# Spoor B1:

## How to enforce optimal transport pricing and investment in a federal state.

S. Van der Loo
S. Proost

KULeuven
Faculteit Economie en Bedrijfswetenschappen

**Algemeen secretariaat – Steunpunt beleidsrelevant Onderzoek Fiscaliteit & Begroting**
Voskenslaan 270 – 9000 Gent – België
Tel: 0032 (0)9 248 88 35 – E-mail: vanessa.bombeeck@hogent.be
www.steunpuntfb.be

# How to enforce optimal transport pricing and investment in a federal state

Saskia van der Loo[1] & Stef Proost[2]

**Abstract**

A federal government tries to force local governments to implement welfare optimal tolling and investment. Local governments have an incentive to charge more than the marginal social cost whenever there is transit traffic. We analyse the pricing and investment issue in an asymmetric information setting and discuss the case of air pollution and of congestion.

Author information

[1] Van der Loo: Department of Economics, Catholic University of Leuven, Naamsestraat 69, B-3000 Leuven, (saskia.vanderloo@econ.kuleuven.be),
[2] Proost: Department of Economics, Catholic University of Leuven, Naamsestraat 69, B-3000 Leuven, (stef.proost@econ.kuleuven.be)..

We would like to acknowledge the GRACE consortium (6 th Framework Research Program of the EC ) (http://www.grace-eu.org) and the Policy research Center Budgetary & Tax Policy, Government of Flanders (B) for financial support.

# 1 Introduction

A well known result of economic theory is that efficiency requires prices equal to the marginal social cost. This, however, is only valid in a first-best setting and amendments to this simple rule are necessary in the presence of additional constraints. This paper focusses on one particular second-best constraint that has not yet been studied in detail in the context of transportation, namely the presence of incentive and information problems when there is more than one policy maker. This problem is relevant for EU – member state conflicts as well as for coordination problems within one country.

Different levels of governments that have conflicting objectives leads to uncoordinated pricing policies. While the upper level (EU, or country) is concerned with the welfare of all the citizens and wants social marginal cost based pricing, a lower level government (a member state or region) may prefer much higher transport charges to extract revenue from transit. This problem is well known. It has been empirically validated for state gasoline taxes in the USA by Levinson [6]. The pricing and investment issues that arise in federations where local governments toll local and transit traffic have been studied in [4] for parallel networks and in [5] for serial networks. These papers did not address the asymmetric information issue.

In the latest White Paper of the European Commission [3], the EU acknowledges that pricing of transport infrastructure should relate to the marginal social cost associated with the use of the infrastructure, but the current interpretation (for tolling roads) is to impose a toll cap related to the infrastructure costs and not to the external costs. One of the arguments why the EU can not impose first-best marginal social cost pricing is that it may lack the necessary information about the marginal social cost of the different member states. The member states can very easily misuse the uncertainty on external costs in their advantage and pretend to have much higher external costs than in reality. To be allowed to charge higher taxes some regions can pretend their ecosystem is very vulnerable or that their urban planning is such that more people are exposed to air pollution than in other regions. In general it is realistic to assume that the member state will have better information about the local marginal external costs than the EU. This is an asymmetric information problem where the lower level governments knows the social marginal cost with more precision.

In this paper we will look whether a federal authority can implement optimal pricing under asymmetric information. Under asymmetric information the federal authority can ask the regions to report their marginal external costs and implements pricing accordingly. We analyze the problem in a very simple setting; one link crossing a single state, we assume all other links perfectly priced such that the simplest partial equilibrium approach is suitable. The conflict between local and central government arises from the fact that the local government wants to extract as much revenue as possible from transit users. For this reason we allow two types of users on the link; local and transit users.

Probably the most important externality in transportation is congestion. This externality differs in two important ways from other externalities: the social marginal cost depends explicitly on the level of usage and the level of congestion has a feedback effect on the level of use. For other externalities such as pollution or noise social marginal costs are constant and do not necessarily influence the level of use.

The paper is organized as follows. In the second section we analyze the first best solution that can be achieved by an omniscient federal government that has to deal with air pollution or road congestion. In the third section we introduce a local government that has different tolling preferences and focuses on air pollution externalities only. Here we find that a revelation mechanism exists that allows the federal government to make the local government implement the first best pricing solution by a well designed transfer scheme. In the fourth section we concentrate on the more difficult case of congestion externalities. We show that, if the transit traffic share is sufficiently large, the federal government is unable to set up a transfer scheme that leads to first best results. In the fifth section we generalize the model by including road capacity decisions and examine the solution advocated by the EU for roads: constrain the local road tolls to be smaller or equal to the average infrastructure costs. In the before last section we look into another generalization of the model: what happens if there are several types of road users, say cars and trucks etc. The last section sums up our findings and adds some caveats.

# 2 First-best pricing

As a benchmark case we analyze the setting where only local traffic is present. When only local traffic is present and neglecting political economy issues, both federal and local government will have the same objective and we get the standard first-best results. We discuss first the case of air pollution and then congestion.

## 2.1 Air pollution

We use a partial equilibrium model to analyze pricing decisions of a single (isolated) link crossing a country (or region). In this section there is only one kind of user, namely local users. The usage is denoted by $X^L$ and is determined by a downsloping inverse demand function $P_L\left(X^L\right)$

$$P_L\left(X^L\right) = a_L - b_L X^L, \quad a_L > 0, b_L > 0. \tag{1}$$

The objective function of both governments (local and federal) is the sum of the surplus of the users (two first terms in 2), plus the toll revenues minus the external costs.

$$W = \int_0^{X^L} P_L\left(x\right) dx - \tau X^L + \tau X^L - e X^L \tag{2}$$

where $\tau$ is the toll levied on transportation, $e$ the constant marginal external cost of one unit of $X^L$. Important in our analysis is that the marginal external cost is constant, does not affect the level of usage and has a purely local impact. Local air pollution damage could be an example, accident externalities imposed by cars on cyclists and pedestrians could be another.

In equilibrium, demand will be equal to the user cost. As we neglect here the other private resource costs, the user cost consists of the toll only. The equilibrium volume is then given by $X^L\left(\tau\right) = \frac{a_L - \tau}{b_L}$. An increase (or decrease) in the toll will reduce (or increase) the traffic volume:

$$\frac{\partial X^L}{\partial \tau} = -\frac{1}{b_L} < 0. \tag{3}$$

Both local and federal government will choose $\tau$ such as to maximize the social welfare function given in (2). The first order condition with respect to $\tau$ is:

$$\frac{\partial W}{\partial \tau} = P\left(X^L\right)\frac{\partial X^L}{\partial \tau} - \tau\frac{\partial X^L}{\partial \tau} + (\tau - e)\frac{\partial X^L}{\partial \tau} = 0$$

Using (1) and (3), the optimal toll is seen to be equal to the marginal environmental damage:

$$\tau^* = e$$

As the marginal air pollution damage is constant, the Pigouvian tax solution is very simple.

## 2.2 Congestion

When the marginal social cost of the externality is not constant, but depends on the usage level of the infrastructure, as is the case with congestion, then the user cost equals the time cost plus the toll where the time cost is an increasing function of the usage. The time cost and the discomfort of travel will in principle increase when a higher volume is loaded on the same infrastructure: average speed will decrease, passengers won't have a seat etc. in the train. We assume that the user cost function is linear in the volume of transport[1]:

$$C\left(X^L\right) = \alpha + \beta X^L + \tau, \quad \alpha > 0, \beta > 0. \tag{4}$$

The objective function for both local and federal government is the sum of the surplus of the users minus the user cost (two first terms in 4), plus the tax revenues (now the external costs are incorporated in the user cost function):

$$W = \int_0^{X^L} P_L(x)\,dx - C\left(X^L\right)X^L + \tau X^L. \tag{5}$$

In equilibrium, demand will equal the user cost ($P_L\left(X^L\right) = C\left(X^L\right)$), and the equilibrium volume is:

$$X^L = \frac{a_L - \alpha - \tau}{\beta + b_L}. \tag{6}$$

Contrarily to the case of air pollution, the level of congestion will now affect the level of usage (feedback effect). If the infrastructure is more easily congestible, say the capacity of the infrastructure is smaller, $\beta$ increases and the usage decreases:

$$\frac{\partial X^L}{\partial \beta} = -\frac{X^L}{\beta + b_L} < 0. \tag{7}$$

[1] The linear user cost function could be seen as the reduced form cost function of a simple bottleneck model with homogeneous users [1].

Again the governments will maximize the social welfare (now given by (5)) with respect to the toll. The first order condition is:

$$\frac{\partial W}{\partial \tau} = P\left(X^L\right)\frac{\partial X^L}{\partial \tau} - \frac{\partial C\left(X^L\right)}{\partial \tau}X^L - C\left(x^L\right)\frac{\partial X^L}{\partial \tau} + X^L + \tau\frac{\partial X^L}{\partial \tau} = 0,$$

and the optimal toll is (using (4), (7) and $P\left(X^L\right) = C\left(X\right)$);

$$\tau^* = \beta X^L. \tag{8}$$

As expected, the more congestible the infrastructure (the higher $\beta$), the higher the marginal external cost and the higher the optimal toll:

$$\frac{\partial \tau^*}{\partial \beta} = \frac{b_L\left(a_L - \alpha\right)}{\left(2\beta + b_L\right)^2} > 0.$$

# 3 Enforcing marginal social cost pricing when air pollution is the only externality

Introducing transit traffic will create a divergence between local and federal government objectives. Transit traffic is traffic by residents of another locality belonging to the federation. In order to concentrate on the asymmetric information issue we neglect the strategic interactions when transit traffic uses networks of several regions as studied in [4] and [5]. The local government maximizes the surplus of the local users plus the revenue it can extract from transit. The federal government is interested in maximizing welfare of all users and wants therefore to control the tolling practices of the local government. To emphasize the difference in local decision making when transit traffic is present or not we start by analyzing the case where there is only transit traffic and generalize later to the case of transit and local traffic. As the type of external cost is crucial for the enforcement of first best pricing, we first focus on air pollution.

## 3.1 The case of only transit traffic

The local government collects the tolls paid by the transit users and is not interested in their welfare. Its objective function is therefore equal to the total toll revenue minus the (local) external cost caused by the traffic:

$$\Pi = \left(\tau - e\right)X^T, \tag{9}$$

where $\tau$ is the toll, $e$ the constant marginal external cost and $X^T$ the transit volume. The demand function for transit is assumed similar to that of local traffic used in the previous section: $P_T\left(X^T\right) = a_T - b_T X^T$.
Contrarily to the local government, the federal government is concerned by the welfare of all citizens, including the transit users and will maximize an objective function similar to (2) where $X^L$ is now replaced by $X^T$. The optimal toll from a federal point of view will therefore be again equal to the Pigouvian tax, namely

$$\tau^* = e.$$

### 3.1.1 The toll preferred by the local government

The local authority will charge a toll $\tau^N$ to the users of the facility that maximizes its welfare given in (9). This toll will solve the first order condition for $\tau$ which is

$$X^T + (\tau - e)\frac{\partial X^T}{\partial \tau} = 0,$$

implying

$$\tau^N = e + b_T X^T = \frac{e + a_T}{2}.$$

The toll increases with the marginal environmental damage. In fact the marginal environmental damage can be considered as a marginal cost for the local government. The toll charged by the local government $\tau^N$ exceeds the social marginal cost because the local government is able to raise revenues by charging transit users,

$$\tau^N > e = \tau^*.$$

Note that we need $a_T > e$ to ensure $X^T > 0$; the maximum willingness to pay for usage of the infrastructure must be at least the damage caused by usage. When the local government is free to set the toll equal to $\tau^N$, its welfare is

$$\Pi = \frac{(a_T - e)^2}{4b_T} > 0,$$

deriving this expression with respect to the damage cost gives us

$$\frac{\partial \Pi}{\partial e} = -\frac{(a_T - e)}{2b_T},$$

which is negative since $a_T > e$ : the higher the damage cost, the lower the local welfare. When $e = a_T$, then the local welfare is zero.

### 3.1.2 Federal toll regulation with asymmetric information

We now suppose that the marginal environmental damage $e$ is unknown to the federal authority: it only knows that the region has either a low marginal environmental damage $\left(e = e^L\right)$ or a high one $\left(e = e^H > e^L\right)$. This uncertainty is not unrealistic. Some regions pretend their ecosystem is very vulnerable or that their urban planning is such that more people are exposed to air pollution than in other regions.

The game is the following: in the first stage, the regional government reports its marginal environmental cost $\tilde{e}^i \in \left\{e^L, e^H\right\}$ to the federal government. In the second stage, the federal government imposes a toll contingent on this report. To ensure truthful reporting we assume that the federal government can make a financial transfer to the regions. These financial transfers $M\left(\tilde{e}^i\right)$ will be such that a region always has the incentive to report its true marginal damage, i.e. the incentive constraints are satisfied. Note that this problem is similar to the problem of regulating a monopoly with unknown costs (see [2]) but since we assume that the monetary transfers do not represent a real cost to society there will be no trade off between efficiency and paying "information rents". Whereas in the classic principal-agent problem the principal will be willing to deviate from the efficient outcome in order to pay less rent, here the principle (= the

federal government) will always implement the first best tolls. Our aim is to check whether it is possible for the federal government to implement first-best tolls while ensuring truthful reporting.
The lower level government knowing that it will have to charge a toll equal to its reported marginal damage will choose to report a marginal damage $\tilde{e}^i$ such as to maximize following function:

$$\max_{\tilde{e}^j} \Pi\left(\tilde{e}^j, e^i\right) = \left[\tilde{e}^j - e^i\right] X^T\left(\tilde{e}^j\right) + M\left(\tilde{e}^j\right), \quad i, j = \{L, H\},$$

$\Pi\left(\tilde{e}^i, e^i\right)$ being the local welfare for a region with marginal damage $e^i$, reporting a marginal damage equal to $\tilde{e}^i$. The transfer scheme $M\left(\tilde{e}^i\right)$ is such that it is beneficial for a region to report its true marginal damage. Since it is the difference between transfers that will be important we can set $M\left(\tilde{e}^H\right) = 0$ and $M\left(\tilde{e}^L\right) = M$ ($M$ can in principle be negative) and the incentive compatibility constraints can be written as:

$$\Pi\left(\tilde{e}^H, e^H\right) \geq \Pi\left(\tilde{e}^L, e^H\right) + M \tag{10}$$

$$\Pi\left(\tilde{e}^L, e^L\right) + M \geq \Pi\left(\tilde{e}^H, e^L\right) \tag{11}$$

These are the incentive compatibility (IC) constraints. The first constraint ensures that a region whose true marginal damage is high will prefer to report a high marginal damage $\tilde{e}^H$ and receive no financial transfer rather than to lie and report a low marginal damage and receive $M$. The second constraint ensures in the same way that a region with a low marginal damage will have no incentive to misreport its marginal damage.
Lets first have a look to the behavior of the local authority when there are no transfers. A region with low marginal damage will have an incentive to misreport its damage because it can increase its welfare by pretending to have a high marginal damage $\left(\Pi\left(\tilde{e}^H, e^L\right) > \Pi\left(\tilde{e}^L, e^L\right)\right)$. A region with high marginal damage, will, on the other hand, have an incentive to tell the truth since $\Pi\left(\tilde{e}^H, e^H\right) > \Pi\left(\tilde{e}^L, e^H\right)$. This is easily seen in Figure 1.
In order for a region with a low marginal damage to tell the truth, it must be compensated with a financial transfer. The lowest transfer needed to induce truthtelling from such a region will be $M = \Pi\left(\tilde{e}^H, e^L\right) - \Pi\left(\tilde{e}^L, e^L\right)$. It remains to check wether the IC of the high marginal damage region (10) is satisfied. Using the fact that $\Pi\left(\tilde{e}^H, e^H\right) = \Pi\left(\tilde{e}^L, e^L\right) = 0$, (10) reduces to

$$\left(e^L - e^H\right)\left(e^H - e^L\right) < 0.$$

This is always true since $e^H > e^L$, which leads to the first proposition:

**Proposition 1** *When there is only transit traffic and when the environmental damage is unknown to the federal government, the federal government can still implement the first-best tolls. For a region with low environmental damage* $e^L$ *to report truthfully, it will, however, need a financial compensation equal to* $M = \left(e^H - e^L\right) X^T\left(e^H\right)$.

When the difference between the two marginal damages is large, the greater is the gain for a low damage region to pretend to have a high marginal damage and the larger the compensation for truthtelling need to be.

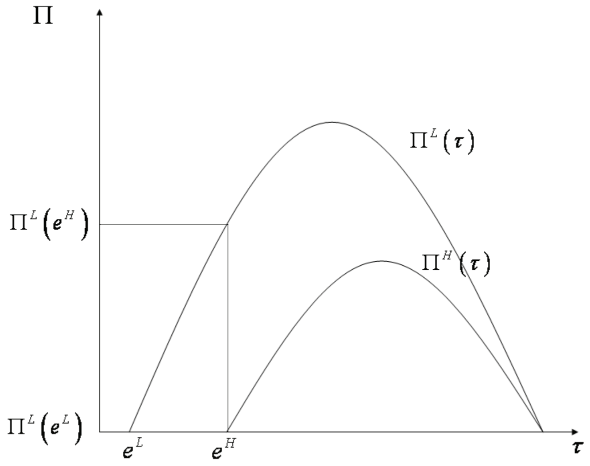


Figure 1: Local welfare functions in the presence of air pollution and in the case that there is only transit traffic. Where $\Pi^I, I = L, H$ stands for the local welfare of a region with low/high marginal cost.

## 3.2 The case with transit and local traffic

When there is both local and transit traffic, the local government will only be concerned about the welfare of the local users and the revenues generated by the transit users. Its objective function is the sum of the surplus of the local users (two first terms), the total toll revenues and the total external costs:

$$\Pi = \int_0^{X^L} P^L(x)\,dx - \tau X^L + (\tau - e)\,X, \tag{12}$$

where $X = X^T + X^L$, the total amount of users. The federal government, on the other hand, takes into account the welfare of both local and transit users:

$$W = \int_0^{X^L} P^L(x)\,dx + \int_0^{X^T} P^T(x)\,dx - \tau X + (\tau - e)\,X.$$

The federal first-best toll is again $\tau^* = e$.

### 3.2.1 The toll preferred by the local government

Solving the first order condition of (12) with respect to $\tau$ yields the preferred toll $\tau^N$, which is of the form

$$\tau^N = e - \frac{X^T}{\frac{\partial X}{\partial \tau}}.$$

Substituting $\frac{\partial X}{\partial \tau}$ in the expression for $\tau^N$ we get a toll level that is excessive:

$$\tau^N = e + \frac{b_L b_T}{b_L + b_T} X^T > \tau^*. \tag{13}$$

Moreover, the more transit users, the higher the locally preferred toll will be. Note that the presence of local users will partly protect the transit users of being excessively tolled since $b_T > \frac{b_L b_T}{b_L + b_T}$ and the toll levied when no local users are present will be even higher.

#### 3.2.2 Federal toll regulation with asymmetric information

As in the case when there was only transit traffic, we now assume that the environmental damage is only known to the local government. Again the local government reports a marginal damage costs $\tilde{e}^i \in \{e^L, e^H\}$ . Doing so it will have to implement a toll equal to $\tilde{e}^i$ and receive a financial transfer $M(\tilde{e}^i)$ which is zero for $\tilde{e}^i = e^H$ and equal to $M$ when $\tilde{e}^i = e^L$. We saw that a region with high environmental damage will charge a toll that is higher than the corresponding marginal damage, which on its turn is larger than the first-best toll for low environmental damage:

$$\tau^{NH} > \tau^{*H} > \tau^{*L},$$

where $\tau^{*i} \equiv \tau^*(e^i), i = L, H$ and $\tau^{Ni} \equiv \tau^N(e^i), i = L, H$. Since the local objective function is a parabolic function of the toll with a maximum for $\tau = \tau^N(e^H)$ and since both $\tau^*(e^H)$ and $\tau^*(e^L)$ are at the same side of the maximizing toll we have that

$$\Pi(\tilde{e}^H, e^H) > \Pi(\tilde{e}^L, e^H).$$

Again, as was the case when no local users were present: a region with high environmental damage has no incentive to lie. Graphically, we have the following situation

The incentive compatibility constraint for a low damage region is

$$\Pi(\tilde{e}^H, e^L) = \Pi(\tilde{e}^L, e^L) + M.$$

In Figure **??** we see that, depending on the relative position of the first-best toll in case of a high damage and the local preferred toll for low damage, a region with low damage will or will not have an incentive to lie when no transfers would be available.

We can show that when the locally preferred toll satisfies following inequality

$$\tau^{NL} < e^L + \frac{e^H - e^L}{2}$$

then a low damage region will never have an incentive to lie about its marginal and the federal government can implement first best tolls without having to make any transfers, i.e.$M = 0$. Since the deviation of the locally first best toll from the first-best toll depends on the volume of transit (see (13)), this inequality tells us that if transit traffic is not very important, then a low damage region will never have an incentive to lie about its marginal damage cost. If transit traffic is important enough, however, a region with low damage costs will have to be compensated in order to report truthfully, the transfer will be equal to

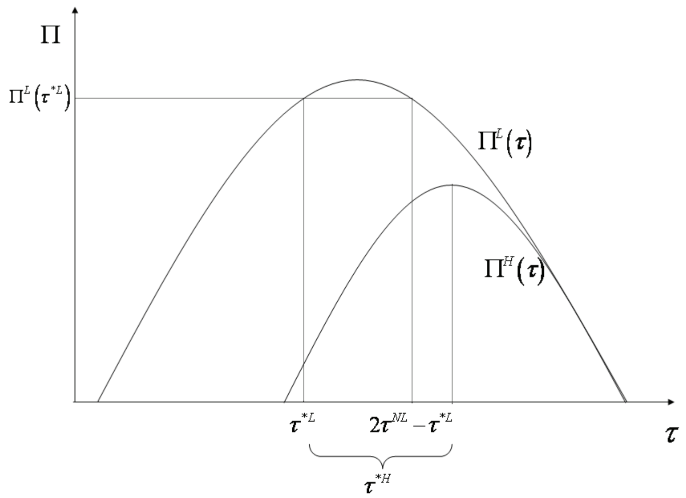


Figure 2: Air pollution case: Local welfare with both local as transit users.

$M = \Pi\left(\tilde{e}^H, e^L\right) - \Pi\left(\tilde{e}^L, e^L\right)$. This transfer could in principle induce a high damage region to mimic a low damage region in order to receive the transfers. It is, however, easy to check that the IC constraints for a high damage region given by:

$$\Pi\left(\tilde{e}^H, e^H\right) > \Pi\left(\tilde{e}^L, e^H\right) + M$$

are equivalent with

$$X\left(e^H\right) > X\left(e^L\right)$$

which is always the case since by assumption $e^H > e^L$.

**Proposition 2** *When there is both local and transit traffic and the marginal environmental damage is unknown to the federal government, a truthful mechanism exists in which each region sets its toll equal to its marginal environmental damage.*
*If*

$$X^T\left(\tau^N, e^L\right) < \frac{\left(e^H - e^L\right)\left(b_L + b_T\right)}{2b_L b_T},$$

*no compensation is needed I.e.* $M = 0$*,*
*if this condition is not satisfied, then a financial compensation is needed in order to induce a "low cost" region to report its cost truthfully. This compensation must be equal to*

$$M = \left[2b_L\left(a_T - e^H\right) - b_T\left(e^H - e^L\right)\right]\frac{\left(e^H - e^L\right)}{2b_L b_T}.$$

# 4 Enforcing marginal social cost pricing when congestion is the only externality

## 4.1 The case of only transit traffic

In the following sections we assume that congestion is the only externality present. A distinctive feature of congestion is that it, contrarily to externalities discussed in the previous sections, it affects the users of the infrastructure and will influence the demand levels. The local government is not interested in the welfare of the transit users, it will only be interested in the congestion costs of transit users in as far as they affect transit demand and the toll revenues. When only transit users are present the objective function of the local government is therefore very simple: it is equal to the total toll revenue

$$\Pi = \tau X^T, \tag{14}$$

where $\tau$ is the toll and $X^T$ the transit volume. The federal first-best toll is $\tau^* = \beta X^T$ (see (8)).

### 4.1.1 The toll preferred by the local government

Solving the first order condition of (14) yields

$$\tau^N = \frac{a_T - \alpha}{2},$$

the toll is independent of the congestion level. The local welfare will however depend on the level of congestion;

$$\Pi\left(\tau^N\right) = \frac{\left(a_T - \alpha\right)^2}{4\left(\beta + b_T\right)},$$

and the more congestible the infrastructure (the higher $\beta$), the lower the local welfare:

$$\frac{\partial \Pi}{\partial \beta} = \frac{-\left(a_T - \alpha\right)^2}{4\left(\beta + b_T\right)^2} < 0. \tag{15}$$

### 4.1.2 Federal toll regulation with asymmetric Information

In this section we suppose that the federal government is not well informed about the marginal external costs of congestion. Again this is not an realistic assumption. The marginal external cost depends on values of time (so on composition of traffic). It also consists of schedule delay costs [1] so that observations on the length of queues etc. are insufficient information. The lack of information concerns the slope of the user cost function, or more precisely, the parameter $\beta$. We assume that the federal government only knows that the slope of the user cost function can be either $\beta = \beta^L$ or $\beta = \beta^H$, where $\beta^H > \beta^L$. The larger the parameter $\beta$, the more easily congestible is the infrastructure and so we will refer to a region with $\beta = \beta^L$ as a region with "low marginal external congestion cost (mecc)" and to a region with $\beta = \beta^H$ as a region with "high mecc". As was the case in section 3.1.2. we will check whether with the help of financial transfers, it is possible to implement the first-best outcome.

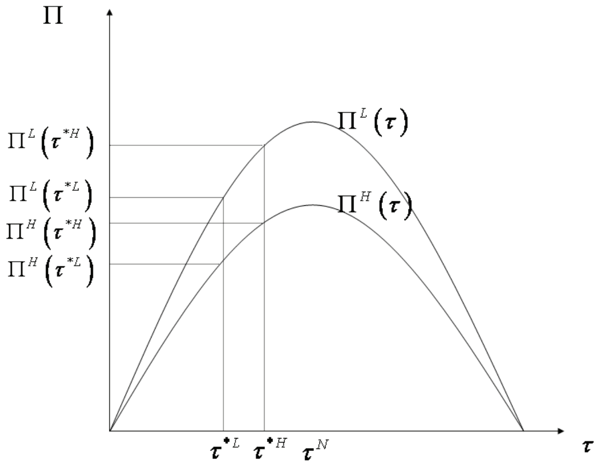


Figure 3: Local welfares for different road congestibility and tolls

The problem for the local government is to choose its reported mecc $\tilde{\beta}^i \in \left\{\beta^L, \beta^H\right\}$ such that it maximizes

$$\Pi\left(\tilde{\beta}^j, \beta^i\right) = \tau\left(\tilde{\beta}^j\right) X^T\left(\tau\left(\tilde{\beta}^j\right), \beta^i\right) + M\left(\tilde{\beta}^j\right),$$

where $\tau\left(\tilde{\beta}^j\right) = \tilde{\beta}^j X^T\left(\tau\left(\tilde{\beta}^j\right), \beta^i\right)$, the first-best toll given that the mecc is equal to the $\tilde{\beta}^j$. Again we can assume $M\left(\tilde{\beta}^H\right) = 0$ and $M\left(\tilde{\beta}^L\right) = M$ , where $M$ has to satisfy the incentive constraints:

$$\begin{aligned} \Pi\left(\tilde{\beta}^H, \beta^H\right) \quad &\geq \Pi\left(\tilde{\beta}^L, \beta^H\right) \quad + M \\ \Pi\left(\tilde{\beta}^L, \beta^L\right) + M &\geq \Pi\left(\tilde{\beta}^H, \beta^L\right). \end{aligned}$$

When no transfers are available, we can see in Figure 3 that a country with a low mecc will have an incentive to misreport its mecc. If a country has high mecc, on the other hand has an incentive to tell the truth and thus $M = \Pi\left(\tilde{\beta}^H, \beta^L\right) - \Pi\left(\tilde{\beta}^L, \beta^L\right) > 0$. A country with low mecc will need to be compensated to be truthful and the IC constraints reduce to

$$\Pi\left(\tilde{\beta}^H, \beta^H\right) - \Pi\left(\tilde{\beta}^L, \beta^H\right) > M.$$

In Figure 3 we see that for an identical toll, the toll revenues of a region with low mecc will be higher than for a region that has a more easily congestible infrastructure since there will be more traffic using its infrastructure: $\Pi\left(\tau, \beta^L\right) \geq$

$\Pi\left(\tau,\beta^{H}\right)$ for all $\tau$. Both welfare functions will be equal to zero when $\tau$ is 0 or $\alpha$. These two properties imply that $\left|\frac{\partial^2\Pi\left(\tau,\beta^H\right)}{\partial\tau^2}\right| < \left|\frac{\partial^2\Pi\left(\tau,\beta^L\right)}{\partial\tau^2}\right|$, which on its turn implies that for every $M > 0$ and for every $0 \leq \tau_1 < \tau_2 \leq \alpha$ :

$$\Pi\left(\tau_1,\beta^L\right) + M \geq \Pi\left(\tau_2,\beta^L\right) \Rightarrow \Pi\left(\tau_1,\beta^H\right) + M \geq \Pi\left(\tau_2,\beta^H\right),$$

where the equality on the right-hand side only holds when the left-hand side holds for equality. This property holds for every $\tau_1$ and $\tau_2$ and thus also for the special case where $\tau_1 = \tilde{\beta}^L$ and $\tau_2 = \tilde{\beta}^H$ and we see that there is a conflict with the IC constraints. This means that the federal government will not be able to find a transfer scheme that induces a region to declare its true mecc and implement the corresponding first-best toll, even if it has access to financial transfers. In fact the result holds for any pair of tolls and financial transfers $M$, and the federal government will never be able to induce a truthful report of the mecc. Note that the major difference with the air pollution type of externalities is that there the second derivative of the local welfare is constant. This difference reflects the fact that congestion has an influence on the demand levels, while air pollution does not (feedback effect).

**Proposition 3** *When there is only transit traffic and the marginal external cost of congestion (mecc) is unknown to the federal government, no truthful mechanism exists that allows the federal government to implement marginal social cost pricing.*

The best thing the federal government can do in this case is to maximize expected welfare and impose a toll cap. No type will have an incentive to go below the toll cap provided that $\tau^{\text{cap}} > \tau^N$.

## 4.2 The case with transit and local traffic

When there are also local users, the welfare function of the local government will be the sum of the user surplus of the local users (first two terms) plus the total toll revenues:

$$\Pi = \int_0^{X^L} P^L(x)\,dx - C(X)\,X^L + \tau X. \tag{16}$$

In contrast, the federal government will also take into account the user surplus of the transit users:

$$W = \int_0^{X^L} P^L(x)\,dx + \int_0^{X^T} P^T(x)\,dx - C(X)\,X + \tau X.$$

Equating the demand functions for transit and local users (equations (**??**) and (1) respectively) to the linear user cost function similar to (4), yields us the transit and local volumes in function of the mecc and the toll. Deriving the resulting expressions for the volumes with respect to the toll yields:

$$\frac{\partial X^L}{\partial\tau} = \frac{-b_T}{B} < 0, \qquad \text{and} \qquad \frac{\partial X^T}{\partial\tau} = \frac{-b_L}{B} < 0$$

where $B \equiv \beta(b_L + b_T) + b_L b_T$. As expected, both user volumes decrease when the toll increases.

### 4.2.1 The toll preferred by the local government

We obtain an expression for the locally preferred toll by solving the f.o.c. with respect to $\tau$ of (16):

$$\tau^N = \beta X^L - \frac{X^T}{\frac{\partial X}{\partial \tau}}.$$

Since $\frac{\partial X}{\partial \tau} < 0$,

$$\tau^N > \text{lmecc}$$

The toll exceeds the local marginal external cost, defined as the marginal external cost imposed on the locals, and the more transit there is, the larger will be the difference between the locally preferred toll and the federal optimal toll (see [4])
Substituting $\frac{\partial X}{\partial \tau}$ in the expression of $\tau^N$ we get

$$\tau^N = \beta X(\tau, \beta) + \frac{b_T b_L}{b_T + b_L} X^T(\tau, \beta) \tag{17}$$

and so

$$\tau^N > \beta X(\tau, \beta) = \text{mecc}. \tag{18}$$

The toll charged by the local government exceeds the social marginal cost[2].
Deriving (17) with respect to $\beta$ yields:

$$\frac{\partial \tau^N}{\partial \beta} = \frac{b_T}{(b_T + 2b_L)} X > 0.$$

For higher mecc (higher $\beta$) the local authority will charge a higher toll and so $\tau^N\left(\beta^H\right) > \tau^N\left(\beta^L\right)$ as expected.

### 4.2.2 Federal toll regulation with asymmetric Information

Take now the case where the exact value of $\beta$ $\left(\beta^L \text{ or } \beta^H\right)$ is unknown by the federal government.
We see in Figure 4 that, as usual, a region with high mecc never has an incentive to lie since $\Pi\left(\tilde{\beta}^L, \beta^H\right) < \Pi\left(\tilde{\beta}^H, \beta^H\right)$, but a country with low mecc will in some cases have an incentive to lie when no transfers exists. Similarly to the case of air pollution type of externalities, when

$$\tau^{*H} > 2\tau^{NL} - \tau^{*L}$$

a low mecc region has no incentive to lie. This last condition can be rewritten:

$$X^T\left(\tau^{NL}, \beta^L\right) < \frac{(Z-\alpha)(b_T+b_L)^2\left(\beta^H - \beta^L\right)}{2\left(\beta^H(b_T+b_L) + B^H\right)\left(\beta^L(b_T+b_L) + B^L\right)}. \tag{19}$$

When there is an incentive for a low mecc region to mimic a region with high mecc, we will again have cases where the monetary transfers needed to induce

[2]Note that the volumes are, however, the volumes for $\tau = \tau^N$ and not the first-best volumes. It can be shown that the first-best toll is lower than the locally preferred toll whenever $\beta(b_T + b_L) + b_T b_L > 0$, which is always the case.

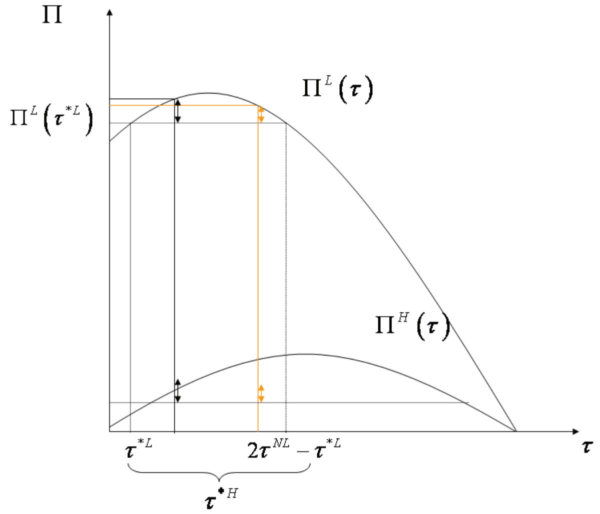


Figure 4: Local welfares for different congestion functions when there is local and transit traffic.

truthtelling will be such that the IC for a hig mecc region will be violated. This will happen whenever

$$\tau^{*H} < \tau^{NL}.$$

When this condition is satisfied, it is impossible for a federal government to implement the first-best outcome with the help of monetary transfers. Writing the inequality in function of the transit volume yields

$$X^T\left(\tau^{NL}, \beta^L\right) > \frac{(Z-\alpha)\left(b_T+b_L\right)^2\left(\beta^H-\beta^L\right)}{\left(\beta^H\left(b_T+b_L\right)+B^H\right)\left(\beta^L\left(b_T+b_L\right)+B^L\right)}, \tag{20}$$

where $B^A \equiv \beta^A\left(b_L+b_T\right)+b_L b_T$ (for a derivation of these two expressions, see Appendix).
This means that only when the transit share is small enough, the regions will declare their true mecc. If the transit share is bigger, a region with low mecc will have an incentive to overstate its mecc. Again, compensation will be needed. This compensation will in some cases induce a high mecc region to declare it has low mecc. From Figure 4 we see that this is the case when $\tau^{*H} < \tau^{NL}$, or when condition (20) is not satisfied.

**Proposition 4** *When there is both local and transit traffic and the marginal external congestion cost is unknown to the federal government, there are three cases*

1. *conditions (19) is satisfied: the federal government can set the toll equal to the mecc corresponding to the declared $\beta$*

2. *condition (20) is satisfied: no mechanism exists where the federal government can induce a region to report its mecc truthfully and impose the corresponding first-best toll.*

3. *neither (19) nor (20) are satisfied: the federal government can set toll equal to the mecc corresponding to the declared $\beta$ but needs to make a financial transfer if a region declares it has a low mecc*

# 5 A toll cap equal to the average infrastructure costs

We have seen in the previous section that in the case of congestion (or any kind of externalities with feedback effects) when there is transit traffic, there is no obvious way to implement first-best tolling when there is some uncertainty about the magnitude of the externality. As said in the introduction, the current practise in the EU is to constrain the toll level by a toll cap which equals the average infrastructure cost. In this section we will investigate to what extend this practise makes sense. The advantage is that such a cap does not require any knowledge about the level of congestion and will therefore not rely on the reporting of the marginal external costs from the regional governments. The federal government needs only to know the total infrastructure costs and the toll revenues. Assuming constant returns to scale in road capacity costs, the total infrastructure costs ($TC$) are equal to

$$TC = \frac{k}{\beta},$$

where $k$ is the unit cost of capacity and $1/\beta$ is the level of capacity. It is cheaper to provide a highly congested or badly serviced road (low capacity, high $\beta$). Note that both the unit cost of capacity and the level of capacity can be unknown to the federal government, we only assume that the total costs are known. The toll revenues collected by the regional government can not exceed the total infrastructure cost, so

$$\tau X < TC$$

The local government has now the freedom to choose not only the toll level $\tau$ but also the capacity level $1/\beta$ .

## 5.1 The case of only transit traffic

The objective functions are as before but include now the infrastructure costs which are born by the local government. With only transit traffic, the objective function equals the toll revenues minus the infrastructure costs:

$$\Pi = \tau X^T - \frac{k}{\beta}. \tag{21}$$

The federal welfare is given by

$$W = \int_0^{X^T} P(x)\,dx - C\left(X^T\right) X^T + \tau X^T - \frac{k}{\beta}. \tag{22}$$

### 5.1.1 The toll and investment level preferred by the federal government

It is interesting to see what would happen if the federal government had the possibility to choose toll and capacity. If this would be the case it would choose $\beta$ and $\tau$ such as to maximize federal welfare. It will have to solve the following two f.o.c. simultaneously

$$\frac{\partial W}{\partial \tau} = 0 \quad \text{and} \quad \frac{\partial W}{\partial \beta} = 0.$$

The first f.o.c. yields the first-best toll

$$\tau^* = \beta X^T (\tau^*, \beta).$$

The f.o.c. for $\beta$ can be rewritten as

$$\left[\tau - \beta X^T\right] \frac{\partial X^T}{\partial \beta} - X^T X^T + \frac{k}{\beta^2} = 0.$$

Using, $\tau = \beta X$ and the fact that $\beta$ should be positive we have that:

$$\beta^* = \frac{\sqrt{k}}{X^T (\tau^*, \beta^*)}. \tag{23}$$

The higher the marginal infrastructure cost, the more congested the infrastructure will be since the government will invest less. The more transit, the more revenues can be extracted and the more can be invested. Substituting $\beta^*$ back in the expression for the first-best toll we get

$$\tau^* = \sqrt{k}$$

and this produces the well known cost-recovery result.

### 5.1.2 The toll and investment level preferred by the local government

It is instructive to see what the local authority would choose as its capacity level $(1/\beta)$ and toll if it is not subjected to regulation. The problem of the local government is to solve the next two equations simultaneously:

$$\begin{aligned} \frac{\partial \Pi}{\partial \tau} &= X^T + \tau \frac{\partial X^T}{\partial \tau} = 0 \\ \frac{\partial \Pi}{\partial \beta} &= \tau \frac{\partial X^T}{\partial \beta} + \frac{k}{\beta^2} = 0 \end{aligned}$$

The first equation gives us the same result as previously:

$$\tau^N = \frac{(a_T - \alpha)}{2}.$$

Using the derivatives with respect to $\beta$; $\frac{\partial X^T}{\partial \beta} = -\frac{X^T}{b_T + \beta}$ and substituting the result of the f.o.c for $\tau = \frac{(a_T - \alpha)}{2} = \frac{X^T}{b_T + \beta}$ in the first order conditions for $\beta$, we get an expression for the capacity level preferred by the local government:

$$\beta^N = \frac{\sqrt{k}}{X^T\left(\tau^N, \beta^N\right)}.$$

This is the optimal capacity from the federal point of view (see eq.(23)) given the level of usage. This means that if the federal government could induce optimal charging, the regional government would automatically opt for the optimal investment level.

### 5.1.3 A toll cap equal to the average infrastructure cost

Now the local government has to observe the following constraint:

$$\tau X^T(\tau, \beta) \leq \frac{k}{\beta}.$$

The optimization problem for the local government becomes:

$$\max_{\tau,\beta} \quad \Pi \tag{24}$$

$$\text{s.t. } \tau X^T(\tau, \beta) - \frac{k}{\beta} \leq 0 \tag{25}$$

where $\Pi$ is given in eq(21). It is clear that in this case the local government will choose toll and capacity levels such that the toll revenues exactly equal the infrastructure costs since otherwise it will have negative welfare and will choose not to invest at all. The local government will be indifferent to all pairs of tolls and capacity levels that yield zero welfare and that satisfy $\tau X^T(\tau, \beta) = k/\beta$. Using the equilibrium expression for $X^T = \frac{a_T - \alpha - \tau}{\beta + b_T}$, we see that the constraint reduces to:

$$\tau\left(a_T - \alpha - \tau\right) = k + \frac{k b_T}{\beta}.$$

so for every $\beta$ in a feasible range there is a $\tau$ that satisfies the constraint and there is an infinity of solutions that satisfy this constraint but only one is optimal from a federal point of view.

## 5.2 The case of transit and local traffic

### 5.2.1 Federal price setting

The federal optimization problem yields the same result as in the case where there is only transit traffic but now the transit flow $X^T$ is replaced by the total flow $X$ :

$$\begin{aligned} \tau^* &= \beta^* X\left(\beta^*, \tau^*\right) \\ \beta^* &= \frac{\sqrt{k}}{X\left(\beta^*, \tau^*\right)} \end{aligned}$$

substituting $\beta^*$ in $\tau^*$, the first-best toll becomes

$$\tau^* = \sqrt{k}$$

### 5.2.2 Local pricing and investment strategy

Without regulation, the local authority would choose its investment level ($\beta$) and $\tau$ such to maximize its welfare $\Pi - \frac{k}{\beta}$, where $\Pi$ is given in (16). It has to solve the next two equations:

$$\begin{aligned} \frac{\partial \Pi}{\partial \tau} &= X + \tau \frac{\partial X}{\partial \tau} + P(X_L) \frac{\partial X^L}{\partial \tau} - \frac{\partial C}{\partial \tau} X^L - C \frac{\partial X^L}{\partial \tau} = 0 \\ \frac{\partial \Pi}{\partial \beta} &= P^L(X_L) \frac{\partial X^L}{\partial \beta} - \frac{\partial C}{\partial \beta} X^L - C \frac{\partial X^L}{\partial \beta} + \tau \frac{\partial X}{\partial \beta} + \frac{k}{\beta^2} = 0 \end{aligned}$$

The first equation gives us a toll

$$\tau^N = \beta X\left(\tau^N, \beta\right) + \frac{b_T b_L}{b_T + b_L} X^T\left(\tau^N, \beta\right)$$

The f.o.c. for $\beta$ is

$$\left(\tau - \beta X^L\right) \frac{\partial X}{\partial \beta} - X X^L + \frac{k}{\beta^2} = 0$$

Substituting $\tau^N$ and using $\frac{\partial X}{\partial \tau} = \frac{1}{X} \frac{\partial X}{\partial \beta}$ gives

$$\beta^N = \frac{\sqrt{k}}{X\left(\tau^N, \beta^N\right)}$$

Again, the capacity will be set optimally. Note that the toll level is larger than the optimum so that flows are smaller. The local authority will thus invest too little and charge too much.
With local traffic, the local government could in principle charge tolls smaller than the average infrastructure cost and still have positive welfare. Under the constraint, the optimization problem for the local government becomes:

$$\max_{\tau,\beta} \Pi, \tag{26}$$

$$\text{s.t. } \tau X - \frac{k}{\beta} \leq 0 \tag{27}$$

where $\Pi$ is given by eq(16).
Define the Langragian

$$\mathcal{L} = \Pi - \lambda\left(\tau X - \frac{k}{\beta}\right)$$

then the Khun-tucker conditions are

$$\frac{\partial \mathcal{L}}{\partial \tau} = 0 \tag{28}$$

$$\frac{\partial \mathcal{L}}{\partial \beta} = 0 \tag{29}$$

$$\lambda \geq 0, \tau X - \frac{k}{\beta} \leq 0 \text{and } \lambda\left(\frac{k}{\beta} - \tau X\right) = 0 \tag{30}$$

the first and the second equation gives

$$\tau = \frac{B}{(b_T + b_L)} X - \frac{b_T b_L}{(1-\lambda)(b_T + b_L)} X^L \tag{31}$$

$$\frac{k}{\beta^2} = \left[\frac{b_T b_L}{(1-\lambda) A}\right] X X^L + \frac{(b_T + b_L)}{B} \tau X \tag{32}$$

where $B = \beta (b_T + b_L) + b_T b_L$
The last equation is satisfied when either $\lambda = 0$ or $\frac{k}{\beta} - \tau X = 0$. It is easy to show that when $\lambda = 0$ (or with other words there is no constraint), that the toll revenues will always equal or exceed the infrastructure costs. So this leads us to conclude that the constraint is binding. If this is so, we have an unique solution, namely we have a solution

$$\tau = \sqrt{k}, \quad \beta = \frac{\sqrt{k}}{X}, \tag{33}$$

which corresponds exactly with the first best solution[3]. This means that when the higher level has no knowledge about the marginal external congestion cost, it can still achieve the first-best by letting the local government decide about capcity levels and tols, provided that the average infrastructure cost cap holds.

# 6 Conclusions

In this paper it is assumed that the federal government lacks information on the external costs created by transit traffic. The local government knows the external costs and uses this asymmetry in information to charge transit and local traffic more than the marginal social cost.
If the external cost does not affect the use of the infrastructure (as in the case of some forms of air pollution), there exist a transfer scheme by which the federal government can induce the local government to charge the right tolls. If no transfer scheme exists all that can be done is to impose a toll cap equal to the expected value of the marginal external cost.
When the external cost is of the congestion type so that the level of congestion affects the level of use, a transfer scheme to induce the local government to implement the right toll only exists if transit traffic is not sufficiently important. These results are summarized in Table 1.

| | **Air pollution** | | **Congestion** | | |
|---|---|---|---|---|---|
| | **low transit** | **high transit** | **low transit** | **medium transit** | **high transit** |
| **Transit** | with transfer | | no mechanism exists | | |
| **Local and Transit** | without transfer | with transfer | without transfer | with transfer | no mechanism exists |

[3]the langargian multiplier $\lambda$ is then equal to the proportion of the transit traffic in the overall traffic $\left(\lambda = \frac{X^T}{X}\right)$.

Table 1: conditions for which a mechanism exists that induces truthfull reporting of the marginal social costs

If there is only type of traffic and there are constant returns to scale in capacity extension, one can achieve a first best outcome for prices and investment by using as toll cap the average infrastructure cost.

In this paper we assumed that there is only one type of users and that their contribution to the externality is identical. When transit and local traffic have the same unitary air pollution or congestion effect, there is only one parameter that is unknown to the federal government.

The propositions derived in this paper can be generalized to the case of several types of users if their relative contribution to the externality is known. This would be the case if the relative emission rates of trucks and cars are known or if the relative congestion contribution of cars and trucks are known.

Having different types of users does however most likely create problems to use the average infrastructure cost as toll cap. In this case the federal government does control neither the toll nor the investment levels. When there are more types of users the federal government can control easily the absence of discrimination between local and transit traffic for each type but this will be insufficient. Whenever the transit share of one type of users is larger, there will be an incentive for the local government to overcharge this group. This is a well known result in the tax exporting literature. The implication is that the first best character of a toll cap equal to the average infrastructure cost most probably will break down. To see this, take an extreme example and assume that all trucks except one are transit trucks but that all passenger cars are local traffic. Tolls for trucks will be inefficiently high and the toll cap can not prevent this and the powerful result that using toll caps equal to the average infrstructure cost is enough to ensure efficient pricing and investment breaks down.

This paper uses a very simple model and several extensions are worth studying. One extension is the use of more complex networks. The competition for transit traffic may limit the pricing power of local government levels in the case of parallel networks [4]. A second extension is to consider a wider range of instruments, besides transfer mechanisms and toll caps; one may also consider quality standards for roads or uniform fixed tariffs.

# A Derivation of conditions (19) and (20)

The equilibrium volume are determined by the Wardrop equilibrium concept where

$$Z - b_T X^T = \alpha + \beta \left( X^T + X^L \right) + \tau = Z - b_L X^L$$

Solving this for $X^T$ and $X^L$ and using $X = X^T + X^L$ we get

$$X\left(\tau, \beta^A\right) = \frac{(Z-\alpha)\left(b_T + b_L\right)}{B^A} - \frac{\left(b_T + b_L\right)}{B^A}\tau \tag{34}$$

where $B^A \equiv \beta^A \left(b_L + b_T\right) + b_L b_T$.

From (8), (18) and solving the equilibrium volumes we know that

$$\tau^{*A} = \beta^A X\left(\tau^{*A}, \beta^A\right) \tag{35}$$

$$\tau^{NA} = \beta^A X\left(\tau^{NA}, \beta^A\right) + \frac{b_T b_L}{b_T + b_L} X^T\left(\tau^{NA}, \beta^A\right) \tag{36}$$

Substituting (35) in (34) yields,

$$X\left(\tau^{*A}, \beta^A\right) = \frac{(Z-\alpha)(b_T + b_L)}{2\beta^A (b_T + b_L) + b_L b_T} \tag{37}$$

Substituting (36) in (34) and using (37) we get

$$X\left(\tau^{NA}, \beta^A\right) = X\left(\tau^{*A}, \beta^A\right) - \frac{b_T b_L}{2\beta^A (b_T + b_L) + b_L b_T} X^T\left(\tau^{NA}, \beta^A\right) \tag{38}$$

With the help of these four last equations (first using (35) and (36), then substituting (38) and after rearranging the terms we substitute (37)) we get the result that the inequality $\tau^{*H} < \tau^{NL}$ can be rewritten as

$$X^T\left(\tau^{NL}, \beta^L\right) > \frac{(Z-\alpha)(b_T + b_L)^2 \left(\beta^H - \beta^L\right)}{\left(\beta^H (b_T + b_L) + B^H\right)\left(\left(\beta^L - 1\right)(b_T + b_L) + B^L\right)}$$

Condition (20) is obtained in a similar way.

# References


[1] Arnott R., de Palma A., Lindsey R. (1993) A Structural Model of Peal-Period Congestion: A Traffic Bottleneck with Elastic Demand. *The American Economic Review* **83** (1), 161-179

[2] Baron P.B., Myerson R.B. (1982) Regulating a Monopolist with Unknown Costs. *Econometrica* **50** (4), 911-930

[3] Commission of the European Communities (CEC) (1998), Fair payement for infrastructure use: A phased approach to a common transport infrastructure charging framework in the EU. White Paper, COM (1998) 466 final. Brussels: European Comission

[4] De Borger B., Proost, S., Van Dender K. (2005) Congestion and tax competition in a parallel network. *European Economic Review* **49** (8), 2013-2040

[5] De Borger B., Dunkerley F., Proost, S. (Forthcoming) Strategic Investment and Pricing Decisions in a Congested Transport Corridor. *Journal of Urban Economics*

[6] Levinson, D. (2001), Why State Toll - An empirical model of finance choice. *Journal of Transport Economics and Policy*, **35**(2), 223-238

Steunpunt beleidsrelevant onderzoek 2007-2011

fiscaliteit en begroting

HOGESCHOOL GENT
LID VAN DE ASSOCIATIE UNIVERSITEIT GENT

UNIVERSITEIT GENT

KATHOLIEKE UNIVERSITEIT LEUVEN