FACULTY OF LAW, ECONOMICS AND FINANCE

CREA
Discussion
Paper
2012-03

Center for Research in Economic Analysis
University of Luxembourg

Modal choice and optimal congestion

available online : http://wwwfr.uni.lu/recherche/fdef/crea/publications2/discussion_papers/2011

Quentin David, CREA, University of Luxembourg
Renaud Foucart, ECARES, Université libre de Bruxelles
.

May, 2012

For editorial correspondence, please contact: crea@uni.lu
University of Luxembourg
Faculty of Law, Economics and Finance
162A, avenue de la Faïencerie
L-1511 Luxembourg

Modal choice and optimal congestion*

Quentin David[†], CREA, Université du Luxembourg.

Renaud Foucart[‡], ECARES, Université Libre de Bruxelles.

May 7, 2012

Abstract

We study the choice of transportation modes within a city where commuters have heterogeneous preferences for a car. As in standard models of externalities, the market outcome never maximizes aggregate welfare. We show that in the presence of multiple equilibria problems of coordination can worsen this result. Hence, a social planner focusing on the marginal impact of policies may miss the largest source of inefficiency. We discuss two policy tools: taxation and traffic separation (e.g. exclusive lanes for public transportation). Setting the optimal levels of taxation and of traffic separation constitutes a necessary but not a sufficient condition to reach the first best equilibrium. Comparing the relative efficiency of both policies, we show that traffic separation should be preferred for large-scale policies while taxation better applies to marginal modifications of commuting patterns.

JEL: R4, L5, H2.
Keywords: Modal choice, Coordination, Network effect, Cross-modal congestion.

1 Introduction

The cost of congestion is an increasingly important issue in urban areas. For instance, Duranton and Turner (2011) estimate that a typical American household spends 161 person-minutes in a car every day. Goodwin (2004) expected the annual cost of congestion in the UK to reach 30 billion £ in 2010. De Palma and Lindsey (2011) report congestion costs between 0.5 and 1.5 % of GDP in urban areas. Most of the congestion is due to the use of private cars. On the one hand, cars generate both congestion - on other cars and on public transportation - and pollution. On the other hand, cars are necessary for the economy. Unfortunately, screening commuters to

*We would like to thank Rabah Amir, Marco Batarce, Paul Belleflamme, Micael Castanheira, Marjorie Gassner, Marc Ivaldi, Moëz Kilani, Martin Koning, Frederic Malherbe, Pierre M. Picard, Jacques François Thisse, Yves Zenou and various conference and seminar participants in Brussels, Luxembourg, Marseille and Istanbul for their helpful comments and suggestions. Quentin David thanks the FNR for financial support.

[†]Contact: quentin.max.david@gmail.com

[‡]Contact: rfoucart@ulb.ac.be

reach the optimal share of cars is a complex policy challenge. This is why policies must identify tools that affect peoples' behavior and improve efficiency. In practice, the most widely-used policies addressing traffic issues are taxation,[1] subsidies and traffic separation.[2]

In this paper, we build a theoretical model in which heterogeneous commuters decide simultaneously whether to use a private car or public transportation. Car users generate congestion on all the commuters and users of public transportation enjoy a positive network externality.[3] We do not specifically model pollution costs, as this externality affects all commuters regardless of their modal choice, and therefore does not affect this decision. In practice, considering the impact of pollution would lower the socially optimal share of car users obtained with our model.

First, we show that the market outcome never maximizes aggregate welfare. This is a classic result, in presence of externalities. Second, we explain how ex-ante similar cities might end up with very different modal shifts. This is a problem of coordination when a large share of commuters have similar preferences. In the presence of such multiple equilibria, the one involving the highest share of public transportation always Pareto dominates all the others. Therefore, the market presents two types of inefficiencies. The first one is at the margin: the market provides a too large share of car users in any decentralized equilibrium. The second is more substantial: coordination failures may lead to the presence of inefficient equilibria.

We study two policies: taxation (which, in our discrete choice setup is equivalent to a fare subsidy) and traffic separation. Both can be used to enforce coordination. We show that the main problem of taxation is when the number of car users is small. Shrinking the tax base can be very detrimental for the remaining car users. The main problem of traffic separation is when the increase in congestion costs outweighs the benefits resulting from the decrease of the share of car users. This happens when the share of car users is high in equilibrium. In practice, a

[1] See de Palma and Lindsey (2011) for a survey of the different methods and impacts of congestion tolls.

[2] For instance through the use of exclusive lanes for public transportation. See Cain et al. (2006) and Echeverry et al. (2005) for larger discussion of the case of Bogota's Transmilenio and its application to other countries.

[3] We show in Extension 5.1 how introducing discomfort externalities increases the likelihood of ending up with multiple equilibria.

social planner considering marginal changes in commuting patterns should focus on taxation, while a social planner interested in more substantial changes should focus on traffic separation.

The question of optimal congestion has been addressed by many scholars from different fields. Among economists, it is rather consensual that pigouvian taxation should be the preferred way to deal with congestion problems (Beesley and Kemp, 1987, Calfee and Winston, 1998). The idea is that, given both the structure of cities and the intrinsic preferences that many consumers have for the car, one should focus on the best way to accommodate traffic flows and make car user pay for the marginal external cost they produce (Anas and Small, 1998). The 'games of congestion' have been largely studied in economic theory (Rosenthal, 1973) and many applied papers 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. The question of public transportation has often been of minor interest though some authors (e.g. Mirabel, 1999, Dobruzkes and Fourneau, 2007) addressed the so-called 'crossed modal externalities' (the impact of the congestion generated by one mode of transportation on the other one). The congestion costs have been shown to be convex both in terms of pollution (De Vlieger et al., 2000) and in terms of perceived cost (Wardman, 2001).[4]

Another group of papers focuses on urban planning. It emphasizes the fact that the structure of the city is the main driver of commuting patterns. The main idea to improve the performance of urban transportation is to have a shift towards 'transit-oriented development'. Belzer and Aultier (2002) define such a development as follows: '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'.[5] Some economists indirectly address this dimension by considering a form of traffic separation (see Berglas et al., 1984, Arnott et al., 1992, de Palma

[4]Time is valued 50% higher when spent in congestion. Hence, the cost of congestion is convex, as congestion (i) increases travel time, and (ii) increases the marginal cost of travel time. This principle is applied by Santos and Bhakar (2005) to assess the benefits of the congestion toll in London.

[5]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 the literature and 'good practices' in transit oriented development.

and Lindsey, 2002, de Palma et al. 2008). These papers propose various approaches for road pricing and tolls in the presence of alternative roads, modes of transportation and consumer preferences. Their settings differ from our in various dimensions, but all have in common to find a unique equilibrium and an optimal policy, corresponding to the idea of pricing the marginal externality. In this paper, we show that in the presence of multiple equilibria, internalizing marginal externalities may not be efficient. Then, some physical planning (traffic separation, in our model) must be used as a coordination device.

The multiple equilibria come from the conjunction of congestion, positive externalities from public transportation, and commuters' heterogeneity. A relatively large literature exists on the network effect of the number of transit users on the efficiency of public transportation. In a seminal contribution, 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 underlying 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 or buses 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. The combination of network externalities in public transportation and congestion by cars is a feature of several economic models (Tabuchi, 1993, Parry and Small, 2009). To repeat, those models focus on a unique equilibrium. Commuters differ in their preference for the use of a private car (Beirão and Cabral, 2007, Handy et al., 2005, Jensen, 1999, Steg, 2005, Hiscock et al., 2002 and Van Vught et al., 1996). Berhoef and Small (2004) encompass this dimension by considering heterogeneous agents in a model of pricing for car use only. Batarce and Ivaldi (2011) test this feature in a model of modal choice applied to Santiago, Chile.

The existence of similar cities characterized by different modal shifts has already been documented in the late eighties by Pucher (1988). He observed that 'Urban transportation and traveler behavior vary widely, even among countries with similar per capita income, technology and urbanization'. Kenworthy and Laube (1999) show that the fraction of workers using transit is 6 times higher in wealthy Asian cities compared to the US.[6] They also find that the commuting time is lower (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. Cities where transit is intensively used appear to need a smaller share of subsidies for operating it.

The paper is organized as follows. In the next section, we provide an illustrative example that contains the basic intuitions behind our main results. Section 3 presents the model, shows that a Nash Equilibrium always exists, gives conditions for the existence of multiple equilibria and discusses their relative efficiency. In Section 4, we derive the optimum of the social planner and study taxation and traffic separation. We extend the model in section 5, addressing the possible existence of capacity constraints and congestion within public transportation, and considering the possibility of underground transit. We conclude in Section 6.

2 An illustrative example

Suppose a continuum of commuters simultaneously choose between using a car or public transportation in order to minimize the cost associated to their modal choice. These costs (T^c and T^{pt}) are respectively given by:

$$T^c = t + f$$

and

$$T^{pt} = t + W + \varepsilon,$$

where t denotes the time spent in congestion. In the benchmark case, it is identical for both modal choices and we assume that $t = 0$ if less than 50% of the population uses public transportation and $t = 1$ otherwise. f is the fixed cost associated with the use of a car, and is set at 1. W is a cost associated with the use of public transportation that is characterized by a network externality such that if less than 50% of the population uses public transportation,

[6] Similarly, Pucher and Renne (2003) computed that, in the US, public transport accounted for less than 2% of urban travel in 2001.

$W = 2$, otherwise, $W = 1/2$. Finally, ε represents the value of an intrinsic preference for the use of a car, compared to public transportation.

Consider three groups (A, B and C) of commuters of equal size (each group represents 1/3 of the population), characterized by different levels of ε. Group A displays strong preferences for the use of a car ($\varepsilon = 2$), group B is indifferent ($\varepsilon = 0$) and group C prefers public transportation ($\varepsilon = -2$).

We define z as the equilibrium share of the population using a car and $\hat{z}$ as the beliefs over the outcome z of the game.

Depending on the expectations, the outcome of the game for an individual belonging to one of the groups is

	if $\hat{z} = \frac{2}{3}$		if $\hat{z} = \frac{1}{3}$	
	T^C	T^{pt}	T^C	T^{pt}
Group A	**2**	5	**1**	2.5
Group B	**2**	3	1	**0.5**
Group C	2	**1**	1	**−1.5**

There exist two Nash Equilibria in Pure Strategy. Group A always uses a car, group C always uses public transportation and group B uses a car if its members believe that the other commuters in the group will do so and public transportation otherwise. With the same exogeneous set of parameters, one may end up either in a world where a majority of people commute either by car or by public transportation. However, the latter equilibrium Pareto dominates the former. We refer to these two equilibria as the "good" and the "bad" equilibrium when $z = 1/3$ and $z = 2/3$, respectively.

We now study the ability of two policies (taxation and traffic separation) to avoid the "bad" equilibrium.

Policy 1: taxation

Consider the simplest taxation scheme: a tax is levied on car users and is thrown away. To ensure that group B uses public transportation in the presence of the tax, its level must be set in such a way that $2 + T > 3$. As a consequence, we remain with the "good" equilibrium only but, this equilibrium is no longer Pareto improving as the group A is worse off with this tax (the cost of taxation outweighs the decrease in congestion).

Policy 2: physical planning

Consider now a policy of traffic separation. This policy consists in separating the traffic lanes for cars and for public transportation. Under the assumption of a fixed number of traffic lanes, this implies an increase of the congestion for car users and a decrease for users of public transportation. Assume that, in the presence of congestion, the cost of congestion for cars is doubled ($t^c = 2$ if $z > 50\%$, $t^c = 0$ otherwise) while it is divided by two for public transportation ($t^{pt} = 1/2$ if $z > 50\%$, $t^c = 0$ otherwise). The 'bad' equilibrium disappears again and the game becomes

	if $\hat{z} = \frac{2}{3}$		if $\hat{z} = \frac{1}{3}$	
	T^C	T^{pt}	T^C	T^{pt}
Group A	**3**	4.5	**1**	2.5
Group B	3	**2.5**	1	**0.5**
Group C	3	**0.5**	1	**−1.5**

This policy is Pareto Improving, since every commuter is better off in the equilibrium with policy than in the "bad" equilibrium without policy.

3 The model

3.1 Basic assumptions

We consider a closed city with a unit mass of commuters who have to make a discrete choice between using a private car or public transportation. The use of a car generates congestion on the other commuters. The space is finite and it is possible to increase neither the number of roads nor the number of the traffic lanes. The degree of separation of public transportation from the rest of the traffic is given by $\alpha \in [0, 1]$.[7] Commuters are heterogeneous as they have different

[7] α is exogenous in this section, but we allow the social planner to choose its level in the next section. Note also that we use a very general definition of α, one that encompasses many possibilities to protect public transportation. The condition being that increasing α decreases congestion for public transportation and increases congestion for car users. This excludes the possibility of building an underground (which is briefly discussed in section 5).

intrinsic preferences for the use of a car (relative to public transportation).[8] The outcome of the game is a share z of car users, and $(1-z)$ of public transportation users.

The utility[9] of a commuter i, traveling in a private car or with public transportation, is respectively given by

$$U_i^c(\alpha, z) = -f_c - t^c(\alpha, z) + \frac{\varepsilon_i}{2} \tag{1}$$

and

$$U_i^{pt}(\alpha, z) = -W(z) - t^{pt}(\alpha, z) - \frac{\varepsilon_i}{2}. \tag{2}$$

The fixed cost associated with the use of the car is denoted by $f_c > 0$.[10]

The functions $t^c(\alpha, z)$ and $t^{pt}(\alpha, z)$ $(\in IR^+)$ represent the congestion faced respectively by cars and public transportation. They are assumed to be equal if there is no traffic separation between cars and public transportation, and equal to zero if there are no users of cars (i.e. $t^c(0, z) = t^{pt}(0, z)$ and $t^c(\alpha, 0) = t^{pt}(\alpha, 0) = 0$ respectively). Both functions are increasing and convex in z, and a higher degree of traffic separation (higher α) generates more congestion for cars (because there is less space for them) and less congestion for public transportation. This last effect is assumed to be amplified by z (that is, separation has an impact only if there is actually a problem of congestion). Hence, we have

$$\begin{aligned}
\frac{\partial t^c}{\partial z}(\alpha, z) &> 0,\ \frac{\partial^2 t^c}{\partial z^2}(\alpha, z) \geqslant 0,\ \frac{\partial t^c}{\partial \alpha}(\alpha, z) > 0,\ \frac{\partial^2 t^c}{\partial z \partial \alpha}(\alpha, z) > 0 \\
\frac{\partial t^{pt}}{\partial z}(\alpha, z) &> 0,\ \frac{\partial^2 t^{pt}}{\partial z^2}(\alpha, z) \geqslant 0,\ \frac{\partial t^{pt}}{\partial \alpha}(\alpha, z) < 0,\ \frac{\partial^2 t^{pt}}{\partial z \partial \alpha}(\alpha, z) < 0.
\end{aligned}$$

The individual parameter, ε_i, is the preference for the use of a car, compared to public transportation. It comes from a cumulative distribution function $\varepsilon_i \sim F(\varepsilon)$. F is assumed to be

[8] This preference can be negative. One can imagine various alternative ways of modelling heterogeneity: different valuation for time and money, different location within the city, ease of access to the public transportation network, etc. We use the simplest formulation for the tractability of the model.

[9] Utility functions are expressed in monetary terms. All components (fixed costs, congestion, individuals' heterogeneity and waiting time) are expressed in monetary terms.

[10] This is the additional cost compared to the use of public transportation, which is normalized to 0.

continuous and differentiable over its support $(-\infty, +\infty)$. This support implies that some individuals love public transportation so much that they would never accept not to use it ($\varepsilon_i \to -\infty$), while others will never use public transportation ($\varepsilon_i \to +\infty$). Without loss of generality, we split ε_i equally between the two utility functions.

The waiting time for public transportation is $W(z) \in \mathbb{R}_0^+$. It displays a positive network externality for public transportation users. The idea is that, if there are more users, the frequency of public transportation increases and the waiting time decreases.[11] For simplicity, we assume this network externality to be linear. If there are $(1-z)$ users of public transportation, the waiting time of each of them is given by $W(z)$, with

$$W'(z) > 0 \text{ and } W''(z) = 0.$$

Further on in the paper, it will be useful to define $G(x) = F^{-1}(\varepsilon)\ \forall \varepsilon \in IR$. Given the assumptions over $F(\varepsilon)$, the support of G is $x \in [0,1]$.[12]

Definition 1 $\Delta(\alpha, z)$ *is the additional congestion faced by car users in comparison to the congestion faced by public transportation, i.e.*

$$\Delta(\alpha, z) = t^c(\alpha, z) - t^{pt}(\alpha, z).$$

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

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

(i) $\frac{\partial \Delta(\alpha,z)}{\partial \alpha} > 0$;

(ii) $\frac{\partial \Delta(\alpha,z)}{\partial z} > 0$, $\forall \alpha > 0$;

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

[11] An alternative interpretation: a lower price for a given quality of service.

[12] With $G(0) = +\infty$, $G(1) = -\infty$ and $G(x) = \varepsilon_i$ such that there is a mass x of commuters with $\varepsilon > \varepsilon_i$.

Proof. (i) and (ii) are straightforward from the properties of $t^c(\alpha, z)$ and $t^{pt}(\alpha, z)$. Property (iii), the supermodularity of $\Delta(\alpha, z)$ is obtained using $\frac{\partial^2 t^c(\alpha,z)}{\partial z \partial \alpha} > 0$ and $\frac{\partial^2 t^c(\alpha,z)}{\partial z \partial \alpha} < 0$. The definition of $\Delta(\alpha, z) = t^c(\alpha, z) - t^{pt}(\alpha, z)$ leads to:

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

■

3.2 The game

The modal choice is a simultaneous game among a unit mass of commuters. It consists in each commuter choosing the mode of transportation (either car of public transportation) that maximizes her utility given her expectation on z. Hence, commuter i commutes by car if $U_i^c(\alpha, z) > U_i^{pt}(\alpha, z)$, i.e.

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

If it is a best response *ex post* for a commuter j with $\varepsilon_j > \varepsilon_i$ to commute using public transportation, it is also a best response for commuter i to do so.

Using Definition (1), the condition for commuter i to use a car becomes

$$\varepsilon_i > f_c - W(z) + \Delta(\alpha, z). \tag{3}$$

3.3 Decentralized Equilibria

In this section, we first show that a Nash equilibrium always exists. Second, we derive the conditions for the presence of multiple equilibria. Third, we characterize the most efficient one.

3.3.1 Existence

The existence of at least one Nash equilibrium is relatively easy to show. Stability *ex post* comes from the fact that there always exists an equilibrium where a share of commuters strictly prefers public transportation while the other prefers to use a car.

Proposition 1 *There exists at least one pure strategy Nash equilibrium.*

Proof. Remember $F(\varepsilon)$ is assumed to be continuous and differentiable over its support $(-\infty, +\infty)$. This implies that there exists at least one commuter k with taste parameter ε_k such that, if all commuters with parameter $\varepsilon_j < \varepsilon_k$ use public transportation, and all commuters with $\varepsilon_k < \varepsilon_i$ take the car,

$$G(z_k) = f_c - W(z_k) + \Delta(\alpha, z_k). \tag{4}$$

Commuter k is indifferent between the private car and public transportation. Sharing the same beliefs, commuters with $\varepsilon_j < \varepsilon_k$ strictly prefer public transportation and $\varepsilon_k < \varepsilon_i$ strictly prefer their car. Thus, it is a Nash Equilibrium. ■

3.3.2 Multiplicity

The intuition behind the existence of multiple equilibria is the following. Assume that there is a large share of commuters with similar preferences (ε) for the use of a car. When they believe that most of them use public transportation, it is a best response for them to do so. This is a Nash equilibrium with a low z. If, on the contrary, most of them believe that they will use a car, they expect public transportation not to be efficient and, indeed, it will not be. This is also a Nash equilibrium, involving a high z.

Proposition 2 *There exist multiple equilibria if and only if there exists a solution z_k such that*

$$\frac{\partial G(z_k)}{\partial z} > \frac{\partial [f_c - W(z_k) + \Delta(\alpha, z)]}{\partial z}. \tag{5}$$

Proof. The proof is presented in Appendix A.1. ■

For this condition to be fulfilled, the difference in the costs between the two modes of transportation must be sufficiently low and a sufficiently high mass of commuters must have similar preferences. Consider the particular case of unimodal preferences: few people with polarized preferences, and a large fraction of people with similar preferences. This is likely to lead to the presence of three equilibria, as plotted in Figure 1 (with $\alpha = 0$). There are two stable[13] equilibria, one with few users of public transportation (a share z_1 of car users) and one

[13] Those equilibria are locally stable in the sense that agents' best-response to any small perturbation to the equilibrium z would bring this share back to equilibrium.

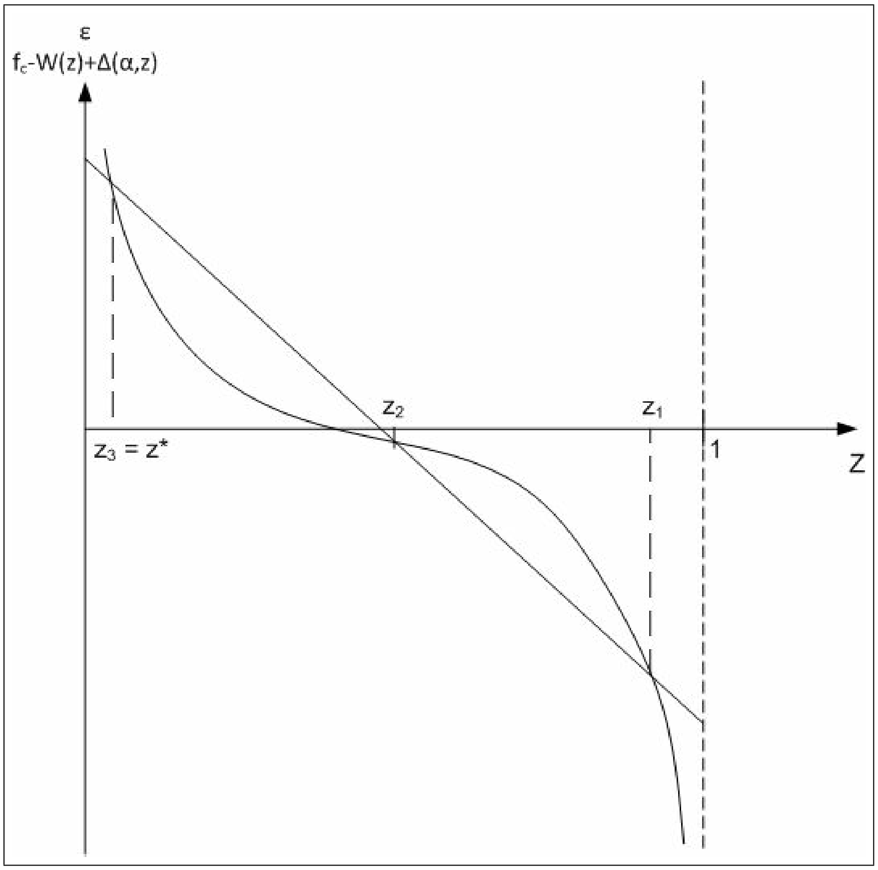

Figure 1: Illustration with multiple equilibria

with a large fraction (a share $z_3 < z_1$ of car users). There is also one unstable equilibrium, z_2.

3.3.3 Efficiency

Proposition 3 *If there are multiple equilibria, the equilibrium involving the higher use of public transportation Pareto dominates all the other equilibria. The Pareto dominant equilibrium is denoted $\widehat{z}$.*

Proof. The formal proof is provided in Appendix A.2. ■

Figure 1 illustrates this proposition. Define three groups of people A, B and C. Group A uses a car in both equilibria, group C uses public transportation in both equilibria, and group B uses public transportation when $z = z_3$ and a car otherwise. As the cost of both public transportation and car use are lower in z_3, groups A and C are strictly better off. By revealed preferences, group

B is also better off in that equilibrium. They are better off by using a car in z_3 than in z_1, but they use public transportation instead.

4 Social planner

In this section, we stress the possible presence of two different sources of inefficiencies. The *coordination failure* implies a move from one type of equilibrium to another (e.g. z_1 to z_3 in the previous section). The *sub-utilization of public transportation* implies that, at the margin, the first best equilibrium requires more users of public transportation at any initial Nash equilibrium. This latter result is standard in the presence of externalities.

First, we derive the first best equilibrium conditions for a social planner maximizing the aggregate utilities. We show that this social planner may miss the coordination failure by focusing on local maximization. Second, we study the effect of two policies, taxation and traffic separation at different initial Nash equilibria. Third, we study the conditions for these policies to be Pareto improving and compare their relative efficiency.

4.1 The social planner's optimum

Assume that the aim of the social planner is to maximize the sum of all commuters' utilities, i.e.

$$\begin{aligned} \underset{z,\alpha}{Max} \int_0^1 \left[\phi U_i^c(\alpha, z) + (1-\phi) U_i^{pt}(\alpha, z)\right] dx \\ \text{such that } \phi &= 1 \text{ if } \varepsilon_i \geq G(z) \\ &= 0 \text{ otherwise} \\ \alpha, z &\in [0,1] \end{aligned}$$

This is equivalent to

$$\min_{z,\alpha} \int_0^z \left[f_c - t^c(\alpha, z) + \frac{G(x)}{2}\right] dx + \int_z^1 \left[W(z) - t^{PT}(\alpha, z) - \frac{G(x)}{2}\right] dx$$

and the first order conditions are

$$G(z^*) = f_c - W(z^*) + \Delta(\alpha, z^*) + z t_z^c(\alpha, z^*) + (1 - z^*) \left[W'(z^*) + t_z^{PT}(\alpha, z^*)\right] \tag{6}$$

and

$$z t_\alpha^c(\alpha, z) + (1 - z) t_\alpha^{PT}(\alpha, z) = 0. \tag{7}$$

Rearranging the terms to compare the private costs and the public benefits for commuter $i : G(z^*) = \varepsilon_i$, equation (6) leads to the following condition:

$$W(z^*) - \Delta(\alpha, z^*) - f_c + G(z^*) = z^* t_z^c(\alpha, z^*) + (1 - z^*) \left[W'(z^*) + t_z^{PT}(\alpha, z^*)\right]. \tag{8}$$

Thus, we have:

Lemma 2 *In any Nash Equilibrium, the share of car users is too high. For the share of car users to be socially optimal, there must exist public transportation users that strictly prefer the car.*

Proof. The right-hand side of equation (8) corresponds to the social cost of increasing the share of car users. This is clearly positive, as all negative externalities of a car are increasing with z. On the left hand side is the individual preference for the car of the swing commuter z^*, such that all commuters with $\varepsilon < G(z^*)$ take public transportation, and the other use a car. For the equality to hold, this must be positive. This implies that the socially optimal swing commuter strictly prefers to use a car rather than public transportation. ■

If the second FOC leads to an interior solution, $\alpha^* \in (0, 1)$, it becomes

$$\frac{z}{(1 - z)} = -\frac{t_\alpha^{PT}(\alpha^*, z)}{t_\alpha^c(\alpha^*, z)}. \tag{9}$$

The right-hand side of equation (9) is positive and represents a measure of the relative efficiency of a traffic separation policy, i.e. the marginal effect of α on the relative commuting time ratio. Defining $\beta(\alpha, z) = -\frac{t_\alpha^{PT}(\alpha^*, z)}{t_\alpha^c(\alpha^*, z)}$, it is reasonable to believe that $\beta_\alpha(\alpha, z) \geq 0 \ \forall \ z \in [0, 1]$.

Indeed, on the one hand, by increasing the share of roads dedicated to public transportation, the incidence of congestion on public transportation is reduced proportionally. On the other hand, the effect on cars is likely to be different. The creation of dissociated traffic lanes for public transportation generates bottlenecks for cars. The creation of these bottlenecks is likely to increase congestion but at a marginally decreasing rate (by increasing the number of bottlenecks, the impact of each one is reduced).

If the second FOC does not yield an interior solution (i.e. if $\beta_\alpha(\alpha, z) = 0$ or $\frac{z}{(1-z)} \neq \beta(\alpha, z)$ $\forall \alpha \in [0,1]$), the social planner will choose $\alpha = 0$ if z is sufficiently high, and $\alpha = 1$ if z is sufficiently small. If there is no interior solution and there is a large share of car users, it is socially beneficial - at the margin - to allow more space for cars. If there is a large share of public transportation users, it is socially beneficial to fully protect public transportation from congestion.

It is important to underline that equations (8) and (9) give the conditions to reach a local maximum, not necessarily a global one. As for the decentralized equilibria, there is no reason for these optima to be unique.

Therefore, these conditions give an insight into what can be socially optimal to solve the *sub-utilization of public transportation* at the margin: in any decentralized state of the world, there are not enough users of public transportation. This is illustrated in Figure (2) where we plot the functions on the left and the right hand sides of equation (6) together with equation (4), the decentralized equilibrium condition. Compared to the decentralized equilibrium, the right hand side of the equation is associated to a higher intercept and slope, for any value of α, i.e. the difference between the two curves is increasing in z.

A social planner could be misled when he tries to reach a maximum using equations (8) and (9). The planner may be missing a larger inefficiency: coordination failure. Indeed, consider a Nash equilibrium that implies a high share of cars, z_k, and assume a taxation scheme able to internalize the marginal externality. If the social planner maximizes the aggregate utility

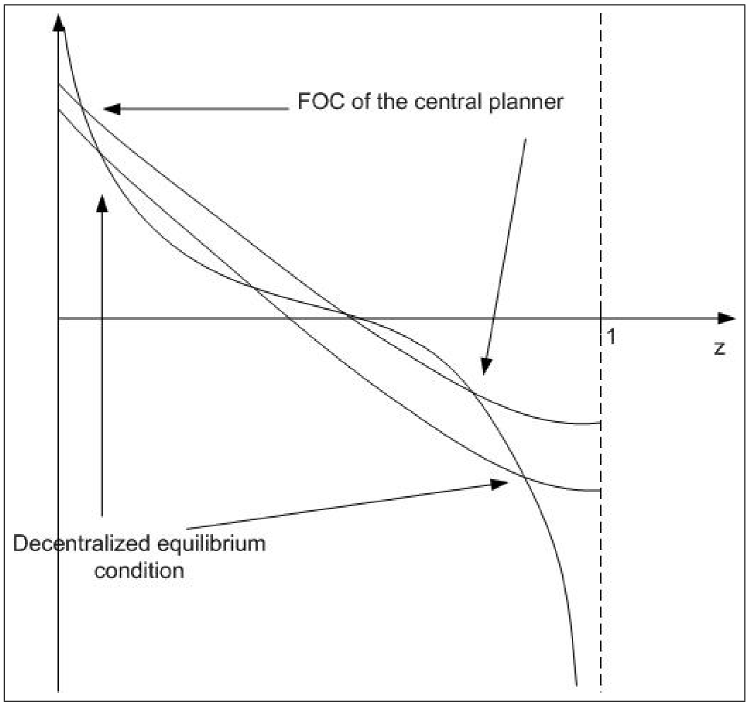

Figure 2: Decentralized equilibrium and social planner's first order condition for z^* ($\alpha > 0$)

by setting $\alpha = 0$ for this value of z_k, there is a possibility that there exists another (lower) value, $\widehat{z} \neq z_k$, another Nash equilibrium, with another optimal value of taxes and $\alpha > 0$. In other words, a social planner must ensure not to target a local maximum when a better, global maximum is reachable.

4.2 Policy tools

In the previous sections, we showed that two different kinds of inefficiencies have to be distinguished: coordination failure and sub-utilization of public transportation. The first one comes from poor coordination between individuals in the presence of multiple equilibria, when the prevailing equilibrium does not involve the best use of public transportation. The second is due to the two considered externalities (congestion and network externalities) leading to the sub-optimal use of public transportation, whatever the prevailing equilibrium. To address the first inefficiency properly, it is required to significantly change individuals' behavior. To address the second one, a central planner should affect the behavior of some marginal individuals only: those that are the most likely to use public transportation after the implementation of new policies. In the presence of a unique equilibrium, the only relevant conditions are given by equations (8) and (9). Otherwise, policies may also have a role by ensuring coordination towards the most efficient equilibrium.

It is worth noting that, in our setting, a government could set a very high tax and then remove it almost instantaneously in order to force commuters to coordinate on the efficient equilibrium. Nevertheless, we believe that this is not realistic. The dynamics of switching from one equilibrium to another is a long, progressive process. To solve this issue, we assume that setting a policy implies that the government will keep it in place forever. This assumption may be considered ad hoc, but it is realistic for an intervention to take effect and be credible.

We assume that the government has two policy tools at its disposal: the **taxation** of car users (T), and the possibility to change the traffic **separation** between cars and public transportation (α). Due to the discrete choice nature of the model, tax is equivalent to a fare subsidy.[14] We do not consider variations of taxation schemes that can have differential effects among car users or time of the day.[15]

We assume that **a taxation policy** implies to levy T on every car user and that this tax is redistributed lump-sum among all commuters.[16] Therefore, every commuter receives a transfer zT and car users pay T. The new utility functions become

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

After the introduction of a taxation policy, a commuter i uses public transportation if and only

[14] An alternative policy would be to allow the social planner to invest in lower W for a given value of z.

[15] One can refer to Parry (2002) for a comparison between a single lane toll, a uniform congestion tax across freeway lanes, a gasoline tax, and a transit fare subsidy for the reduction of congestion.

[16] As will be made clear below, assuming the tax is lost only marginally affects the results.

if

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

and, by proposition 1, there is always at least one 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 involves a lower share of car users than in $\hat{z}$.

The **traffic separation** is the other available policy tool. Assume that the government sets a new traffic separation, $\alpha' : \alpha' > \alpha$. The new utility functions become

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

Now, a commuter i uses public transportation if

$$\varepsilon_i < \Delta(\alpha', z) - W(z) + f_c.$$

Here, one cannot theoretically claim that the impact of traffic separation is sufficient to keep only one equilibrium. It depends on the size of the effect of traffic separation on the difference in commuting times. The effect of these two policies is presented in Figure 3. The optimal level of traffic separation has been discussed in the previous section.

Comparing equations (6) and (4), an optimal taxation scheme can easily be obtained. If a social planner wants to reach an optimal modal split, z^*, it must set a level of taxation corresponding to the social marginal effect of the use of a car, computed at the targeted optimum, z^*.

For every locally optimal modal split z^*, there exists an optimal level of taxation T^* corresponding to the social marginal impact of the use of a car in z^*, this level of taxation corresponds to the sum of the social marginal congestion for cars and public transportation and of the social marginal opportunity cost in terms of network externality of car users not using public transportation.

$$T(\alpha, z^*) = z^* t_z^c(\alpha, z^*) + (1 - z^*)\left[W'(z^*) + t_z^{PT}(\alpha, z^*)\right]$$

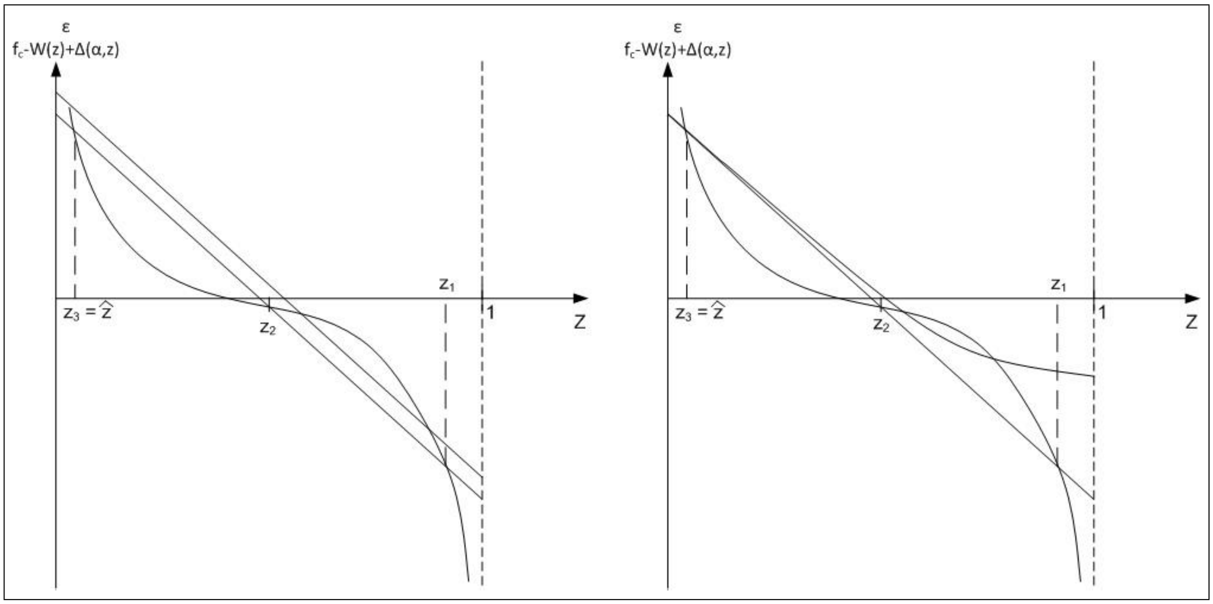

Figure 3: Effect of taxation (left) and traffic separation (right).

It is interesting to note that the optimal tax is increasing in z^* ($T_z(\alpha, z^*) \geq 0$). This means that, comparing two similar cities, if the optimal share of car users is higher in one of the two cities, the level of taxation in that city must also be higher. This result is due to the marginal cost of car use (both in terms of congestion and in terms of network externalities) which is increasing in the share of car users.

This optimal taxation is a necessary condition for the decentralized equilibrium to be optimal. It is not a sufficient condition. Indeed, consider a city where the decentralized equilibrium is located in z_1 and the global optimum slightly to the left of $\widehat{z}$. If a social planner sets a taxation compatible with $\widehat{z}$, it is very likely that the equilibrium in the city does not end up at $\widehat{z}$, but rather somewhere between z_1 and $\widehat{z}$, as shown in Figure 4, where we duplicate the curves of Figure 2 and add the new decentralized equilibrium condition taking the optimal taxation into consideration.

The same approach applies to traffic separation. From the first order condition (7), the optimal traffic separation is decreasing in z, the share of car users. In the presence of a large

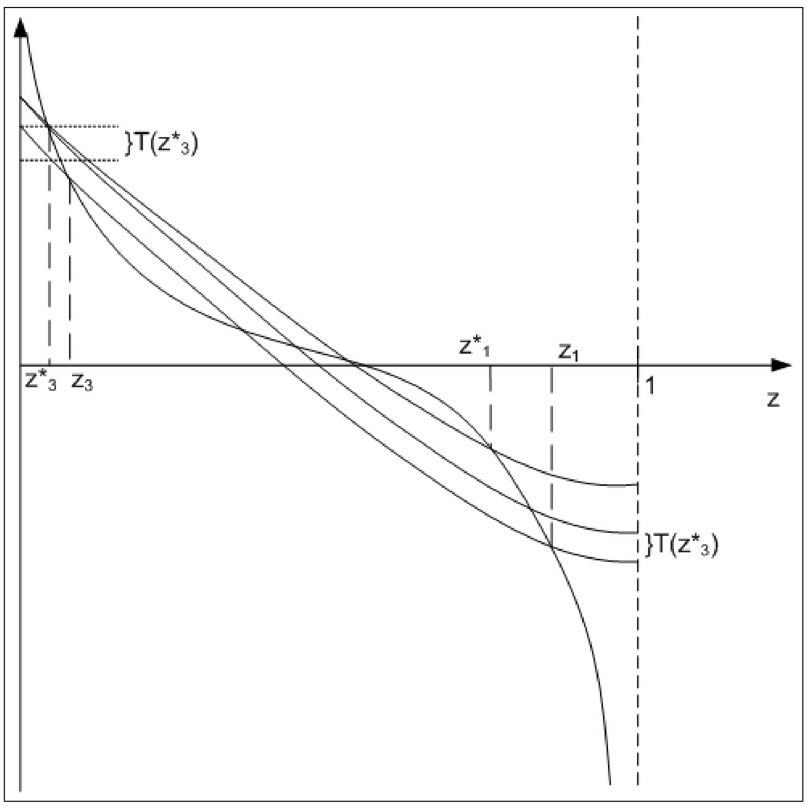

Figure 4: Effect of a taxation computed at z_3^* on the different decentralized equilibria.

share of car users, the marginal impact of α is to dramatically increase the transportation time for the majority of commuters and to decrease it for a minority. Hence, in the global welfare maximum, the traffic separation is very high, but this is not necessarily the case in the local maximum, and setting the level of traffic separation of the global optimum is not a sufficient condition to reach this equilibrium. This can lead to too much traffic separation in an equilibrium where most of the commuters use their car.

4.3 Efficiency of policy tools

Even setting the optimal policy is not a sufficient condition to reach the first best. This is why it is necessary to provide a more general tool to assess policies: efficiency. In this section, we first define the conditions for either taxes or traffic separation to be Pareto-improving. Second, given those conditions, we provide a general result to compare their relative efficiency. If one is not strictly better than the other, then taxation should be preferred for smaller changes of commuting patterns, while traffic separation should be preferred for bigger changes.

4.3.1 Absolute efficiency of policy tools

Definition 2 *For a given Nash equilibrium, z_k, we define the* ***initial swing commuter*** *as the commuter indifferent between using a car or public transportation before the introduction of a policy, and the* ***final swing commuter*** *as the commuter indifferent between using a car or public transportation after the introduction of a policy.*

Definition 3 z^T *and* z^α *are the equilibria after the implementation of, respectively, a policy of taxation and of traffic separation.*

In our setting, the study of policy implementation requires an analysis of its effect on three types of agents. The *initial users of* public transportation (those on the right of the initial swing commuter), *the switching users* (those located between the initial and the final swing commuters), and the *remaining car users* (those on the left of the final swing commuter).

We know, by definition of our externalities, that all the *initial users of public transportation* are better off with the implementation of either policy. Indeed, they enjoy higher network externalities (more users of public transportation), face less congestion (less cars) and, in case of a taxation policy, receive a lump sum transfer. In the case of traffic separation, they enjoy an additional decrease in congestion (since public transportation benefits from a higher share of roads).

Now we study the effect of the policies on *the switching users* and the *remaining car users*. We show that measuring the welfare of the remaining car users is sufficient to assess the Pareto efficiency of these policies. First, we identify the conditions under which the implementation of a policy can be Pareto improving and, second, we compare the effect of those policies on remaining car users.

Lemma 3 *Conditions for the two considered policies to be Pareto improving:*
(i) A policy of ***taxation*** *is Pareto improving if and only if*

$$\left[t^c\left(\alpha, z_k\right) - t^c\left(\alpha, z^T\right)\right] > \left(1 - z^T\right) T. \tag{10}$$

(ii) A policy of ***traffic separation*** *is Pareto improving if and only if*

$$t^c(\alpha, z_k) > t^c(\alpha', z^\alpha). \tag{11}$$

Proof. See Appendix A.3. ■

A policy of taxation is Pareto improving if the reduction of congestion compensates for the cost of taxation. A policy of traffic separation is Pareto improving if it reduces congestion for the *remaining car users*. This implies a trade-off between fewer car users ($z^\alpha < z_k$), concentrated over fewer traffic lanes ($\alpha' > \alpha$). The combination of these two effects must reduce congestion for the policy of traffic separation to be Pareto improving.

From equation (10), if z is sufficiently large, the condition is not extremely restrictive. Indeed, the tax levied on car users is largely compensated by the payoffs resulting from the lump-sum benefit of the considered tax. However, when z decreases, the tax base shrinks, making this condition more and more restrictive.

One can conveniently rewrite equation (11) by separating two effects: a positive effect (decrease in congestion due to the lower number of cars) and a negative effect (increase in congestion due to the increase of traffic separation):

$$t^c(\alpha, z) - t^c(\alpha', z^\alpha) = [t^c(\alpha, z) - t^c(\alpha, z^\alpha)] - [t^c(\alpha', z^\alpha) - t^c(\alpha, z^\alpha)]. \tag{12}$$

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

4.3.2 Relative efficiency of policy tools

It is not possible to compare the absolute efficiency of these two policies without considering specific functional forms for congestion. Nevertheless, a general intuition of the relative efficiency of these two policies can be derived from the following proposition.

Proposition 4 *Assume there exist two distinct policies, α_1 and T_1, that yield the same equilibrium, z_1, and that car users enjoy the same utility under either of these two policies. Then, for any other two distinct policies, α_2 and T_2, yielding another equilibrium, z_2 ($z_2 < z_1$), car users always prefer the policy of traffic separation to the policy of taxation.*

Proof. See Appendix A.4. ■

It follows from this proposition that even though we cannot theoretically exclude the possibility that one of the two policies is always better than the other, if this is not the case, taxation should be preferred for small changes in z, while separation should be preferred for larger changes.

This relates to the two schools of thought we presented in the literature review. If a social planner is convinced that the city is car-dependent, and that any policy can only have a marginal impact on the modal split, then a policy of taxation may be the best policy. But, if one believes that a large shift can take place, traffic separation might be a better choice.

5 Extensions

5.1 Capacity constraints and discomfort externalities

There are two possible types of congestion in public transportation. First, public transportation is a source of congestion. For instance, there may be so many buses on a bus lane that the travel time on that lane increases with the number of public transportation users. Second, congestion can occur when commuters cannot enter in the first bus and face a queue to access public transportation. This would increase the waiting time for commuters.

Remember that an equilibrium is such that $\varepsilon_i = \Delta(\alpha, z) - W(z) + f_c$. Congestion between public transportation leads $\Delta(\alpha, z)$ to decrease after some threshold, say, $\overline{z}$. Congestion within public transportation implies that instead of enjoying network externalities among the users of public transportation, $W(z)$ decreases below a certain threshold.

Capacity constraints, viewed in a strict way, would be that there is no mean to serve the demand for public transportation if this demand is higher than a given threshold, say $(1-\overline{z})$. In this case, for people exceeding this threshold, waiting time goes to infinity. In Figure 5, we illustrate the case of a strict capacity constraint, i.e. no possibility to transport more than a share $(1-\overline{z})$ of the population by public transportation. Note that any line located between the dashed line (no congestion within or between public transportation) and the vertical line could be achieved in the presence of different degrees of capacity constraints and/or congestion in public transportation.

Another representation of congestion within the public transportation system is to explicitly model discomfort externalities. Crowding costs (Kraus, 1991) are imposed by every marginal passenger on other passengers while the Mohring effect could be seen as a discrete process. Consider a bus lane where a bus is added when a given number of passengers per bus is reached. Every time a bus is added, the quality of public transportation "jumps" by discretely reducing travel time. However, within any bus, there is marginal congestion. The discomfort externality locally reduces the incentive to use public transportation when the number of passengers per

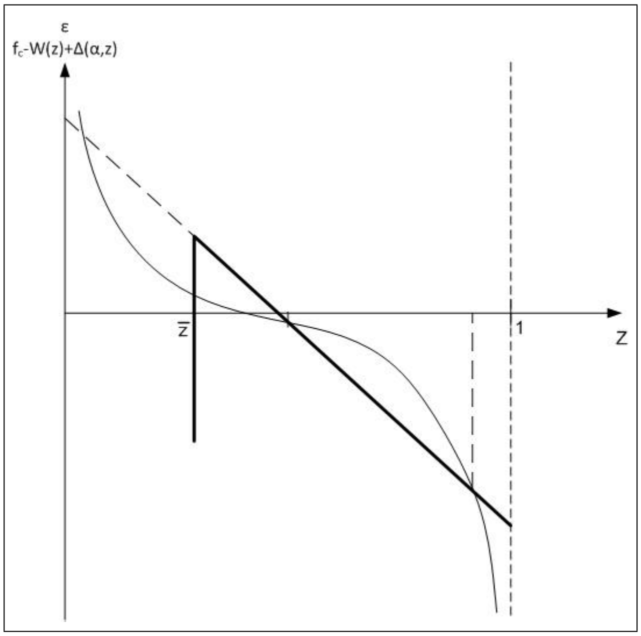

Figure 5: Strict capacity constraint in public transportation

bus increases. Assuming the network effect outweighs the congestion effect, the cost function is modified as in Figure 6. W(z) is locally decreasing in z and stepwise increasing in z. With this modified setup, the likelihood of facing multiple equilibria increases. The incentive to remain locally with an equilibrium share of car users is increased by the fact that, at the margin, increasing the share of public transportation users decreases the average quality of public transportation.

5.2 Building an underground

For the moment, we have considered traffic separation corresponding to bus lanes or light rail: more space for public transportation and less space for cars. Another way to prevent public transportation from congestion is to build an underground. In comparison to delimiting bus lanes (which is almost free), an underground is much more costly to build. Hence, even if one cannot deny that an underground can be an efficient way to provide fast public transportation and decrease congestion within public transport, it is likely to be counter-productive when dealing with coordination problems. Indeed, the underground may actually decrease the travel time

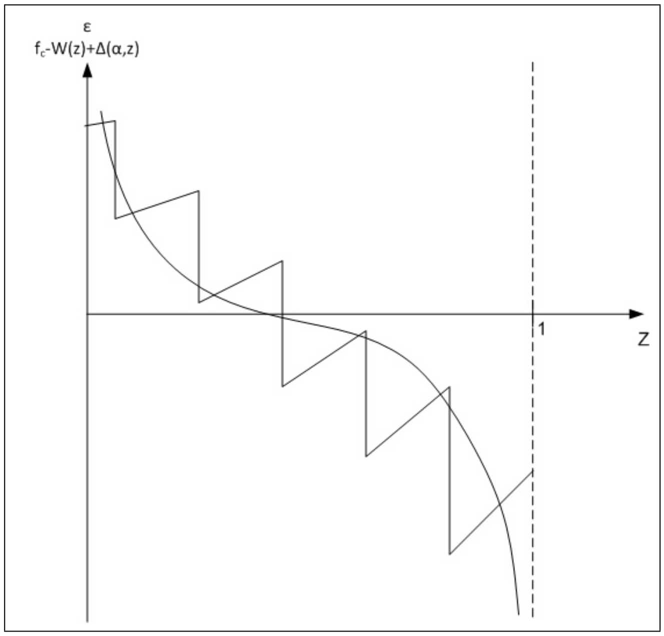

Figure 6: Discomfort externalities in public transportation

both for public transportation and for cars. If, as in Tabuchi (1993), the cost of infrastructure is supported only by public transportation users, building an underground may actually decrease the share of public transportation users in the modal shift.

Assume that the underground is the only available type of public transportation. $\widetilde{\alpha}$ is now the investment in the underground. This investment is associated with $M(\widetilde{\alpha})$, the lump sum cost paid by all commuters, independent of their modal choice.

The effect of $\widetilde{\alpha}$ on time spent commuting by car is now $\frac{\partial t^c(\widetilde{\alpha},z)}{\partial \widetilde{\alpha}} \leq 0$, because the underground does not reduce (and potentially increases) the space for cars in the city. The marginal impact of $\widetilde{\alpha}$ on $t^{pt}(\widetilde{\alpha}, z)$ remains negative: more investments in public transportation reduce the commuting time when using public transportation.

Defining $\Delta^u(\widetilde{\alpha}, z) = t^c(\widetilde{\alpha}, z) - t^{pt}(\widetilde{\alpha}, z)$, the utilities of both type of commuters become

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

and the new equilibrium is defined by

$$\varepsilon_i = f_c - W(z) + \Delta^u(\widetilde{\alpha}, z).$$

Since $\forall\alpha = \widetilde{\alpha}$, ($\alpha$ and $\widetilde{\alpha}$ $\in [0,1]$), we have $\Delta^u(\widetilde{\alpha}, z) \leq \Delta_1(\widetilde{\alpha}, z)$, the equilibrium comes with a lower share of public transportation users and, in the case of multiple equilibria, it is even more difficult to avoid the 'bad' equilibrium. This result would be even stronger if the cost of the underground were supported by public transportation users only, as in Tabuchi (1993). However, an argument in favor of the existence of an underground could be linked to congestion in public transportation (as discussed in the previous extension). In that case, an underground can be seen as a means to expand the supply of public transportation in the presence of capcity constraints.

6 Conclusion

We show that the combination of externalities of congestion, cross modal externalities and network externalities with heterogeneous commuters can lead to multiple equilibria. This may explain why a priori similar cities end up with very different patterns of car use. We also show that policy tools, namely taxation and traffic separation, are not equivalent in terms of welfare. When one of the two is not strictly more efficient than the other, separation should be preferred for large-scale policies while taxation should be preferred for smaller modifications of commuting patterns. This result partly explains the differences in policy recommendations from the two schools in the literature - physical planning and car-dependant cities - suggesting different reforms to improve the modal choice within a city. On the one hand, a policymaker that believes (as "physical planners") that there must be an important change in the modal split should focus on the allocation of space (α in our model) and increase the share of roads devoted to public transportation only. On the other hand, a social planner only concerned with marginal changes in the cost of cars - or that simply believes that car dependence is the best possible state of the world for a given city - should privilege taxation.

References

[1] Anas, A, Arnott, R and K Small (1998). "Urban Spatial Structure" ; *Journal of Economic Literature ;* 36 ; 3 ; pp. 1426-1464

[2] Arnott, R & de Palma, A and R Lindsey (1992). "Route choice with heterogeneous drivers and group-specific congestion costs," Regional Science and Urban Economics, Elsevier, vol. 22(1), pages 71-102, March.

[3] Batarce, M and M Ivaldi (2011). "Travel Demand Model with Heterogeneous Users and Endogenous Congestion: An application to optimal pricing of bus services" ; *IDEI Working Papers,* Toulouse

[4] Beesley, M and M Kemp (1987). "Urban Transportation" ; *Handbook of Regional and Urban Economics;* vol. 2 ; chpt. 26 ; pp. 1023-1052

[5] Beirão, G and Cabral, S (2007). "Understanding attitudes towards public transportation and private car: A qualitative study" ; *Transport Policy ;* 14 ; pp. 478-489

[6] Belzer, D and G Aultier (2002). "Transit oriented development: moving from rhetoric to reality" ; *Brookings Institutions Center on Urban Metropolitan Policy*; Discussion paper ; 55pp.

[7] Berglas, E, Fresko, D and D Pines (1984). "Right of Way and Congestion Toll", Journal os Transport Economics and Policy, Vol. 18, N°2, pp. 165-187.

[8] Calfee, J and C Winston (1998). "The value of automobile travel time: implications for congestion policy"; *Journal of Public Economics ;* 69 ; pp.83-102

[9] Cain A, Darido G, Baltes M, Rodriguez P and J Barrios (2006). "Applicability of Bogotá's TransMilenio BRT System to the United States", *Federal Transit Administration Working Paper*

[10] Cervero R., Ferrell C 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.

[11] de Palma, A and R Lindsey (2002). "Private roads, competition, and incentives to adopt time-based congestion tolling," Journal of Urban Economics, Elsevier, vol. 52(2), pages 217-241, September.

[12] de Palma, A, Kilani, M and R Lindsey (2008). "The merits of separating cars and trucks," Journal of Urban Economics, Elsevier, vol. 64(2), pages 340-361, September.

[13] De Palma A and R Lindsey (2011). "Traffic congestion pricing methodologies and technologies", *Transportation Research Part C*

[14] De Vlieger I, De Keukeleere D and JG Kretzschmar. (2000). "Environmental effects of driving behaviour and congestion related to passenger cars" ; *Atmospheric Environment ;* 34 ; pp.4469-4655

[15] Dittmar H and G Ohland (2003). "The new transit town: best practices in transit-oriented development", *Island Press*.

[16] Dobruszkes F and Y Fourneau (2007). "Coûts directs et géographie des ralentissements subis par les transports publics bruxellois" ; *Brussels Studies* ; N°7.

[17] Duranton, G and M Turner (2011). "The Fundamental Law of Road Congestion: Evidence from US Cities," *The American Economic Review*, 101 (6), 2616–2652.

[18] Echeverry JC, Ibáñez AM, Moya A, Hillón LC, Cárdenas M and A Gómez-Lobo (2005). "The Economics of TransMilenio, a Mass Transit System for Bogotá", *Economía*, Vol. 5, No. 2 (Spring, 2005), pp. 151-196

[19] Goodwin, P (2004). "The Economic Costs of Road Traffic Congestion", ESRC Transport Studies Unit, University College London Working Paper, 26pp

[20] Handy S, Weston L and P Mokhtarian (2005). "Driving by choice or necessity?", *Transportation Research Part A: Policy and Practice* ; pp. 183-203

[21] Hiscock R, Macintyre S, Kearns A and A Ellaway (2002). "Means of transport and ontological security: Do cars provide psycho-social benefits for their users ?" ; *Transportation Research part D* ; 7 ; pp. 1119-1135

[22] Jensen M (1999). "Passion and Earth in transport: a sociological analysis on travel behaviour" ; *Transport Policy* ; 6 ; pp. 19-33

[23] Kenworthy J and F 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

[24] Mirabel F (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

[25] Mohring H (1972). "Optimization and Scale Economies in Urban Bus Transportation" ; *The American Economic Review* ; 62 ; 4 ; pp. 591-604

[26] Parry I. (2002). "Comparing the efficiency of alternative policies for reducing traffic congestion" ; *Journal of Public Economics* ; 85 ; 3 ; pp. 333-362

[27] Parry I and K Small (2009). "Should Urban Transit Subsidies Be Reduced ?" ; *The American Economic Review* ; 99 ; 3 ; pp. 700-724

[28] Pucher J (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

[29] Pucher J and J Renne (2003). "Socioeconomics of Urban Travel: Evidence from the 2001 NHTS" ; *Transportation Quarterly* ; 57 ; 3 ; pp. 49-77

[30] Rosenthal R (1973). "A class of Games Possessing Pure-Strategy Nash Equilibria" ; *International Journal of Game Theory* ; 2 ; 1 ; pp. 65-67

[31] Santos G and J Bhakar (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

[32] Steg L (2005). "Car use: lust and must. Instrumental, symbolic and affective motives for car use" ; *Transportation Research part. A* ; 39 ; pp. 147-162

[33] Tabuchi, T (1993). "Bottleneck Congestion and Modal Split" ; *Journal of Urban Economics* ; 34 ; pp. 414-431

[34] Van Vugt M, Van Lange P and Meertens R (1996). "Commuting by car or public transportation ? A social dilemna analysis of travel mode judgements" ; *European Journal of Social Psychology* ; 26 ; pp. 373-395

[35] Verhoef E and K Small (2004). "Product Differentiation on Road. Constrained Congestion Pricing with Heterogeneous Users" ; *Journal of Transport Economics and Policy* ; 38 ; 1 ; pp. 127-156

[36] Vickrey W (1963). "Pricing in Urban and Suburban Transport" ; *The American Economic Review* ; 53 ; 2 ; pp. 452-465

[37] Wardman, M (2001). "A review of British evidence on time and service quality valuations" ; Transportation Research Part E: Logistics and Transportation Review, Volume 37, Issues 2?3, April?July 2001, pp 107-128

A Proofs

A.1 Proof of Proposition 2

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

$$G(z_k) = f_c - W(z_k) + \Delta(\alpha, z_k).$$

First, we show that the existence of a Nash equilibrium, z_k, satisfying

$$\frac{\partial G(z_k)}{\partial z} > \frac{\partial f_c - W(z_k) + \Delta(\alpha, z_k)}{\partial z}, \tag{13}$$

is a sufficient condition for the existence of multiple equilibria. Second, we show that this is a necessary condition.

(i) If there exist such a z_k, then for any $\eta > 0$ arbitrarily small, we have

$$\begin{aligned} G(z_k) &= f_c - W(z_k) + \Delta(\alpha, z_k) \\ G(z_k + \eta) &> f_c - W(z_k + \eta) + \Delta(\alpha, z_k + \eta) \\ G(z_k - \eta) &< f_c - W(z_k - \eta) + \Delta(\alpha, z_k - \eta). \end{aligned}$$

Since the support of F is $(-\infty, \infty)$ and therefore $G(1) < f_c - W(1) + \Delta(\alpha, 1)$ and $G(0) > f_c - W(0) + \Delta(\alpha, 0)$. It implies that the functions must cross at least three time and there exists at least three equilibria. So, condition (13) is a sufficient condition for the existence of multiple equilibria.

(ii) Assume that at any Nash Equilibrium z_k, we have

$$\frac{\partial G(z_k)}{\partial z} < \frac{\partial f_c - W(z_k) + \Delta(\alpha, z_k)}{\partial z}$$

then for any $\eta > 0$, we have

$$\begin{aligned} G(z_k) &= f_c - W(z_k) + \Delta(\alpha, z_k) \\ G(z_k + \eta) &< f_c - W(z_k + \eta) + \Delta(\alpha, z_k + \eta) \\ G(z_k - \eta) &> f_c - W(z_k - \eta) + \Delta(\alpha, z_k - \eta). \end{aligned}$$

Since the support of F is $(-\infty, \infty)$, for any $z', z'' : z' < z_k < z''$, $G(z') > f_c - W(z') + \Delta(\alpha, z')$ and $G(z'') < f_c - W(z'') + \Delta(\alpha, z'')$. This implies that the functions cross only once and condition (13) is necessary.

$>$From (i) and (ii), condition (13) is, indeed, a necessary and sufficient condition. ∎

A.2 Proof of Proposition 3

Proof. Assume there exist T equilibria $z_1 < z_2 < ... < z_T$

(1) We want to show that z_1 Pareto dominates any equilibrium z_j, $j = \{2, ..., T\}$

(2) For any pair z_j, z_1 with $z_j > z_1$, there are three categories of commuters:

(a) Commuters with ε_i such that $F(\varepsilon) < 1 - z_j$. Their best response is to use public transportation in both equilibria. Those users are better off in equilibrium z_1 as

$$t^{pt}(\alpha, z_1) < t^{pt}(\alpha, z_j) \text{ and } W(z_1) < W(z_j),$$

then

$$t^{pt}(\alpha, z_1) + W(z_1) + \frac{\varepsilon_i}{2} < t^{pt}(\alpha, z_j) + W(z_j) + \frac{\varepsilon_i}{2}$$

(b) Commuters with ε_i such that $1 - z_j < F(\varepsilon) < 1 - z_1$. Their best response is public transportation in equilibrium z_1 and car in equilibrium z_j. Those users are better off in equilibrium z_1. Indeed, as commuters reveal their preferences by choosing their mode, then for any $F(\varepsilon_j)\epsilon[1 - z_j, 1 - z_1]$:

$$W(z_j) + t^{pt}(\alpha, z_j) + \varepsilon_j > f_c + t^c(\alpha, z_j) \quad (14)$$

and

$$W(z_1) + t^{pt}(\alpha, z_1) + \varepsilon_j < f_c + t^c(\alpha, z_1). \tag{15}$$

As congestion increases in z,

$$f_c + t^c(\alpha, z_j) > f_c + t^c(\alpha, z_1).$$

Hence, it is straightforward that

$$f_c + t^c(\alpha, z_j) > W(z_1) + t^{pt}(\alpha, z_1) + \varepsilon_j$$

(c) Commuters with ε_i such that $1 - z_1 < F(\varepsilon)$. Their best response in both equilibria is to take the car. Those users are better off in equilibrium z_1 as:

$$f_c + t^c(\alpha, z_j) > f_c + t^c(\alpha, z_1).$$

■

A.3 Proof of Lemma 3

Proof. Let us divide the population into three families: those who use public transportation 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 public transportation afterward (C-PT). We show the effect of both policies on the differents families described above.

In the case of taxation:

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

$$U_T^{pt} - U^{pt} = W(z_k) - W\left(z^T\right) + t^{PT}(\alpha, z_k) - t^{PT}\left(\alpha, z^T\right) + z^T T$$

which can be decomposed into three effects, all welfare enhancing (as long as $z^T < z_k$): $W(z_k) - W\left(z^T\right)$ corresponds to the reduction of waiting time; $t^{PT}(\alpha, z_k) - t^{PT}\left(\alpha, z^T\right)$ comes from the reduction of congestion; and $z^T T$ comes from the lump sum transfer from car users to the user of public transportation.

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

$$U_T^c - U^c = t^c(\alpha, z_k) - t^c\left(\alpha, z^T\right) - \left(1 - z^T\right) T > 0$$

i.e. it increases their welfare if

$$t^c(\alpha, z_k) - t^c\left(\alpha, z^T\right) > \left(1 - z^T\right) T$$

3) C-PT: It is easy to show that if car users are better off, swing commuters are also better off. Swing commuters prefer public transportation to the car under the policy leading to z^T. Hence, if their utility when using a car is higher then in z_k, so is their utility while using public transportation.

In the case of separation, increasing α to α':

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

$$U_\alpha^{pt} - U^{pt} = W(z_k) - W(z^\alpha) + t^{PT}(\alpha, z_k) - t^{PT}\left(\alpha', z^\alpha\right)$$

which can be decomposed into two effects, both being welfare enhancing (as long as $z^\alpha < z_k$): $W(z_k) - W(z^\alpha)$ corresponds to the reduction of waiting time a,d $t^{PT}(\alpha, z_k) - t^{PT}(\alpha', z^\alpha)$ 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

$$U_\alpha^c - U^c = t^c(\alpha, z_k) - t^c\left(\alpha', z^\alpha\right) > 0$$

i.e. it increases their welfare if

$$t^c(\alpha, z_k) > t^c\left(\alpha', z^\alpha\right)$$

3) C-PT: The reasoning is similar as for the taxation policy. If car users are better off, swing commuters are also better off. ■

A.4 Proof of Proposition 4

Proof. Starting from $\alpha_0 \geq 0$ and $T_0 = 0$ and given the definition and the properties of $\beta(\alpha, z)$ (note that $\beta(\alpha, z) = -\frac{t_{\alpha}^{PT\prime}(\alpha^*,z)}{t_{\alpha}^{c\prime}(\alpha^*,z)}$ with $\beta'_{\alpha}(\alpha, z) \geq 0 \; \forall z \in [0,1]$), it is possible to define $\gamma(\alpha)$ such that

$$t^{pt}(\alpha_1, z_1) - t^{pt}(\alpha_0, z_1) = -\gamma(\alpha_1)\left[t^c(\alpha_1, z_1) - t^c(\alpha_0, z_1)\right], \tag{16}$$

with $\gamma'(\alpha) > 0$.

(i) Consider two policies: α_1 ($\alpha_1 > \alpha_0$ is associated with $T = 0$) and T_1 (associated with α_0) yielding the same equilibrium z_1. By definition, a commuter indifferent between the two policies has an ε_j such that

$$f_c + T_1 - W(z_1) + \Delta(\alpha_0, z_1) = \varepsilon_j = f_c - W(z_1) + \Delta(\alpha_1, z_1).$$

This simplifies to

$$T_1 = \Delta(\alpha_1, z_1) - \Delta(\alpha_0, z_1),$$

which can be conveniently rewritten as

$$T_1 = \left[t^c(\alpha_1, z_1) - t^c \alpha_0, z_1\right] - \left[t^{pt}(\alpha_1, z_1) - t^{pt}(\alpha_0, z_1)\right].$$

By assumption, this leads to

$$T_1 = (1 + \gamma(\alpha_1))\left[t^c(\alpha_1, z_1) - t^c(\alpha_0, z_1)\right]. \tag{17}$$

(ii) Car users are indifferent between these two policies if $\exists \; \alpha_1$, T_1, z_1 such that

$$(1 - z_1) T_1 = t^c(\alpha_1, z_1) - t^c(\alpha_0, z_1).$$

These conditions imply

$$\begin{aligned}(1-z_1)[1+\gamma(\alpha_1)][t^c(\alpha_1,z_1)-t^c(\alpha_0,z_1)] &= t^c(\alpha_1,z_1)-t^c(\alpha_0,z_1)\\ (1-z_1) &= \frac{1}{1+\gamma(\alpha_1)}.\end{aligned}$$

Now consider two alternative policies associated with a higher use of public transportation: α_2, T_2, z_2 with $z_2 < z_1$. Car users are now better off with traffic separation than with taxation iff

$$(1-z_2)\,T_2 > t^c(\alpha_2,z_2)-t^c(\alpha_0,z_2).$$

>From the expression of T in equation (17),

$$\begin{aligned}(1-z_2)[1+\gamma(\alpha_2)][t^c(\alpha_2,z_2)-t^c(\alpha_0,z_2)] &> t^c(\alpha_2,z_2)-t^c(\alpha_0,z_2)\\ (1-z_2) &> \frac{1}{1+\gamma(\alpha_2)}.\end{aligned}$$

This is always true as $z_2 < z_1$, $\gamma(\alpha_2) \geq \gamma(\alpha_1)$ and given that $(1-z_1) = \frac{1}{1+\gamma(\alpha_1)}$ ■