

#### 3.1 Policy tools

We assume that the government have two policy tools at disposal and study their effects on individual's welfare: the **taxation** of CU ( $T$ ) and the possibility to change the traffic **separation** between CU and PT ( $\alpha$ ).

We assume that a **taxation policy** implies to levy  $T$  on every CU and that this tax is redistributed lump-sum among all commuters<sup>3</sup>. Therefore, every commuter receives a transfer  $zT$  and CU pay  $T$ . The new utility functions become

$$\begin{aligned} U_i^c(\alpha, T, z) &= -f_c - t^c(\alpha, z) - (1-z)T + \frac{\varepsilon_i}{2} \\ U_i^{pt}(\alpha, T, z) &= -W(n) - t^{pt}(\alpha, z) + zT - \frac{\varepsilon_i}{2} \end{aligned}$$

After the introduction of a taxation policy, a commuter  $i$  uses public transportation iff

$$\varepsilon_i < \Delta(1\alpha, z) - W(n) + T + f_c$$

---

<sup>3</sup>As will be made clear below, assuming the tax is lost only affects marginally the results presented in the next subsection.

and, by proposition 1, there is always a Nash Equilibrium. As we do not limit the size of the tax, there always exist a  $T$  such that the only Nash Equilibrium with taxation is on the right-hand side of  $n^*$ .

The **traffic Separation** is the other considered policy. Assume that the government set a new traffic separation,  $\alpha' : \alpha' > \alpha$ , the new utility functions are given by

$$\begin{aligned} U_i^c(\alpha', T, z) &= -f_c - t^c(\alpha', z) + \frac{\varepsilon_i}{2} \\ U_i^{pt}(\alpha', T, z) &= -W(n) - t^{pt}(\alpha', z) - \frac{\varepsilon_i}{2}. \end{aligned}$$

Now, a commuter  $i$  uses public transportation if

$$\varepsilon_i < \Delta_1(\alpha', z) - W(n) + f_c$$

Here, one cannot theoretically affirm that the impact of traffic separation is sufficient to keep only one equilibrium.

### 3.2 Efficiency of policy tools

**Definition 3** For a given Nash equilibrium,  $n_k$ , we define the **initial swing commuter** as the commuter indifferent between CU and PT before the introduction of a policy and the **final swing commuter** as the commuter indifferent between CU and PT after the introduction of a policy.

For the sake of simplicity, we start by comparing these two policies by considering their effect in a setting where they produce the same instantaneous utility differential (before commuter adjusts their modal choice) computed for the initial swing commuter.

**Definition 4** If  $\bar{z}$  denote the decentralized equilibrium number of car users before the introduction of a policy, we define  $\xi_i$  as the utility differential computed for the initial swing commuter ( $i$ ) holding modal choices unchanged, i.e.

$$\xi_i(\alpha', T, \bar{z}) = U_i^{pt}(\alpha', T, \bar{z}) - U_i^c(\alpha', T, \bar{z})$$



In other words,  $\xi_i$  is the instantaneous utility differential faced by the initial swing commuter when we hold commuters' behaviour constant. Note that according to the definition of  $\xi(\alpha', T, \bar{z})$ , the entire expression should be:  $\xi = [U_i^{pt}(\alpha', T, \bar{z}) - U_i^{pt}(\alpha, \bar{z})] - [U_i^c(\alpha', T, \bar{z}) - U_i^c(\alpha, \bar{z})]$ , but by definition, at equilibrium,  $U_i^c(\alpha, \bar{z}) - U_i^{pt}(\alpha, \bar{z}) = 0$  for the swing commuter.

Consider two policies (taxation and traffic separation) yielding to the same  $\xi$ , i.e. consider  $T$  and  $\alpha'$  such that  $\xi(\alpha', 0, \bar{z}) = \xi(\alpha, T, \bar{z})$  and denote the new utilities (after the implementation of any of these two policies) by  $U'$ .

We call  $z^T$  and  $z^\alpha$  the equilibria after the implementation of policies involving, respectively, **taxation** or **separation**.

We know by definition of our externalities that all the initial public transportation users are better off with the implementation of any of those policies. Indeed, they enjoy higher network externalities (more users of PT) and face less congestion (less CU) and, in case of taxation policy, they receive a lump sum transfer, while, in case of traffic separation, they have an additional decrease in congestion (since PT enjoy a higher share of roads exempted from congestion).

We consider as an objective of the social planner to make the remaining car users and commuter who change their modal choice as good as possible. As it will be clear further; this corresponds to a Rawlsian welfare function.

**Lemma 2** *For any utility differential ( $\xi$ ) obtained with taxation only or traffic separation only ( $\xi = \xi(\alpha', 0, \bar{z}) = \xi(\alpha, T, \bar{z})$ ), the number of car users in equilibrium after the tax  $z^T$  is lower than the number of car users in equilibrium after the separation  $z^\alpha$ , i.e.*

$$z^T \leq z^\alpha \quad \forall \xi > 0$$

**Proof.** The entire proof is provided in appendix. The idea is that **taxation** implies

$$\xi(\alpha, T, \bar{z}) = T$$

while **separation** implies

$$\xi(\alpha', 0, \bar{z}) = \Delta_1(\alpha', \bar{z}) - \Delta_1(\alpha, \bar{z}).$$

Since this last equation is strictly decreasing in  $z$  (and as it is not the case of the previous one), for any  $\xi$ , the equilibrium level of car users after the implementation of any of these two policies is such that  $z^\alpha > z^T$ . ■

**Lemma 3** *A policy of **taxation** is Pareto improving iff*

$$[t^c(\alpha, \bar{z}) - t^c(\alpha, z^T)] > (1 - z^T) T$$

*while a policy of **traffic separation** is Pareto improving iff*

$$t^c(\alpha, \bar{z}) > t^c(\alpha', z^\alpha)$$

**Proof.** See appendix A.5. ■

We can see that a policy of traffic separation is Pareto improving if it reduces the congestion for car users. It implies a trade off between less CU (lower  $z$ ), concentrated over less traffic lines (lower  $\alpha$ ). The combination of these two effects must reduce congestion for the policy to be Pareto improving.

The reduction of congestion due to a policy of taxation must compensate the cost of that taxation.

For the first part of lemma 3, we can already see that, if  $z$  is sufficiently large, the condition is not extremely restrictive. Indeed, the tax levied on car users is largely compensated by the payoffs coming from the lump-sum benefit of the considered tax. However, when  $z$  decreases, the tax base shrinks, making this condition more and more restrictive.

For the second part of lemma 3, one can conveniently rewrite this result by separating two effects, the positive effect on cars (decrease in congestion due to the lower number of cars) and the negative effect (increase in congestion due to the separation).

$$t^c(\alpha, z) - t^c(\alpha', z^\alpha) = [t^c(\alpha, z) - t^c(\alpha, z^\alpha)] - [t^c(\alpha', z^\alpha) - t^c(\alpha, z^\alpha)] \quad (9)$$

The ability of a traffic separation policy to be Pareto improving depends on the relative importance of these two forces.

**Proposition 4** *If there exist an utility differential  $\xi$  such that car users are better off in equilibrium under **separation** than under **taxation**, then for any  $\xi' > \xi$ , car users are better off in equilibrium under separation than under taxation.*

**Proof.** The formal proof is provided in appendix A.3 We show that cars are better off under **separation** then under **taxation** iff

$$(1 - z^T)T > [t^c(\alpha', z^\alpha) - t^c(\alpha, z^\alpha)] + [t^c(\alpha, z^\alpha) - t^c(\alpha, z^T)]$$

The left-hand side is increasing in  $\xi$  while the right-hand side is decreasing in  $\xi$ . Hence, if there exist a  $\xi$  such that this inequality is true, then it is true for any  $\xi' > \xi$  ■

This result is important in terms of policy and comes from two sides: supermodularity and the size of the tax base<sup>4</sup>.

Consider an increase in **taxation** for a share of car users  $z$ . For this value of  $z$ , the utility of a CU decreases by  $T(1 - z)$ . But the tax mechanically decreases  $z$ , as it changes the preferences of the switching users. On the one hand, this decreases the time spent in congestion. But on the other hand this also increases the tax burden of the remaining car users. Indeed, those keep on paying the same nominal tax, but the tax base shrinks in the meantime, reducing the amount received as a lump sum transfer.

Consider now a policy of **separation**  $\alpha' > \alpha$ . There is also an initial shock  $t^c(\alpha', z) - t^c(\alpha, z)$  for car users. The separation also mechanically decreases  $z$  but, as opposed to taxation, this change in  $z$  only yields positive effects on car users. First, the congestion decreases. Second, by supermodularity, the marginal cost of  $\alpha'$  on congestion of CU also decreases when  $z$  decreases.

We cannot theoretically exclude the possibility that one of the two policies is always better. But, if it is not the case, **taxation** is better for small changes in  $z$ , while **separation** should be preferred for larger changes. We illustrate this graphically in the next subsection.

---

<sup>4</sup>any of those would be enough to support this result.

### 3.3 Graphical illustration

We propose here three different illustrations of policy implementation under multiple equilibria<sup>5</sup>. These cases are depicted in Figure 3 to 5. The formal welfare conditions have been computed in the previous subsection. The goal of this subsection is to provide some relevant intuitions for policymakers.

1. Suppose  $n_j < n^*$  which corresponds to the case where the initial equilibrium involves too much CU and,  $n^\alpha = n^T < n^*$ . This corresponds to a situation where the implementation of our considered policies is not sufficient to reach the Pareto dominating equilibrium. We assume that these two policies lead to the same equilibrium ( $n^T = n^\alpha$ ) for convenience only. By supermodularity of the policy tool  $\alpha$ , one can easily show that  $n^T = n^\alpha$  implies the ‘good’ equilibria does not vary with the same amplitude.

For such a policy, intuitively, one should be more kind to use **taxation**, as the tax base is large (and then the monetary cost per car user is low) while the taxation doesn’t affect negatively the travel time. On the other hand, **separation** should be avoided, as it increases the travel time for all of the many car users. This effect is exacerbated by the fact that the costs of separation are supermodular.

2. Suppose  $n_j < n^*$  and  $n^* < n^\alpha \leq n^T$  is observed if the implemented policy allows to switch toward the best equilibrium. It implies even more PT than the decentralized equilibrium. We clearly see how the relationship between  $n^\alpha$  and  $n^T$  ( $n^\alpha \leq n^T$ ) is due to supermodularity in Figure 4. Here, we observe that the shock is important (and potentially Pareto-improving). However, the impact of **taxation** is very strong. Indeed, the remaining car users pay the nominal tax, but as the tax base is extremely small, they

---

<sup>5</sup>One can extrapolate cases 1 and 3 to cases of unique equilibria, with respectively low (high) share of PT.

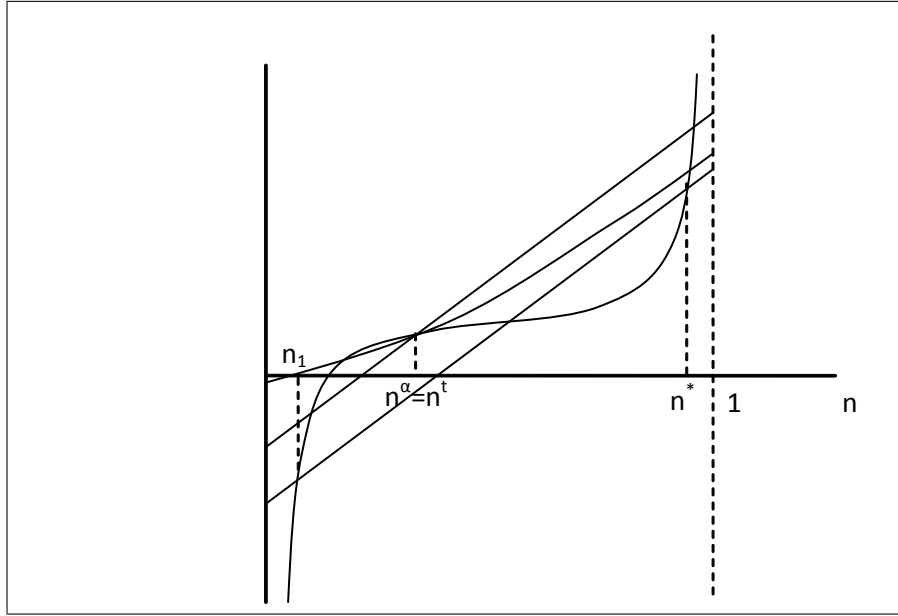


Figure 3: Increasing  $\alpha$  and  $T$  in the case of two decentralized equilibria, but not enough to reach the ‘good’ equilibrium

receive almost no transfer from the tax (similarly, the PT users do not significantly benefit from the tax). On the other hand, the impact of **separation** on the remaining car users is not important, as congestion is very small.

3. Suppose that  $n_j = n^*$ , the ‘good’ decentralized equilibrium, and a social planner implements a policy such that  $n^* < n^T = n^\alpha$ . Here, a **separation** policy is of small impact, as congestion is small. However, **taxation** has an important impact on car users, as its positive effects (congestion) are almost inexistent, while its negative effect are high (small tax base). It can be clearly seen as a redistribution device either than a policy instrument.

## 4 Discussion

### Building an underground

For the moment, we have considered  $\alpha$  as a way to dissociate traffic by diminishing the

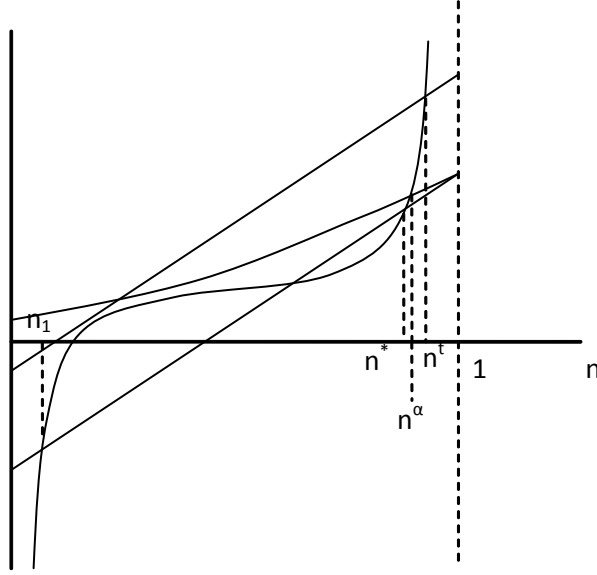


Figure 4: Increasing  $\alpha$  and  $T$  in the case of two decentralized equilibria in order to move close to the ‘good’ decentralized equilibrium

fraction of the road dedicated to cars. Intuitively, this can correspond to bus lines or light rail. Another way to prevent public transportation from congestion is building an underground. An underground is known to be more costly to build, while delimiting bus lines is almost free.

Assume  $\alpha$  is now the investment in the underground (with cost  $M(\alpha)$  paid lump sum by all commuters, independently of the chosen mode of transportation). One should expect the following properties: (i) The time spent commuting by car is given by  $t_c(\alpha, z)$  with  $\frac{\partial t_c(\alpha, z)}{\partial \alpha} \leq 0$ , because the underground does not reduce at all (and potentially increases) the space for cars in the city and (ii) the time spent commuting using public transportation is given by  $t_{pt}(\alpha, z)$  with  $\frac{\partial t_{pt}(\alpha, z)}{\partial \alpha} < 0$ , because it protects public transportation users from congestion.

This implies that  $\Delta_t(\alpha, z)$  is smaller than using bus lines, making it more difficult to get rid of the ‘bad’ equilibrium. The potential new equilibrium is very close to  $n^\alpha$  (as congestion is low when only few commuters use car), but every single commuter has to bear a cost  $M(\alpha)$ . However, an argument to defend the existence of an underground can be the very fact of congestion in

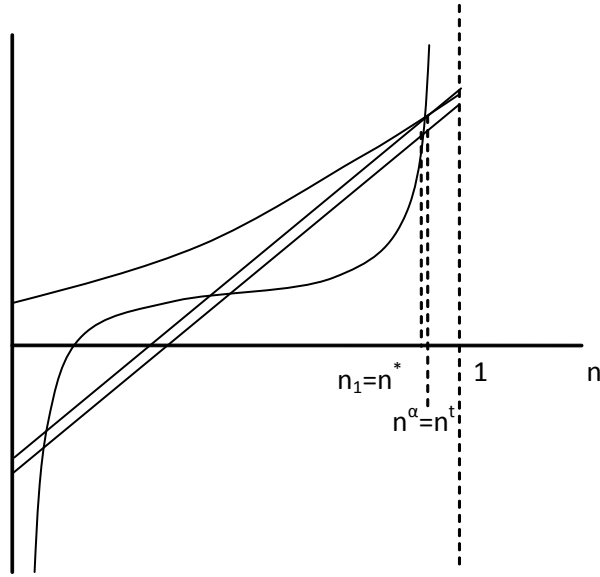


Figure 5: Increasing  $\alpha$  and  $T$  when already in the decentralized Pareto dominant equilibrium

the public transport (as a way to expand the supply).

### The outside option

In the model, we have considered a game implying only two choices for a commuter: the car or public transportation. However, there also exist outside options: staying at home, taking a bike or leaving the city. In this framework, the focus of our welfare analysis should be even more put on the remaining car users.

### Conclusions

In this paper, we have shown that the conjunction of externalities of congestion, cross modal externalities and network externalities with heterogeneous agents can lead to multiple equilibria. This can help explaining the fact that cities that are similar in many ways end up with different patterns of car use. We also show that political tools are not equivalent, and discuss the welfare implications of taxation and road planning. We show that, when one of the two is not strictly more efficient than the other, then separation should be preferred for large-scale policies, while taxation should be preferred for marginal modifications of commuting patterns.

This strongly relates to the two ‘schools’ presented in our introduction. A policy maker that believe, as the so-called ‘physical planners’ that there must be an important change in the modal split should focus on the allocation of space, by reducing the one devoted to cars. Social planners only concerned by marginal changes in the cost of cars - or that simply believe in the car dependency of their city - should focus on taxation.



## References

- [1] ANAS, Alex, ARNOTT, Richard, SMALL, Kenneth (1998). Urban Spatial Structure ; *Journal of Economic Literature* ; 36 ; 3 ; pp.1426-1464
- [2] BEESLEY, Michael and KEMP, Michael (1987). Urban Transportation ; *Handbook of Regional and Urban Economics* ; vol. 2 ; chpt. 26 ; pp.1023-1052
- [3] BEIRÃO, Gabriela and CABRAL, Sarsfield (2007). Understanding attitudes towards public transportation and private car: A qualitative study ; *Transport Policy* ; 14 ; pp.478-489
- [4] BELZER, Dena and AULTIER, Gerald (2002). 'Transit oriented development: moving from rhetoric to reality' ; Brookings Institutions Center on Urban Metropolitan Policy ; Discussion paper ; 55pp.
- [5] CALFEE, John and WINSTON, Clifford (1998). The value of automobile travel time: implications for congestion policy ; *Journal of Public Economics* ; 69 ; pp.83-102
- [6] CERVERO, R., C. FERRELL and S. MURPHY. (2002). Transit-Oriented Development and Joint Development in the United States: A Literature Review. TCRP Research Results Digest Number 52, National Research Council: Washington, D.C.
- [7] DE VLIIEGER, I ; DE KEUKELEERE, D ; KRETZSCHMAR, J.G. (2000). Environmental effects of driving behaviour and congestion related to passenger cars ; *Atmospheric Environment* ; 34 ; pp.4469-4655
- [8] DITTMAR, Hank, and OHLAND, Gloria (2003). The new transit town: best practices in transit-oriented development, Island Press.
- [9] DOBRUZKES, Frédéric and FOURNEAU, Yves (2007). Coûts directs et géographie des ralentissements subis par les transports publics bruxellois, Brussels Studies, N°7.

- [10] HANDYA, Susan, WESTOMB, Lisa and MOKHTARIANC Patricia (2003). Driving by choice or necessity?, *Transportation Research Part A: Policy and Practice* ; pp. 183-203
- [11] HISCOCK, Rosemary ; MACINTYRE, Sally ; KEARNS, Ade and ELLAWAY, Anne (2002). *Means of transport and ontological security: Do cars provide psycho-social benefits for their users ?* ; *Transportation Research part D* ; 7 ; pp.1119-1135
- [12] JENSEN, Mette (1999). Passion and Earth in transport: a sociological analysis on travel behaviour ; *Transport Policy* ; 6 ; pp.19-33
- [13] KENWORTHY and LAUBE (1999). patterns of automobile dependance in cities: an international overview of key physical and economic dimensions with some implications for urban policy ; *Transportation Research part A* ; 33 ; pp.691-723
- [14] MIRABEL, François (1999). Répartitions modales urbaines, externalités et instauration de péages: Le cas des externalités de congestion et des ‘externalités modales croisées’ ; *Revue économique* ; 50 ; 5 ; pp.1007-1027
- [15] MOHRING, Herbert (1972). *Optimization and Scale Economies in Urban Bus Transportation* ; *The American Economic Review* ; 62 ; 4 ; pp.591-604
- [16] NEWMAN, Peter ; KENWORTHY, Jeffrey and VINTILA, Peter (1995). Can we overcome automobile dependance ? Physical planning in an age of urban cynicism ; *Cities* ; 12 ; 1 ; pp.53-65
- [17] NEWMAN, Peter and KENWORTHY, Jeff (2000). The Ten Myth of Automobile Dependence ; *World Transport Policy and Practice* ; 6 ; 1 ; pp.15-26
- [18] OATES, Wallace (1999). An Essay on Fiscal Federalism ; *Journal of Economic Literature* ; 37 ; 3 ; pp.1120-1149

- [19] Pucher, John (1983). Distribution of Federal Transportation Subsidies: Cities, States, and Regions, *Urban Affairs Review*, 19: 191-216
- [20] Pucher, John (1988). 'Urban Travel Behavior as the Outcome of Public Policy: The Example of Modal-Split in Western Europe and North America,' *Journal of the American Planning Association*, Vol. 54, No. 4, pp. 509-520
- [21] PUCHER, John and RENNE, John (2003). Socioeconomics of Urban Travel: Evidence from the 2001 NHTS ; *Transportation Quarterly* ; 57 ; 3 ; pp.49-77
- [22] ROSENTHAL, Robert (1973). A class of Games Possessing Pure-Strategy Nash Equilibria ; *International Journal of Game Theory* ; 2 ; 1 ; pp.65-67
- [23] Santos, Georgina and Bhakar, Jasvinder (2005). The Impact of the London congestion charging scheme on the generalized cost of car commuters to the city of London from a value of travel time savings perspective ; *Transport Policy* ; 13 ; pp.22-23
- [24] STEG, Linda (2005). Car use: lust and must. Instrumental, symbolic and affective motives for car use ; *Transportation Research part. A* ; 39 ; pp.147-162
- [25] TABUCHI, Takatoshi (1993). Bottleneck Congestion and Modal Split ; *Journal of Urban Economics* ; 34 ; pp.414-431
- [26] VAN VUGT, Mark ; VAN LANGE, Paul and MEERTENS, Ree (1996). Commuting by car or public transportation ? A social dilemma analysis of travel mode judgements ; *European Journal of Social Psychology* ; 26 ; pp.373-395
- [27] VERHOEF, Erik and SMALL, Kenneth (2004). Product Differentiation on Road. Constrained Congestion Pricing with Heterogeneous Users ; *Journal of Transport Economics and Policy* ; 38 ; 1 ; pp.127-156

[28] VICKREY, William (1963). Pricing in Urban and Suburban Transport ; *The American Economic Review* ; 53 ; 2 ; pp.452-465

## A Technical appendices

### A.1 Proof of proposition 2: multiple equilibria

(i) We know from proposition (1) that an equilibrium is a solution to

$$n^* = F(f_c - W(n^*) + \Delta_1(\alpha, z)) \quad (10)$$

(ii) Applying a transformation  $F^{-1}$  to (4) one can rewrite an equilibrium as

$$F^{-1}(n^*) = f_c - W(n^*) + \Delta_1(\alpha, z)$$

First, we show that if there exist a solution  $n^*$  satisfying the following condition:

$$\frac{\partial[f_c - W(n^*) + \Delta_1(\alpha, z)]}{\partial n^*} > \frac{\partial F^{-1}(n^*)}{\partial n^*}$$

then one can find multiple equilibria. Second, we show that this is a necessary condition.

If there exist such a  $n^*$ , than for any  $\eta$  as small as possible, and  $\eta > 0$ , we have

$$\begin{aligned} F^{-1}(n^* + \eta) &< f_c - W(n^* + \eta) + \Delta_1(\alpha, z) \\ F^{-1}(n^* - \eta) &> f_c - W(n^* - \eta) + \Delta_1(\alpha, z) \end{aligned}$$

, which implies at least three equilibria as, by definition,  $F$  has support on  $(-\infty, \infty)$  and  $f_c - W(1) + \Delta_1(\alpha, 1) < \infty$  and  $f_c - W(0) + \Delta_1(\alpha, 0) > -\infty$ .

Multiplicity happens only if this condition is true. Assume

$$\frac{\partial[f_c - W(n^*) + \Delta_1(\alpha, z)]}{\partial n^*} < \frac{\partial F^{-1}(n^*)}{\partial n^*}$$

then we have

$$\begin{aligned} F^{-1}(n^* + \eta) &> f_c - W(n^* + \eta) + \Delta_1(\alpha, z) \\ F^{-1}(n^* - \eta) &< f_c - W(n^* - \eta) + \Delta_1(\alpha, z) \end{aligned}$$

, which implies only one equilibrium as, by definition,  $F$  has support on  $(-\infty, \infty)$  and  $f_c - W(1) + \Delta_1(\alpha, 1) < \infty$  and  $f_c - W(0) + \Delta_1(\alpha, 0) > -\infty$ .

## A.2 Proof of proposition 3: Efficiency

Assume there exist  $T$  equilibria  $n_1 < n_2 < \dots < n_T$

(1) We want to show that  $n_T$  Pareto dominates any equilibrium  $n_i$

(2) While comparing  $n_i$  to  $n_T$ , we have

(a) For any  $\varepsilon_i$  such that  $F(\varepsilon) < n_i$ , the best response is (PT) in both equilibria. We want to show that those users are better off in equilibrium  $n_T$ , i.e. that:

$$\begin{aligned} W(n_i) + t^{pt}(\alpha, n_i) + \varepsilon_i &> W(n_T) + t^{pt}(\alpha, n_T) + \varepsilon_i \\ W(n_i) - W(n_T) &> t^{pt}(\alpha, n_T) - t^{pt}(\alpha, n_i) \end{aligned}$$

, which is true as the left hand side is strictly positive while the right hand side is strictly negative.

(b) There exist  $\varepsilon_i$  such that  $n_i < F(\varepsilon) < n_T$ . Those use public transportation in equilibrium  $n_T$  and car in equilibrium  $n_i$ . We want to show that they are better off in equilibrium  $n_T$ . We know by preference revelation that  $\forall \varepsilon_i \in [n_i, n_T]$ :

$$W(n_T) + t^{pt}(\alpha, n_T) + \varepsilon_i < f_c + t^c(\alpha, n_T) \quad (11)$$

and

$$W(n_i) + t^{pt}(\alpha, n_i) + \varepsilon_i > f_c + t^c(\alpha, n_i) \quad (12)$$

We want to show that:

$$W(n_T) + t^{pt}(\alpha, n_T) + \varepsilon_i < f_c + t^c(\alpha, n_i)$$

Which, as we know by definition that

$$f_c + t^c(\alpha, n_i) > f_c + t^c(\alpha, n_T)$$

comes directly from (1).

### A.3 Proof of proposition 4

By lemmas 3, cars are better off under **separation** then under **taxation** iff

$$t^c(\alpha, z) - t^c(\alpha', z^\alpha) > [t^c(\alpha, z) - t^c(\alpha, z^T)] - (1 - z^T)T$$

Using 9 and separating the right-hand side,

$$\begin{aligned} [t^c(\alpha, z) - t^c(\alpha, z^\alpha)] - [t^c(\alpha', z^\alpha) - t^c(\alpha, z^\alpha)] &> [t^c(\alpha, z) - t^c(\alpha, z^\alpha)] - [t^c(\alpha, z^T) - t^c(\alpha, z^\alpha)] - (1 - z^T)T \\ (1 - z^T)T &> [t^c(\alpha', z^\alpha) - t^c(\alpha, z^\alpha)] + [t^c(\alpha, z^\alpha) - t^c(\alpha, z^T)] \end{aligned} \quad (13)$$

Assume a shock  $\xi$  and an initial equilibrium  $z$ . We see immediatly that:

$$(1 - z^T)\xi > (1 - z)\xi$$

And, by lemma 1

$$[t^c(\alpha', z^\alpha) - t^c(\alpha, z^\alpha)] < [t^c(\alpha', z) - t^c(\alpha, z)]$$

(we know by assumption that the right hand side of this equation is a constant share of  $\xi$ )

$$[t^c(\alpha, z^\alpha) - t^c(\alpha, z^T)]$$

is positive by lemma 2 and increasing in  $z$  by lemma 1

The left-hand side of 13 is increasing in  $\xi$  while the right-hand side is decreasing in  $\xi$ . Hence, if there exist a  $\xi$  such that inequality 13 is true, then it is true for any  $\xi' > \xi$

### A.4 Proof of Lemma (2)

We have

$$\xi = (U_i^{pt} - U_i^{pt}) - (U_i^{tc} - U_i^c)$$

which implies, in the case of **taxation**

$$\xi(0, T) = zT + (1 - z)T = T \quad (14)$$

while it implies, in the case of **separation**

$$\begin{aligned}\xi(\alpha', 0) &= (t^{pt}(\alpha, z) - t^{pt}(\alpha', z)) - (t^c(\alpha, z) - t^c(\alpha', z)) \\ &= \Delta_1(\alpha', z) - \Delta_1(\alpha, z)\end{aligned}\tag{15}$$

Since equation (14) while, by lemma (1), equation (15) is strictly decreasing in  $z$ . Hence, for any  $\xi$ , the equilibrium level of car users after the implementation of any of these two policies is such that  $z^\alpha > z^T$ .

### A.5 Proof of Lemma (3)

Let divide the population into three families: those who use PT before and after the implementation of the new policy (PT-PT), those who use their car before and after the policy (C-C) and the swing commuters, those who used their car before the policy and PT afterward. We show the effect of both policy for the different families described above.

In the case of taxation

1. PT-PT: the variation of their welfare is given by

$$U^{pt'} - U^{pt} = W(\bar{n}) - W(n^T) + t^{PT}(\alpha, \bar{z}) - t^{PT}(\alpha, z^T) + z^T T$$

which can be decomposed into three effects, all being welfare enhancing (as long as  $z^T < \bar{z}$ ):  $W(\bar{n}) - W(n^T)$  corresponds to the reduction of waiting time;  $t^{PT}(\alpha, \bar{z}) - t^{PT}(\alpha, z^T)$  comes from the reduction of congestion; and  $z^T T$  comes from the lumps sum transfer from CU to the user of PT.

2. C-C: the policy of taxation increases their welfare if

$$U^{c'} - U^c = t^c(\alpha, \bar{z}) - t^c(\alpha, z^T) - (1 - z^T) T > 0$$

i.e. it increases their welfare if

$$t^c(\alpha, \bar{z}) - t^c(\alpha, z^T) > (1 - z^T) T$$

3. C-PT: Comparing  $U^{pt'}$  to  $U^c$  leads to an expression which is difficult to handle. We know, by the revealed preferences property, that if the final swing commuter increases his welfare after the implementation of a policy, all the other swing commuters (those switching from CU to PT) also increase their welfare. Since, by definition, the final swing commuter is such that  $U^{pt'} = U^{c'}$ , we can compare  $U^{c'}$  to  $U^c$  for the final swing commuter and show that if the policy is welfare enhancing for the car users, so it is for the swing commuter and, by revealed preferences, so it is for all commuters.

In the case of traffic separation:

1. PT-PT: the variation of their welfare is given by

$$U^{pt'} - U^{pt} = W(\bar{n}) - W(n^\alpha) + t^{PT}(\alpha, \bar{z}) - t^{PT}(\alpha', z^\alpha)$$

which can be decomposed into two effects, both being welfare enhancing (as long as  $z^\alpha < \bar{z}$ ):  $W(\bar{n}) - W(n^\alpha)$  corresponds to the reduction of waiting time and  $t^{PT}(\alpha, \bar{z}) - t^{PT}(\alpha', z^\alpha)$  comes from the reduction of congestion due to two forces: (i) less car and (ii) more traffic lines devoted to PT only.

2. C-C: the policy of separation increases their welfare if

$$U^{c'} - U^c = t^c(\alpha, \bar{z}) - t^c(\alpha', z^\alpha) > 0$$

i.e. it increases their welfare if

$$t^c(\alpha, \bar{z}) > t^c(\alpha', z^\alpha)$$

3. C-PT: Comparing  $U^{pt'}$  to  $U^c$  leads to an expression which is difficult to handle. We know, by the revealed preferences property, that if the final swing commuter increases his welfare after the implementation of a policy, all the other swing commuters (those switching from CU to PT) also increase their welfare. Since, by definition, the final swing commuter is



such that  $U^{p^t} = U^{c'}$ , we can compare  $U^{c'}$  to  $U^c$  for the final swing commuter and show that if the policy is welfare enhancing for the car users, so it is for the swing commuter and, by revealed preferences, so it is for all commuters.

## B Figures

The shape of  $t^c(\alpha, z)$ ,  $t^{pt}(\alpha, z)$  and  $\Delta_1(\alpha, z)$

