Policies to reduce traffic externalities in cities

Bruno DE BORGER and Stef PROOST

Energy, Transport and Environment

Faculty of Economics
And Business

Policies to reduce traffic externalities in cities (*)

Bruno De Borger
Stef Proost

Abstract

This paper considers various policy measures to reduce traffic externalities in cities, including externality-reducing investments, tolls, emission standards, low emission zones, and bypass capacity to guide traffic around the city center. Using a simple model that distinguishes local and through traffic, we study the optimal use of these instruments by an urban government that cares for the welfare of its inhabitants, and we compare the results with those preferred by a federal authority that takes into account the welfare of all road users. Our results include the following. First, compared to the federal social optimum, we show that the city government will over-invest in externality-reducing infrastructure whenever this infrastructure increases the generalized cost of transit traffic. Second, comparing emission standards and road tolls, we find that cities with a lot of commuters will favor tolls, even though from the federal perspective standards are better. Third, when implementing low emission zones, the urban government will set both the fee for non-compliance and the standard at a higher level than the federal government. Moreover, at sufficiently high transit levels the urban government will prefer imposing a toll instead of implementing a low emission zone. Fourth, if the city can toll the urban infrastructure, it will only invest in bypass capacity when it is allowed to earn extra toll revenues on the bypass that exceed investment costs. Although the paper focuses on non-congestion externalities, most insights also hold in the presence of congestion.

JEL: R48, Q58,H77

Keywords: externalities, urban policy, urban transport investment, environment

Introduction

This paper considers various policy measures that governments can use to reduce pollution-type traffic externalities in cities. The model is sufficiently rich to allow for a large variety of such policy instruments. It can obviously deal with standard instruments such as speed limits or urban tolls, but it also easily accommodates a wide range of specific local externality-reducing investments such as traffic lights, road bumps, building noise walls, using 'quiet' asphalt, investment in bypass capacity or in technology to guide traffic around the city, and investment in traffic structures for incoming traffic (e.g., parking at the edge of the city). Moreover, a slight reformulation of the model allows us to look at the economics of low emission zones (*LEZ*'s). Such zones have been implemented in a number of European cities – and introduction is considered in many others – in response to the European Union Clean Air Directive. A low emission zone defines an area that a vehicle is allowed to enter only if it is classified as a low emission vehicle (Wolff and Perry (2010)).[1] All high-polluting vehicles are banned from driving into the *LEZ*, or they pay a fee for non-compliance. Low emission zones can be viewed as combining standards with non-compliance fees.

Of course, there is a large existing literature on general policies to reduce traffic externalities, but these studies either do not specifically focus on urban transport, or they do not cover the wide range of instruments available to tackle pollution-type externalities in cities[2]. This is the purpose of the current paper. We develop a simple model with fixed location that allows us to analyze the optimal use of the broad set of policy instruments referred to above. As in many cities a large share of all transport externalities comes from incoming traffic (such as commuting traffic) and from pure transit traffic (i.e., through traffic that has neither origin nor destination within the city), the model distinguishes between different types of transport users. Moreover, we compare optimal policy choices from the perspective of both the urban and the federal (or national) government. It is assumed that, unlike the urban government, the federal authority takes into account the welfare of all road users, including that of incoming and transit traffic. We study

---

[1] Currently, the focus of low emission zones is mainly on reducing small particulates. To date, LEZ's have been implemented in 152 cities in nine EU countries. Germany and the Netherlands have been particularly active in establishing LEZ's. (Wolff and Perry (2010), Annema and Proag (2011)).

[2] For example, congestion pricing has been extensively studied by Arnott, de Palma and Lindsey (1993), Verhoef, Nijkamp and Rietveld (1996), De Borger and Proost (2001), Parry and Bento (2001), and Van Dender (2003). Optimal emission policies towards vehicles were analyzed in, among others, Fullerton and West (2002), Fischer, Harrington and Parry (2007) and Proost (2011). The use of fuel taxes to cope with transport externalities has been investigated by Parry and Small (2005) and Bento, Goulder, Jacobsen and Von Haefen (2009).

the possible distortions implied by decisions of the urban government and investigate the need for federal intervention. Note that, in order to focus on other types of externalities in cities, the model we develop often ignores congestion; this allows us to simplify expressions and to obtain highly transparent results. However, most results continue to hold in the presence of congestion, see below.

The main results of this paper are easily summarized. First, compared to the federal optimum, we show that the urban government will over-invest in externality-reducing infrastructure whenever this infrastructure increases the generalized cost of transit traffic. This implies that it will over-invest in 'capacity reduction' through road bumps, taking away road capacity for biking paths or walkways, red light phasing, etc. Second, comparing fees (for example, parking fees) or road tolls with imposing emission standards on vehicles, the urban government is more favorable towards tolls than the federal government. When implementing low emission zones, the urban government will set both the fee for non-compliance and the standard at a higher level than the federal government. At sufficiently high levels of through traffic, the urban government will prefer imposing a fee or toll only, even when a formal low emission zone would be the optimal federal policy to pursue. Third, we show that the urban government will underinvest in bypass capacity compared to the federal optimum. If it has the authority to levy an urban toll, we show that it will only invest in bypass capacity when it is allowed to also toll the bypass capacity and earn extra toll revenues that exceed the investment cost. Fourth, we show that, if there is uncertainty about the level of the locally generated externality and through traffic is important, implementing the optimal federal policy requires financial incentives to induce honest reporting of externalities by the city. Finally, if the high transaction cost prevents the federal government from imposing the social optimum on the city, it may in fact be better to leave the city government unregulated than not to allow the city to use the instrument at all. This will be the case if externalities are important, there is not too much transit, and the externality-reducing investment is very effective at reducing the external cost.

The paper has the following structure. In a first section, we present a taxonomy of measures governments can use to tackle inner city pollution-type traffic externalities. In Section 2, we describe the basic version of the model. We use the model in Section 3 to analyze optimal externality-reducing investment decisions by an urban and a federal government. Section 4 then explicitly zooms in on the issue of implementing low emission zones in cities, combining

emission standards and fees for non-compliance. The economics of investing in bypass capacity from the urban and federal perspective are discussed in Section 5; moreover, this section considers the he extension of the model to take account of congestion. The need for federal intervention in urban transport policies is analyzed in Section 6, explicitly considering issues of uncertainty and imperfect information. A final section concludes.

1. A taxonomy of externality-reducing measures

In a first-best world, the optimal way to reduce the emissions of urban traffic is a combination of a reduction in the emissions per kilometer (making cars cleaner) and a decline in the number of kilometers driven. It is well known that an optimal emission tax equal to the marginal external damage implements the first-best combination (Fullerton and West (2002), Proost (2011)). In deciding on the car type that determines the emissions per kilometer, it leads users to choose the car type that minimizes the social cost per kilometer; this includes the driving cost per kilometer plus the remaining emission damage. Moreover, the appropriate emission tax makes sure the number of kilometers driven is such that, at the optimum, the unit cost of driving equals the marginal social cost. However, since emission taxes are difficult to implement, we focus in this paper on policies that are second-best[3].

In Table 1, we present a taxonomy of policy measures this paper deals with. The rows indicate the different measures, the columns give the effect of the policy instruments on a number of relevant indicators such as transport volumes, external (pollution-type) cost per car kilometer, etc. Three characteristics of externality-reducing investments deserve to be mentioned in view of the analysis in Section 3 below: (i) Is the instrument costly to implement? (ii) Does it affect the generalized user cost of traffic? (iii) Does it reduce the external pollution cost per vehicle kilometer? For example, some instruments reduce the external effect per vehicle but they raise the generalized cost of traffic by reducing the speed of vehicles (new traffic lights, bumps, bike paths, etc.); other instruments (e.g., quiet asphalt, noise walls) do not increase the user cost of car transport. Although most measures are costly to the government, some policies have a negligible cost but they do reduce the speed of vehicles (e.g., speed limits or increasing the red

---

[3] When the population of car users is homogenous, the first-best can also be implemented using an emission standard on cars and a mileage tax that is function of the remaining emissions (Fullerton and West (2002)). We briefly discuss this case in Section 3 below.

phase of traffic lights) or they increase other components of the user cost of vehicles (e.g., emission standards for cars).

	Reduces traffic volume in city	Speed reducing effect	Requires large public investment	Reduction of external cost per car kilometer	Impact on urban traffic by the local population
Tolls	+	0	0	0	+
Noise walls, investment in quiet asphalt	0	0	+	+	0
Speed restriction, increasing the red phase of traffic lights	+	+	0	+	+
New traffic lights, road bumps, etc.	+	+	+	0	+
Emission standards for cars	+	0	0	+	+
Low emission zones	+	0	0	+	+
Bypass capacity	+	0	+	0	+

Table 1. Taxonomy of policy measures that address external effects of through traffic (Legend: 0= no (or negative) effect; +: positive effect)

2. Reducing pollution-type transport externalities: structure of the model

In this section, we present the model that will be used to study a variety of policies urban governments can use to reduce urban traffic externalities of the pollution-type. Unlike congestion-type externalities, pollution (and noise) does not have feedback effects on the demand for transport. As our focus is on such externalities that are separable in the utility function, congestion is ignored in most of the analysis that follows. Introducing it complicates the derivation of the results considerably without affecting the qualitative results. Also note that the model of this section is able to capture most of the policies we considered in Table 1, but there is one major exception. The possibility of investing in an urban bypass requires a somewhat

4

different methodology; it is studied in a separate section below (Section 5). There we also briefly report on extension of the model to capture congestion-type externalities.

The city network is used by local traffic (inhabitants of the city) and by incoming and through traffic. The welfare of local users is a major concern of the city's policy-makers. Demand for incoming traffic and through traffic comes from individuals that live outside the city. It is assumed that they do not participate in the local decision-making process, so that the urban government cares for them only because they pay local taxes and they cause external costs in the city[4]. We distinguish the demand by local users (denoted Y) from the demand by incoming and through traffic (X). Demand for both types of traffic is a downward sloping function of the generalized price P. Inverse demands are written as $P^Y(Y), P^X(X)$, respectively.

The total external cost is specified as a function of the total traffic volume and the externality-reducing measures taken by the urban government (such as noise walls, speed restrictions etc.). Specifically, we write the total external cost E associated with a traffic flow $V=Y+X$ as:

$$E=\Psi\{e(1-\alpha z)(V)\} \tag{1}$$

In this expression, e is the externality level (say, emissions) per unit of traffic (per trip, per kilometer) in the absence of externality-reducing measures by the city government. Furthermore, z is the level of investment in externality-reduction. This investment is normalized such that an investment z reduces the per unit externality from e to $(1-\alpha z)e$. The coefficient α captures the effectiveness of the investment in affecting the externality per kilometer[5]. Finally, total external cost is not necessarily strictly proportional to total emissions $e(1-\alpha z)(V)$. Instead, we assume $\Psi'>0$, $\Psi''\geq 0$; in other words, E is weakly convex in total emissions.[6]

---

[4] We take a medium term view of the city: location of households and firms are assumed to be fixed. When relocation is possible, the policy mechanisms may be different as agglomeration economics and interjurisdictional competition start to play a role. See Glaeser (2008) for a review.

[5] It would be more realistic to assume that, for large z, the externality converges to some minimum attainable level. However, the linear specification strongly facilitates the technical derivations, and it has no implications for the qualitative results.

[6] Our assumptions on the environmental damage function are justified for most pollutants, but not for all. For example, some secondary pollutants (tropospheric ozone) are known to have strong non convexities (Mayeres, Proost and Miltz (1993)); in that case only an elaborate air pollution model that tracks diffusion and chemical transformation can detect effective policies. Some accident externalities may also not fit our simple dose-effect relationship.

Investment in externality reduction z may have cost implications for both the government and the car user. First, the cost C to the government is assumed to be independent of the traffic flow. It is given by the following quadratic expression

$$C(z) = \frac{b}{2} z^2 \tag{2}$$

Note that for policy measures that do not require public investment (for example, imposing speed limits), b may be zero. Second, the effect of investment in externality-reduction on the user is captured by the generalized user cost. Both for local and incoming or through traffic in the city this generalized cost is assumed to be given by:

$$G(z, \tau_C) = c + \frac{1}{2} dz^2 + \tau_C \tag{3}$$

In (3), c is the generalized (money plus time) user cost in the absence of externality-reducing investment and assuming there are no tolls. As explained above, given our focus here on pollution-type externalities, the generalized user cost of using the city road is assumed independent of traffic levels; congestion is ignored. However, for many policy measures captured in Table 1, the generalized cost does depend on investment z in externality-reducing infrastructure. For example, traffic lights and road bumps reduce travel speed and accident risks, even if there is no congestion. The relation between generalized cost and investment z is assumed to be quadratic. Finally, the generalized cost can also include a toll τ_C for the use of the city infrastructure. We use a broad interpretation of such a local toll: it includes a standard road toll or congestion charge, but it may also capture parking fees, etc.

Three important parameters govern the range of policy instruments reported in Table 1. First, the parameter α captures the effectiveness of the externality reduction per unit of car use (emissions per vehicle km, noise per vehicle km etc.). Second, the importance of the investment cost born by the public budget associated with z is measured by the parameter b. Third, the effect of the policy measure on the generalized cost of a trip is given by the parameter d. Together they allow much flexibility in describing different measures. For example, some measures (e.g., long red phase at traffic lights) may have $d>0$ but $b=\alpha=0$. They are not costly to implement and leave the external cost per unit of traffic unaffected, but the total volume of traffic decreases through an increase of the generalized cost. Other measures not only increase the generalized cost but also reduce the external cost per unit of traffic (hence $\alpha>0$; for example, think of speed

restrictions) or they may require substantial public investment (then *b>0*; for example, consider noise walls). Finally, for some investment (e.g., quiet asphalt), the generalized cost will be independent of investment *z*. This is easily captured by imposing *d=0*.

In this paper, welfare for the urban local government will be measured as net consumer surplus for local users plus (in the case a toll is imposed) toll revenue, minus total external costs and the cost of externality-reducing investment. Welfare for what we call the 'federal' government also takes into account the welfare of through traffic. We assume that the valuation of the external damage by the local government is identical to the valuation by the federal government. We return to this assumption later (see Section 6). Lastly, note that we disregard the political process by which decisions are made. It is implicitly assumed that all effects of policies are known ex ante and that the population is homogeneous; as a consequence, all inhabitants share equally in all positive and negative effects of implemented policies.[7]

3. Optimal choice of externality-reducing measures

In this section, we study optimal externality-reducing policies to tackle inner city traffic externalities. We do this from the viewpoint of the urban government, and we compare the outcomes with those preferred by the federal authority. We first deal with policy measure *z* as the only instrument that is optimally chosen. Although tolls or parking fees may be in place, they are considered exogenous in the analysis. Then we reconsider optimal decisions on *z* in the presence of an optimal urban road toll.

3.1. Optimal investment in an externality-reducing measure

We first study the decisions of the urban local government; towards the end of this subsection, we turn to the optimal choices from the viewpoint of the federal authority.

To determine its optimal investment in externality-reduction *z*, we assume the urban government solves the following problem

$$\underset{z}{Max} \quad \int_0^Y P^Y(y)dy - \left[c + \frac{1}{2}dz^2 + \tau_C\right]Y + \tau_C(Y+X) - \Psi\left\{e(1-\alpha z)(Y+X)\right\} - \frac{b}{2}z^2 \quad (4)$$

In (4), the first two terms together capture the net consumer surplus of local traffic (gross surplus minus total generalized cost), and the third term is toll (or parking fee) revenue on local and on

---

all incoming and through traffic[8]; the toll τ_C is assumed to be exogenous. The last two terms capture external costs and the cost of investment, respectively. Note that the formulation of the per-unit externality as $e(1-\alpha z)$ implies that z is not only constrained to be non-negative, there is also a maximal possible reduction level: $0 \leq z \leq \dfrac{1}{\alpha}$.

The first-order conditions associated with problem (4) are analyzed in Appendix 1[9]. An internal solution implies the condition

$$\Psi'\alpha e(Y+X) + \{\tau_C - \Psi'e(1-\alpha z)\}\left[\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P}\right]dz = bz + dzY \qquad (6)$$

The left-hand side captures the benefit of extra investment z. It reduces the total externality level of a given traffic flow (first term), and it reduces demand. The reduction in demand leads to lower external costs but also reduces toll revenues (second term); the net benefit is $-\{\tau_C - \Psi'e(1-\alpha z)\}$. The right-hand side of (6) is the overall marginal cost of investment z. It consists of the monetary cost of extra investment plus the increase in user cost for all local users Y, caused by the externality reducing measure.

Expression (6) is quadratic in z, and an explicit solution is not very informative. Using the implicit function theorem does show that, under plausible conditions, the optimal investment in externality-reduction is an increasing function of the traffic volume, of the unit externality e, and of the effectiveness of the investment as captured by α; it is a declining function of the investment cost b and the contribution of the investment to the generalized user cost d.

Some special cases facilitate the interpretation. For example, assume that the measure is of the noise wall type. It affects the externality level per kilometer and is costly to implement, hence we have $\alpha > 0, b > 0$; however, the volume of traffic is not affected ($d=0$). Using (6), the optimal value of z tells us "how good the noise walls should be". We find:

$$z^u = \frac{\Psi'\alpha e(Y+X)}{b}$$

---

[8] Toll revenues are to be understood in a broad sense: they may include parking fees and local sales taxes on traffic related goods (gasoline) etc.

[9] Corner solutions are briefly considered in Appendix 1 as well. It is shown that, for an investment measure that reduces the externality per vehicle ($\alpha>0$), a corner solution with zero investment ($z=0$) can never be optimal. A solution where the city uses the instrument z to the full extent possible may well be optimal; this will be the case if, evaluated at $z = 1/\alpha$, the marginal benefit of investment still exceeds the marginal cost (see Appendix 1).

It is increasing in traffic levels and in the severity of the externality, but it decreases in the government budgetary cost b. Another special case arises when demand is totally inelastic; then we have:

$$z^u = \frac{\Psi'\alpha e(Y+X)}{b+dY} \tag{7}$$

The optimal investment level is lower for measures that strongly affect the generalized cost of local traffic (large d).

Let us now turn to optimal decisions from the perspective of the federal authority that takes a more global view, in the sense that its objective function also captures the welfare of through traffic X. Its optimal investment maximizes

$$\int_0^Y P^Y(y)dy + \int_0^X P^X(x)dx - \left[c + \frac{1}{2}dz^2 + \tau_C\right](Y+X) + \tau_C(Y+X) - \Psi\{e(1-\alpha z)(Y+X)\} - \frac{b}{2}z^2 \tag{8}$$

with respect to z. As mentioned before, we assume the evaluation of external costs is the same for the federal authority as for the city. The first-order condition for z is:

$$\Psi'\alpha e(Y+X) + \{\tau_C - \Psi'e(1-\alpha z)\}\left[\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P}\right]dz = bz + dz(Y+X) \tag{9}$$

The only difference with the optimal condition for the local city government (compare (9) and (6)) is that the cost of externality-reducing measures now also includes the marginal welfare loss for transit traffic (see the term dzX in (9)). The federal authority will therefore invest less in z than the city; we have $z^f < z^u$, where z^f stands for the federally optimal value. Note from (9) that these results hold, unless the measure z does not affect the generalized cost ($d=0$). We have therefore shown the following proposition.

Proposition 1. Compared to the federal social optimum, the city government over-invests in externality-reducing infrastructure whenever this infrastructure increases the generalized cost of transit traffic.

The intuition is clear: the city government ignores that higher investment z raises the user cost for through and incoming traffic X of passing through the city. The implication is that, compared to the social optimum, a city government will invest too much in traffic lights, road bumps etc. In

the case of policy measures that do not raise the generalized cost of transit traffic (like for example noise walls) we have that $d=0$, so that the city government will invest the efficient level.

3.2. Externality-reducing investment and optimal road tolls

How does the availability of an optimal road toll affect investment in z? Obviously, optimal tolls have been studied several times before. In line with De Borger, Dunkerley, Proost (2007) and Ubbels and Verhoef (2008), we expect the city government to "tax export" and to set the toll higher than what is socially optimal. This is easily shown to be the case. The optimal toll from the city's perspective maximizes the objective function given in (4) with respect to τ_C. We find, for a given level of z:

$$\tau_C^u = \Psi'e(1-\alpha z) - \frac{X}{\dfrac{\partial Y}{\partial P} + \dfrac{\partial X}{\partial P}} \tag{10}$$

The urban government (hence the superscript u) charges more than the marginal external cost that remains at investment level z; moreover, the toll rises in through traffic X.

Similarly, the socially optimal toll maximizes (8) with respect to τ_C. We easily derive

$$\tau_C^f = \Psi'e(1-\alpha z) \tag{11}$$

From (10) and (11), it immediately follows that the city sets the toll too high from a social perspective.

What happens if policy makers jointly use externality-reducing investment and tolls to control inner-city externalities? Substituting the optimal toll rule (10) in (6), we find that the city government's optimal investment z is given by:

$$z^u = \frac{\Psi'\alpha e(X+Y)}{b+(X+Y)d} \tag{12}$$

Using the federally optimal toll rule (11) in their optimal investment rule (9) yields the same optimal investment rule[10]. It follows that, conditional on a given transport volume, the preferred investment levels of z would be identical for the city and federal governments[11]. But because the city charges higher tolls, conditional on z, the lower transport volume reduces both the extra

---

[10] Combining optimal investment z with an optimal toll, it is easy to show that the federal optimum achieves the first-best. For example, consider an emission standard ($b=0$) combined with a toll. Our results then reproduce the first-best reported in Fullerton and West (2002) or Proost (2011).

[11] This result is analogous to results on welfare optimal quality choice in models of optimal pricing and quality in industrial organization (see, for example, Spence (1975), Brueckner (2004)).

benefit and the extra cost of investment in *z*. It is easy to show that, if the city toll is higher than the federal toll, this implies a lower (rather than higher) level of *z* for the city. So the city under-invests in externality-reducing activities but charges a high toll. We summarize findings in the following proposition.

Proposition 2. Suppose a toll on the urban road is the only instrument used. Then, compared to the social optimum, the city government sets the urban toll too high. If governments combine tolls with externality-reducing investment, the high toll leads the urban government to invest less, rather than more, in externality-reducing investment compared to what is socially optimal.

4. Standards, fees and Low Emission Zones

In this section, we use the model of the previous sections to study urban and federal incentives to implement low emission zones (*LEZ's*), a combination of a standard with a fee for noncompliance, in urban areas. Recent European legislation strongly advocates cities to implement such zones, and a number of cities have actually done so (Wolff and Perry (2010)). Although some cities have restricted the application to trucks, we here consider the possibility of implementing low emission zones for cars. A *LEZ* effectively means that road users have the choice between either complying with the urban pollution standard, or paying a fee for noncompliance. This combination of an emission standard with a system of fees for non-compliance makes it a second-best policy by definition: a standard without a mileage tax makes sure the car becomes cleaner, but the remaining emission damage is not priced; a fee introduced per kilometer or per trip only affects total mileage but does not induce car users to choose cleaner cars.

To study emission standards and *LEZ*'s, note that the model of the previous sections can easily be adapted to deal with these policy instruments. To do so requires a slight reinterpretation of the model on three accounts. First, we let z capture an emission standard; complying with the standard reduces the emission of the car per kilometer from *e* to $e(1-\alpha z)$. Second, the implementation cost of the standard to the government is assumed to be zero; in the notation of our model, we therefore have *b=0*. Third, for car users, there is a cost of complying with the standard. Strictly speaking, this compliance cost is likely to be a fixed cost (investing in a cleaner

11

car). However, to maximize the analogy with the model developed above, we interpret the compliance cost as a variable cost that raises the generalized cost of road use in the urban area. A sufficient assumption that makes a fixed cost increase equivalent to an increase in the generalized user cost per kilometer is that cars have fixed annual mileage in the urban area. This is of course a strong assumption, but it substantially facilitates the analysis of *LEZ*'s within the framework of the model developed above. It allows us to reinterpret $\frac{1}{2}dz^2$ as the per kilometer compliance cost. The more stringent the standard, the higher the compliance cost.

In the remainder of this section we work in two stages. In a first step, we will assume the compliance costs of all road users to be the same. As we will see below, in this case the problem of introducing a low emission zone effectively boils down to choosing between an optimal standard and an optimal fee. In a second step, we allow for compliance cost differences between local and through traffic; this gives low emission zones as an intermediate case between a standard and a fee.

4.1. Identical compliance costs for all users: standards versus fees

We start with a simple model with homogeneous users that all face the same compliance cost. Under these assumptions, either all users will comply with the standard and hence face a higher user cost, or no one will comply; in that case, users are subject to a fee for non-compliance for each kilometer driven within the low emission zone. Therefore, the problem of low emission zones with identical compliance costs for all users boils down to choosing between an optimal standard and an optimal fee. If the standard and the fee are set in such a way that everyone complies, the problem is the same as that of determining the optimal standard (with the fee set prohibitively high). If the combination of the standard and fee imply that no one complies but everyone prefers to pay the fee, the problem is identical to finding the optimal fee (and setting the standard prohibitively stringent). In the latter case, making abstraction of possible psychological aspects (such as stigma associated with not complying and being fined), the fee plays exactly the same role as a toll.

As this assumption does not affect the qualitative results but drastically simplifies their derivation, we assume throughout this section that the external cost function is just given by $E = e(1-\alpha z)(Y+X)$; in other words, we assume $\Psi'=1$.

The federal government

Let us focus on the federal government. We consecutively consider the case of an optimal standard and an optimal fee.

First, in Appendix 2 we show that the optimal standard that makes all users comply satisfies the following condition

$$(\frac{\alpha e}{d} - z_s^f)(Y_s^f + X_s^f) = e(1 - \alpha z_s^f)\left[(\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P})z_s^f\right] \tag{13}$$

The subscript '*s*' refers to the use of the standard, the superscript '*f*' refers to the values selected by the federal authority.

To interpret (13), assume first that the demand for traffic is totally inelastic. The right hand side then equals zero, and the optimal standard is

$$z_s^f = \frac{\alpha e}{d}$$

This in fact achieves the first-best. To see this, note that -- as mileage cannot be affected -- in first-best all one can do is to have cleaner cars that minimize the sum of the resource cost of a clean car plus the remaining emissions. Doing so, i.e., minimizing

$$\left[c + \frac{1}{2}d(z_s^f)^2 - e(1 - \alpha z_s^f)\right]$$

indeed immediately gives $z_s^f = \frac{\alpha e}{d}$. Introducing elastic demands, it then directly follows from

(13) that $z_s^f > \frac{\alpha e}{d}$. As there is no mileage tax to tax remaining emissions, the optimal standard has to be stricter with price-sensitive demand. Finally, note that (13) immediately implies that an exogenous increase in local demand *Y* or in transit demand *X* has exactly the same effect on the optimal federal standard. Loosely speaking, the relative importance of through traffic does not matter for the optimal standard.

Second, turn to the case where no one complies. The optimal federal fee, denoted F^f, then simply equals marginal external cost (see Appendix 2): we have $F^f = e$. The fee acts like a road toll, a 'product' tax on kilometers; it reduces car use but has no incentive to adopt cleaner cars.

We are interested in comparing welfare when an optimal standard is used with welfare when having an optimal fee. To do so, denote federal welfare with the optimal standard and the optimal fee as W_s^f, W_F^f, respectively. Assuming linear demands to keep the analysis transparent, we then show in Appendix 2 that the welfare difference can be written as:

$$W_s^f - W_\tau^f = \frac{1}{2}\alpha e z_s^f \left(Y_s^f + X_s^f\right)$$
$$+ \frac{1}{2}\left(\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P}\right)\left\{\left[e - \frac{1}{2}d\left(z_s^f\right)^2\right]^2 + e(1-\alpha z_s^f)d\left(z_s^f\right)^2\right\} \tag{14}$$

Begin again with assuming zero elastic demand. In that case, the last term in (14) drops out so that $W_s^f - W_\tau^f > 0$: a standard is always better. This makes sense: a fee does not reduce demand, so that all one can achieve is to make cars cleaner per kilometer. Next consider the case with elastic demands. Elaborating the final term on the right hand side then immediately implies that a standard remains superior as long as the costs of a cleaner car are low (small d) and a clean car is highly effective in reducing emissions per unit (α is large, so that $(1-\alpha z_s^f)$ is low and $\alpha e z_s^f$ is high). An optimal federal fee or toll can only be welfare-superior to a standard if the latter does not reduce emissions per mile much further (so that there remains a large un-priced externality) and the cost of cleaner cars is substantial.

Finally, it is easily verified that (14) implies that the relative desirability of standards versus fees or tolls does not depend on the share of transit traffic in total demand: the optimal standard does not depend on the transit share, and an exogenous increase in transit traffic affects the right hand side of (14) in exactly the same way as an increase in local demand.

The urban government

We now turn to the choices of standards and fees from the perspective of the urban government. First, if the urban authority sets the standard and fee so that everyone complies, the optimal standard satisfies (see Appendix 2)

$$(\frac{\alpha e}{d} - z_s^u)(Y_s^u) = e(1-\alpha z_s^u)\left[(\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P})z_s^u\right] - \frac{\alpha e}{d}X_s^u \tag{15}$$

The subscript 's' refers, as before, to the standard, the superscript 'u' refers to the values at the optimum from the urban viewpoint[12].

For totally inelastic demand, we would have

$$z_s^u = \frac{\alpha e}{d} \left(\frac{Y_s^u + X_s^u}{Y_s^u} \right) > \frac{\alpha e}{d}$$

so that the urban government's standard will be more stringent than the federal one. Note here the crucial importance of transit: the optimal urban standard is increasing in the share of transit, unlike the federal one. This continues to hold in the case of elastic demand, see (15). As in the federal case, elastic demand makes the standard more stringent.

Next, if the urban government opts for an outcome whereby no one complies with the standard but everyone prefers to pay the fine or fee, the optimal fee satisfies (Appendix 2):

$$F^u = e - \frac{X_F^u}{\dfrac{\partial X}{\partial F^u} + \dfrac{\partial Y}{\partial F^u}} \tag{16}$$

The notation Y_F^u, X_F^u refers to the values of Y and X when the urban government uses an optimal fee. The fee increases with transit traffic.

Denote urban welfare with the optimal standard and optimal toll as W_s^u, W_F^u, respectively. We then show in Appendix 2 that, given linear demands:

$$W_s^u - W_F^u = \left[\alpha e z_s^u - d \left(z_s^u \right)^2 \right] Y_s^u + \frac{1}{2} \frac{\partial Y}{\partial P} \left[F^u - \frac{1}{2} d \left(z_s^u \right)^2 \right]^2 + e Y_F^u$$
$$- e(1 - \alpha z_s^u) X_s^u - (F^u - e)(X_F^u - Y_s^u) \tag{17}$$

To interpret this expression first note that, if there were no local demand, (17) reduces to

$$W_s^u - W_F^u = -e(1 - \alpha z_s^u) X_s^u - (F^u - e)(X_F^u)$$

As $(F^u - e) > 0$, this means that the city necessarily prefers the fee over the standard. Second, although an increase in the relative importance of transit demand raises the optimal standard and the optimal fee, (17) suggests that the more important is through traffic X (for given local demand), the more likely the city prefers the fee over the standard. Third, just as for the federal government, a fee becomes more favorable when α is small (no effective emission reduction) and d is large (high compliance costs).

---

[12] If X were zero, the urban optimum is obviously the same as the federal one (compare (15) and (13)).
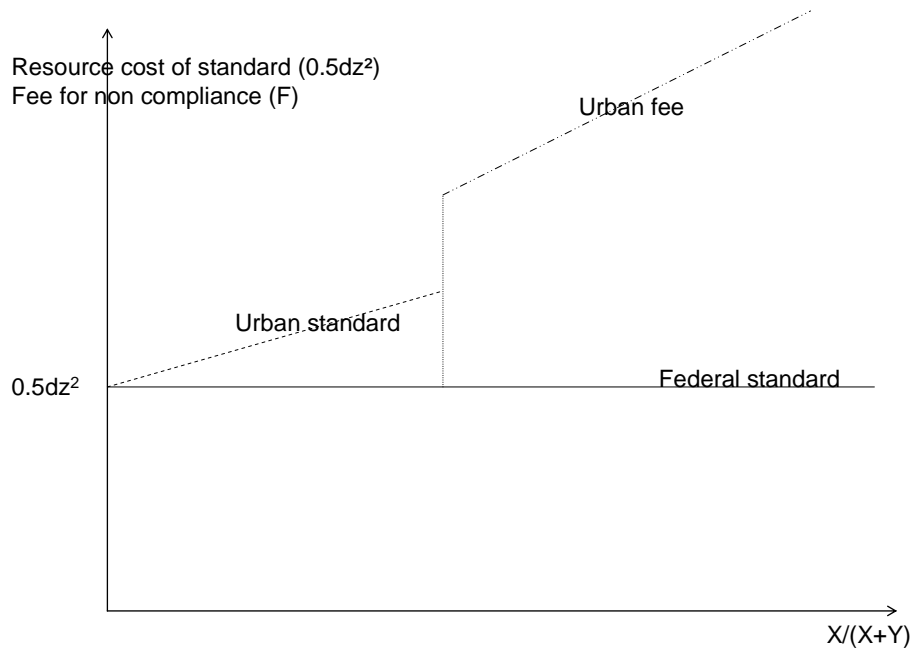
Comparing the urban and federal outcomes

Our analytical results point at two important qualitative conclusions. First, the optimal fee for the urban government rises with the importance of through traffic, and it exceeds the fee that is optimal from a federal perspective. Second, whereas the urban government strongly favors tolls over standards when transit is relatively important in total traffic, the share of transit does not affect the federal government's choice of fee versus standard. We summarize in the following proposition.

Proposition 3. The urban government will set both the standard and the fee above the level preferred by the federal government. Moreover, cities with a lot of commuters or through traffic will favor a fee over a standard, even though from the federal perspective standards may be better.

Figure 1 illustrates the case where the federal optimum turns out to be a standard. The vertical axis measures the additional cost per vehicle-kilometer resulting from a given policy. This can be a fee for non-compliance or the cost of a standard for cleaner cars. The horizontal axis represents the share *X/(X+Y)* of transit traffic *X* in total traffic *(X+Y)*. The horizontal line at $0.5dz^2$ represents the additional resource cost per kilometer of the greener car that meets the federal standard. As argued above, the optimal federal standard does not depend on the relative importance of transit. Consider now the option preferred by the city government. If there is no transit, the urban and federal optimal standards coincide. Starting at the federal value in the absence of transit traffic, the preferred standard for the urban government increases when the share of transit increases (according to (14)). This is represented by the dashed line. When transit becomes sufficiently important, the city will no longer be interested in a standard but prefers a fee. This fee exceeds the marginal external damage and also increases in the share of transit (see equation (16)).

Figure 1: fees and standards with identical compliance costs

Resource cost of standard (0.5dz²)
Fee for non compliance (F)

Urban fee

Urban standard

Federal standard

0.5dz²

X/(X+Y)

4.2. Low emission zones

Consider now the case of unequal compliance costs; as will become clear, this gives rise to low emission zones as intermediate between standards and fees. Specifically, assume that, in the case of compliance with a standard z, the generalized costs per kilometer for local and for transit users are

$$(c + \frac{1}{2}d_Y(z)^2), (c + \frac{1}{2}d_X(z)^2),$$

respectively. We will assume that it is easier to comply for local than for transit users, hence $d_Y < d_X$. One reason may be that transit users drive a larger fraction of all kilometers outside the urban area, raising the compliance cost per kilometer driven in the city more than for locals. Note that in case of non-compliance the generalized costs are $(c + F)$ for all users, where F is the fee to be paid per kilometer in case of exceeding the emission standard.

As all local users are assumed to be identical, they will either all comply or none will comply; the same holds for transit users. Local users will comply if

$$\frac{1}{2}d_Y(z)^2 < F \quad \rightarrow \quad z < \sqrt{\frac{2F}{d_Y}} \tag{18}$$

17

Similarly, transit users comply provided

$$\frac{1}{2}d_X(z)^2 < F \quad \rightarrow \quad z < \sqrt{\frac{2F}{d_X}} \tag{19}$$

In principle, there are many possibilities to consider. To see this, it is clear that -- depending on the parameter values -- there are three possibilities for the federal optimum. It is possible that the federal optimum is (i) a standard that is satisfied by all users (the fee is so high that everyone complies); (ii) a standard to which locals comply but transit users prefer to pay the fee; (iii) a fee only (the standard is so strict that no one complies). Conditional on each of the three possible federal optima, we can analyze the urban government's optimal behavior[13]. However, the main intuition on the difference in behavior by urban and federal governments can be gained by looking at one particularly relevant case, viz. the situation where for the federal government a low emission zone is optimal. We easily show that the urban government will prefer a different policy than the federal authority. In fact, if transit is sufficiently important, the urban government will prefer to set the standard so high that all users prefer to pay the fee.

To show these statements, suppose that from the viewpoint of the federal government a *LEZ* is optimal: in other words, the optimal federal fee and standard are such that it is optimal to have the locals complying with the standard, but transit prefers not to comply and to pay the fee. In Appendix 3, we formally show that this internal optimal solution for the federal government is described by

$$(\frac{\alpha e}{d_Y} - z_s^f)Y^f = e(1 - \alpha z_s^f)\frac{\partial Y}{\partial P}z_s^f \tag{20}$$
$$F^f = e$$

Moreover, by assumption, the solution satisfies (see (18)-(19))

$$\frac{1}{2}d_Y(z_s^f)^2 < F^f = e < \frac{1}{2}d_X(z_s^f)^2 \tag{21}$$

Why will the urban government want to pursue a different policy? The reason is that with sufficiently high transit demand, contrary to the federal case, no internal solution will exist for the urban government. To see this, first note that an optimal internal solution for the urban government would be described by (see Appendix 3):

---

[13] A detailed analysis of the different possibilities is available from the authors.

$$(\frac{\alpha e}{d_Y} - z_s^u)Y^u = e(1-\alpha z_s^u)\frac{\partial Y}{\partial P}z_s^u$$

$$F^u = e - \frac{X^u}{\frac{\partial X}{\partial F}}$$

(22)

If transit were zero, (20) and (22) give the same solution. With positive transit, however, (20) and (22) imply that the city and the federal level will choose the same standard z, but the city will put the penalty fee for non-compliance higher[14]. We have $z_s^f = z_s^u$; $F^u > F^f$. Further note that for an internal solution to the urban government's problem, the standard and fee should satisfy:

$$\frac{1}{2}d_Y(z_s^u)^2 < F^u = e - \frac{X^u}{\frac{\partial X}{\partial F}} < \frac{1}{2}d_X(z_s^u)^2$$

It is clear, then, that for a sufficiently high transit level this condition will not be satisfied. The city will face a corner solution and it will have to reduce the fee and/or raise the standard in order to prevent transit to opt for the standard. Eventually, of course, at even higher transit levels the city will find it optimal to have a fee only, and set the standard prohibitively high.

The situation is shown in Figure 2. It gives both the federally optimal standard for the city inhabitants and the federally optimal fee for transit traffic; both are independent of the relative importance of transit traffic, but the fee exceeds the resource cost of the standard for the inhabitants. At zero transit, urban policy is the same as federal policy: a standard is set for inhabitants of the city. As shown above, when transit is positive, this will lead the urban government to set a higher fee for transit than the federal one (see the highest dashed curve). However, the fact that a standard is in place for locals limits the extent of the fee increase: transit will chose the standard when the resource cost of doing so is less than the fee. This implies that an increase in the transit share forces the city government to also raise the standard for the locals, because otherwise further increases in the fee are impossible. Ultimately, when the transit share is sufficiently high, the city government will raise the standard prohibitively high so as to force all car users, including city habitants, to pay the fee. In other words, at sufficiently high transit shares, the urban government prefers a fee only. This is illustrated on Figure 2. We see that the urban standard (that de facto applies only to locals and is represented by the dashed line) is

---

[14] To see this, note that for an internal solution local demand only depends on the stringency of the standard; it does not depend on the fee, as the latter applies only to transit.

identical to the federal standard for low transit shares but increases above the federal standard when the transit share increases.

Figure 2: fees and standards with differences in compliance costs
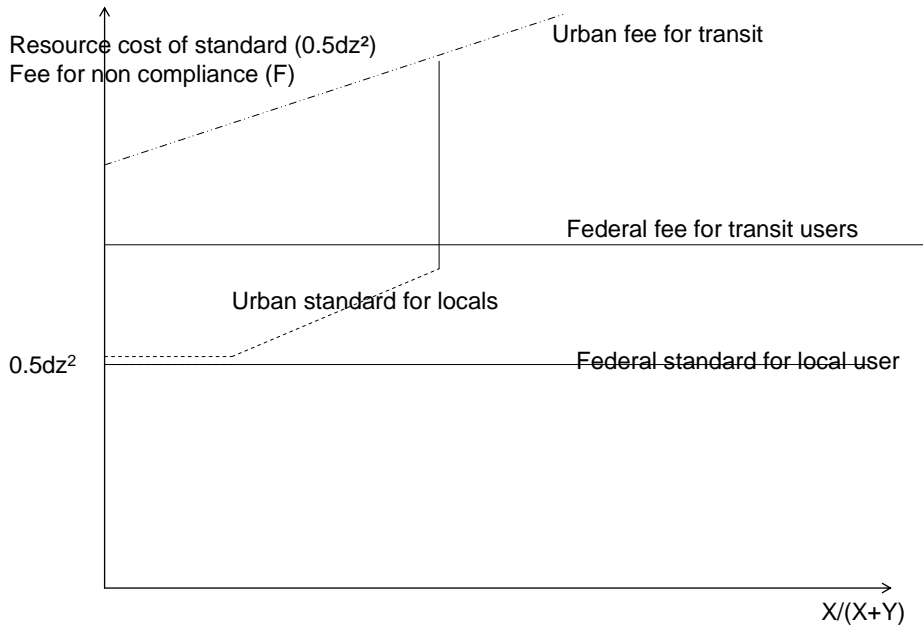
This intuitive argument suggests that, first, when a low emission zone (with locals obeying the standard and transit paying the fee) is optimal from the federal perspective, there is a range of transit levels such that the city will also prefer a low emission zone; however, both the fee and the standard will be higher than at the federal optimum. Second, at sufficiently high transit levels the city goes for a fee or toll on all users. We summarize in the following proposition.

Proposition 4. The urban government has incentives to impose too stringent standards and too high fees for non-compliance compared to the federal optimum. Moreover, it may prefer fees or tolls only when a low emission zone is federally optimal.

5. Investment in a city bypass

In the previous sections we consecutively studied possible biases in local decision making with respect to externality-reducing investments (Section 3) and policies directed towards the type of vehicles that will use the urban network (Section 4). In this section, we turn to a third category of policies, viz. investing in bypass capacity as an externality-reducing instrument for cities. Note that this does not only capture building extra capacity to get around (or under) the city center (for example, building or expanding a ring road, bringing capacity for through traffic underground by a system of tunnels, etc.), it may also include investing in signaling and guidance techniques that lead through traffic around the center.

Of course, a bypass is typically considered to reduce congestion as well as pollution in the city. However, given our focus on pollution-type externalities and to obtain highly transparent results, we initially ignore congestion in our model and work in several steps. We first analyze a bypass as a single policy instrument, then combine it with the possibility to charge an urban toll; next, we study a full set of policy instruments, including the possibility to toll the bypass itself. Finally, the implications of introducing congestion in the model are discussed towards the end of this section; a more complete treatment is provided in appendix, see below.

To facilitate the algebra, we make two simplifying assumptions. First, we assume that local traffic Y will only use the urban road infrastructure, it will never make use of the bypass capacity. Second, as in the previous section we assume $\Psi'=1$ so that the total external cost simplifies to $E = e(1-\alpha z)(X + Y)$.

5.1. Investment in bypass capacity as a single instrument

Let us first focus on bypass capacity in isolation: although an urban toll and capacity-reducing investment z may be in place, in this subsection these are treated as exogenous. To get started, assume that the current cost to use an alternative route that gets users around the city is h_0. Since there is no congestion, all through traffic will use the urban road through the city as long as the generalized cost of the urban road is below that of the initial bypass capacity; using earlier notation (see (3)), this will be the case if $G(z, \tau_C) < h_0$. We assume this is the case in the initial situation.

The government can reduce the user cost of using bypass capacity by extra investment, which we denote k. The user cost of using the bypass is given by a function $h(k)$,

where $h' < 0; h(0) = h_0$. This last equality means that if no extra investment is made in the bypass, the user cost to get around the city remains at its current level. The cost of providing the extra capacity is assumed to be quadratic: $\dfrac{fk^2}{2}$.

The urban government

Will the city invest in extra bypass capacity if it has to finance the investment out of its own funds and if it does, how much will it invest? The city government knows that local traffic *Y* will always use the urban road and that, given investment in the bypass, through traffic *X* will use the road with the lowest generalized cost. Consequently, as there is no congestion, either no or all transit traffic passes through the city. It is then obvious that only two possible optimal solutions exist for the urban authorities. The first one is to not invest in *k* at all. This yields toll revenues (if $\tau_C > 0$) on all traffic (*X+Y*), and the local population faces an external cost $e(1 - \alpha z)(Y + X)$. The second possible optimum is one where the urban government invests just enough to have all transit traffic switch to the bypass; indeed, investing even more does not yield any additional benefits. Optimal bypass investment k^u in this case is the solution of the equation

$$\left[c + \frac{1}{2}dz^2 + \tau_C \right] = h(k) + \varepsilon \tag{23}$$

for *k*, where ε is infinitesimally small so that all through traffic uses the bypass. Of course, the city now loses potential toll revenues on transit, but the external cost in the city drops to $e(1 - \alpha z)Y$.

Whether it is better to invest k^u or not to invest at all depends on the comparison of welfare for *k=0* (and all traffic (*Y+X*) going via the city) and welfare for k^u (where only *Y* passes through the city and *X* goes via the bypass). The former is given by

$$\int_0^Y P^Y(y)dy - \left[c + \frac{1}{2}dz^2 + \tau_C \right]Y + \tau_C(Y + X^{city}) - e(1 - \alpha z)(Y + X^{city}) - \frac{b}{2}z^2$$

where X^{city} is the demand for using the urban road by through traffic in the case of zero extra investment. If the city does invest in extra bypass capacity, urban welfare is

$$\int_0^Y P^Y(y)dy - \left[c + \frac{1}{2}dz^2 + \tau_C \right]Y + \tau_C Y - e(1 - \alpha z)Y - \frac{b}{2}z^2 - \frac{f}{2}(k^u)^2$$

Since there is no congestion in the model, extra bypass capacity does not affect the net consumer surplus of local users: Y is independent of k. As the toll and the level of externality reducing investment z are exogenously fixed by assumption, we then immediately see that the outcome of the comparison only depends on whether or not the savings in external costs, minus the lost toll revenues due to transit now using the bypass, outweigh the capacity costs. Investing k^u, rather than not investing at all, will be optimal provided that:

$$\left[e(1-\alpha z)-\tau\right]X^{city} > \frac{f(k^u)^2}{2} \tag{24}$$

Interpretation is easy. The city will only be interested in investing in a bypass when through traffic is very important and the urban toll it currently charges to all users of the urban infrastructure is below marginal external cost. For example, if the city does neither impose toll nor parking fees (so $\tau_C = 0$) and externality levels per unit are high, investing in bypass capacity makes sense, provided the cost of investment is not too high. If the inequality in (24) goes the other way around, the city will not invest at all in bypass capacity or in guiding transit traffic around the city center.

The federal government

Let us turn to the socially optimal investment in bypass capacity. Note that it can never be socially optimal to invest a positive amount k as long as $c + \frac{1}{2}dz^2 + \tau_C < h(k)$, because in that case no transit user would use the bypass. But from the federal viewpoint, investing more than the level that just induces all transit users to use the bypass may well be optimal: although extra investment is costly, it also reduces the generalized cost of bypass use and, therefore, it benefits bypass users X. This was ignored by the city government. Denoting the optimal capacity from the federal perspective as k^f, it then also immediately follows that $k^f \geq k^u$. The local level never invests more in a city bypass (that it finances itself) than is socially optimal, and it may well invest less.

Finally, note that if it is socially optimal to invest in extra bypass capacity (so $k^f > 0$), optimal investment k^f maximizes:

$$\int_0^Y P^Y(y)dy + \int_0^X P^X(x)dx - \left[c + \frac{1}{2}dz^2 + \tau_C\right]Y - \left(h(k)\right)X + \tau_C Y - e(1-\alpha z)Y + e_B X - \frac{b}{2}z^2 - \frac{f}{2}k^2$$

where $X = X(P), P = h(k)$, and e_B is the external cost per unit on the bypass. This yields the first-order condition

$$-h'\left[X + (h + e_B)\frac{\partial X}{\partial P}\right] = fk \tag{25}$$

The left hand side is the net benefit of extra bypass capacity. Extra capacity reduces the user cost for bypass users, and it leads to more users. The right hand side captures the marginal investment cost.

5.2. Combining an optimal city toll with bypass investment

Of course, the above story was developed for given values of z and the city toll τ_C. Now suppose that the local government has the authority to levy the optimal city toll. What does this imply for the incentives to invest in bypass capacity? We found above that the optimal city toll exceeds marginal external cost (see (10)), so that $\left[e(1 - \alpha z) - \tau_C\right] < 0$. This then implies automatically that (24) can never hold; we will necessarily have:

$$\left[e(1 - \alpha z) - \tau_C\right]X^{city} < \frac{f(k^u)^2}{2}$$

In other words, if the urban government can charge its optimal toll on the urban road, it will never invest in a bypass or in technology to guide traffic around the city. Tolling strengthens the under-investment of the local government in bypass capacity, since investing in bypass capacity would go at the expense of substantial toll revenues.

5.3. Tolling the bypass: using all instruments

Finally, let us see what happens if governments have a full set of instruments, implying that they can also toll the bypass. This is in fact not an unrealistic situation, as tolling the new capacity is often seen as a way of (partially or fully) financing the investment cost. We assume in what follows that the authorities can determine all instruments optimally: investment z, investment in bypass capacity k, and tolls on both the urban road and the bypass. We denote the toll on the bypass as τ_B. Our main interest in this subsection is to find out under what conditions the urban government has an incentive to invest in bypass capacity if it can use all instruments.

First, if building the bypass were optimal from the city's perspective it would determine the policy variables so as to:

$$\underset{z,\tau_C,k,\tau_B}{Max} \quad \int_0^Y P^Y(y)dy - \left[c + \frac{1}{2}dz^2 + \tau_C\right]Y + \tau_C Y + \tau_B X - e(1-\alpha z)Y - \frac{b}{2}z^2 - \frac{f}{2}k^2$$

Denote the optimal values as $(z^*, \tau_C^*, \tau_B^*, k^*)$. Optimal traffic levels are denoted Y^*, X^*, and optimal welfare from the viewpoint of the urban government is given by $W^u(z^*, \tau_C^*, \tau_B^*, k^*)$. Second, if not investing in bypass capacity were optimal, the values for the urban road toll and externality-reducing investment in the city, which we denote $\tilde{\tau}_C, \tilde{z}$, are the solution to

$$\underset{z,\tau_c}{Max} \quad \int_0^Y P^Y(y)dy - \left[c + \frac{1}{2}dz^2 + \tau_C\right]Y + \tau_C(Y+X) - e(1-\alpha z)(Y+X) - \frac{b}{2}z^2$$

Let optimal welfare be written as $W^u(\tilde{z}, \tilde{\tau}_C, 0, 0)$.

The question is under what conditions investing in extra bypass capacity is better than not investing at all. When is

$$W^u(z^*, \tau_C^*, \tau_B^*, k^*) > W^u(\tilde{z}, \tilde{\tau}_C, 0, 0) ? \tag{26}$$

Although a general answer to this question is not straightforward, an interesting necessary condition for this to hold is easily derived. In Appendix 4, we show that

$$W^u(\tilde{z}, \tilde{\tau}_C, 0, 0) > W^u(z, \tau_C, \tau_B, k; \tau_B X - \frac{f}{2}k^2 = 0) \tag{27}$$

where $W^u(z, \tau_C, \tau_B, k; \tau_B X - \frac{f}{2}k^2 = 0)$ is optimal city welfare under a break-even constraint for the bypass. This requires the toll revenues earned on the bypass to equal bypass investment costs. Expression (27) shows that not investing in bypass capacity is always better than investing in the bypass when the bypass has to break even. This has two implications. The first one is that a necessary condition for (26) to hold is that the optimal toll revenues on the bypass must exceed the investment cost. If the optimal bypass government so as not to exceed the investment cost on the bypass, then the urban government will never invest in the bypass.

Intuition for the previous findings is obvious. Because the urban road is also tolled, a bypass road is only interesting if it generates net revenues; if not, the urban government is better off by extracting revenues from transit traffic X via the city toll, saving on the investment cost of the bypass.

5.4. Summary of findings

We summarize our findings from this section in the following proposition.

Proposition 5.

  a. **If it is socially optimal to invest in bypass capacity or in technology to guide traffic around the city, the city government will necessarily under-invest compared to the social optimum.**

  b. **If the urban government does not have the authority to implement optimal tolls on neither the urban road nor the bypass, it will invest in bypass capacity if the external cost is large and the investment cost is low.**

  c. **If the urban government has the authority to implement its optimal toll policy on the urban road but it cannot toll the bypass, it will never invest in bypass capacity.**

  d. **If the urban government can charge both a toll on traffic using the urban road and a toll on bypass users, but it has to self-finance the bypass investment, then a necessary condition for the city to invest in the bypass is that the toll revenues on the bypass exceed the investment costs.**

This proposition has some interesting consequences. First, note that part *b.* of the Proposition implies that cities may well be willing to invest in guidance systems (low investment cost) but not in costly bypass road capacity. Second, the finding that a city will never invest in bypass capacity if it does not have the authority to impose tolls on the extra capacity is highly relevant, because many cities are only interested in building a bypass because tolling this new infrastructure yields potentially large net revenues. For example, suppose the city can impose tolls on both the urban road and on the bypass. Then a federal regulation that forces the city to invest the toll revenues on the bypass in bypass capacity will imply that the city will never construct the bypass. Moreover, it is well known that under specific constant returns to scale and homogeneity conditions socially optimal tolls are exactly cost recovering (Mohrung and Harwitz (1962), de Palma and Lindsey (2007)). This implies that, if the federal government would force the city to charge socially optimal tolls on the bypass, it will never be in the city's interest to build the bypass.

5.5. Extension of the model for congestion

Finally, note that we developed the analysis in this section in the absence of congestion. This is a severe limitation, because bypass capacity is typically not only justified on the basis of pollution concerns, but also to reduce excessive congestion in the city. And indeed, the highly transparent but somewhat extreme results obtained in this section have to be qualified when congestion is explicitly introduced on both the urban road and the bypass. To analyze this further, we extended the model to allow for congestion. Specifically, instead of assuming that all transit uses either the urban road or the bypass, we let traffic distribute itself on the two-road network according to the Wardrop principle: in equilibrium, for transit the generalized cost of the two options will equalize.

The extension to capture congestion in the model is worked out in Appendix 5. Although the modeling becomes extremely complex and the results are less transparent, one can show that some of the main insights continue to hold. For example, the urban government invests less in bypass capacity than is socially optimal. The city now has the opportunity to underinvest in bypass capacity in order to raise bypass toll revenues. Moreover, the urban government's optimal toll on the bypass implies revenues exceeding capacity costs. For more details, we refer to Appendix 5.

6. Should the federal government intervene in urban policies?

Not surprisingly, we found in the previous sections that local and federal governments prefer different policies. The question arises to what extent the distortions implied by local government policies can be corrected by federal intervention. In principle, the distortion can be avoided by having the federal government impose its optimal policy on the city.

Such federal intervention raises several issues. First, the federal level may not have the information to implement its optimal policy. For example, it may not know the true level of the locally generated externality e and it may have to design a truth-revealing mechanism that induces the urban government to provide the correct information (Van der Loo and Proost (2011)). Second, even if the federal level has perfect information about the externality e that is locally generated, implementation costs may be prohibitively high. The federal government then

faces two alternatives: it may decide not to intervene at all and let the city implement its optimal local policy, or it may simply not allow the city to use the instrument at all.

In the remainder of this section, we discuss the federal government's problems in correcting urban policy failures in more detail. Throughout, we focus on externality-reducing investment z as the only policy instrument, as considered in Section 3. However, similar analysis can be applied to the case of tolling or low emission zones. We begin with a particularly simple case in a first subsection, then we discuss the role of asymmetric information and implementation costs.

6.1. A simplified case: no need for intervention

In Sections 2 and 3, we considered two types of externality-reducing investments, those that do and those that don't affect the generalized cost. We showed in section 3.1 that in cases where the generalized cost is independent of investment, the level of z chosen by the city is welfare-optimal. A good example are noise walls: they are a public good, and the optimal level is achieved when the marginal cost equals the total marginal damage. As all the damage is local and the cost is borne by the local level (city government), the local government has the right incentives.

Proposition 6. It is always welfare improving to allow the local government to make externality-reducing investments as long as they are costless for through traffic.

6.2. Asymmetric information and federal intervention

Now look at investments that do raise the cost of both local and through traffic. Then we know that the optimal level of the investment wanted by the city exceeds that which is federally optimal (see Proposition 1). In principle, it is then welfare-improving for the federal level to impose the optimal federal level on the city. To do so, the federal level needs information on e. Suppose, however, that the federal level does not know e and depends on truthful reporting by the local authorities to implement the federally optimal policy.

To study this case, consider the following setup. Assume the true level can be high or low, denoted as e^H, e^L, respectively. The urban government can report a high or low level, denoted as $\bar{e}^i$ $(i = H, L)$. Welfare for the urban government W^u is given by

28

$$W^u(\overline{e}^i, e^j) = \int_0^Y P^Y(y)dy - \left[c + \frac{1}{2}dz^2 + \tau_C\right]Y + \tau_C(Y+X) - e^j(1-\alpha z)(Y+X) - \frac{b}{2}z^2 \quad (28)$$

In this expression, all demands are evaluated at the appropriate generalized cost. Moreover, as before, we have assumed for simplicity that total external cost is given by $E = e(1-\alpha z)V$, where $V = Y + X$ is total traffic demand. Finally, z is the federally optimal level, given the reported value of e. So $z = \left(z^f(\overline{e}^i)\right)$

Truthful reporting requires the incentive compatibility constraints

$$\begin{aligned} W^u(\overline{e}^H, e^H) &\geq W^u(\overline{e}^L, e^H) \\ W^u(\overline{e}^L, e^L) + S &\geq W^u(\overline{e}^H, e^L) \end{aligned} \quad (29)$$

where S is a possible financial transfer. The first constraint is always satisfied. As the city wants a higher level of z than the federal authority, it has an incentive to always say the truth when the true e is high. The second constraint may require a positive financial transfer: when the true e is low, the urban government may have an incentive to cheat and report that e is high.

We will investigate the incentives for the urban government to misreport the true externality level for the case of zero elastic demand; little is gained by considering the more general case. The desired level by the city for low e is then (see (6)):

$$z^u(e^L) = \frac{\alpha e^L V}{b + dY} \quad (30)$$

The federally implemented levels are in this case, for low and high reported e, respectively (see (9)):

$$z^f(\overline{e}^L) = \frac{\alpha e^L V}{b + dV} \qquad z^f(\overline{e}^H) = \frac{\alpha e^H V}{b + dV} \quad (31)$$

Note that it is a priori unclear whether honest reporting or cheating gives the highest city welfare. We have $z^f(\overline{e}^L) < z^u(e^L)$ but both $z^f(\overline{e}^H) > z^u(e^L)$ and $z^f(\overline{e}^H) < z^u(e^L)$ are possible However, using (30) and (31), it immediately follows that we have $z^f(\overline{e}^H) < z^u(e^L)$ if

$$\alpha(b + dY)(e^H - e^L) - \alpha e^L dX < 0$$

As $e^H > e^L$, this can only happen for large through traffic. So if there is much transit through the city the federally imposed investment level when the city reports high e is less than what the city wants, but it is by definition higher than what it is for low reported e. Consequently, it is in the

city's interest to cheat and to report high *e*. In that case a financial transfer will be needed to induce honest behavior. This is not necessarily the case for low transit.

This intuitive argument can easily be shown more formally. To keep the analysis transparent, we stick to the assumption of perfectly inelastic demand; no extra intuition is gained by considering the more general case. Using (28), the second incentive compatibility constraint in (29) can then, after simple algebra, be reformulated as follows:

$$S \geq \frac{1}{2}(b+dY)\left[\left(z^f(\overline{e}^L)\right)^2 - \left(z^f(\overline{e}^H)\right)^2\right] + \alpha e^L V\left[\left(z^f(\overline{e}^H)\right) - \left(z^f(\overline{e}^L)\right)\right]$$

Working out, we find

$$S \geq \left[\left(z^f(\overline{e}^H)\right) - \left(z^f(\overline{e}^L)\right)\right]\left\{\frac{1}{2}(b+dY)\left[\left(z^f(\overline{e}^L)\right) - \left(z^f(\overline{e}^H)\right)\right] + dX\left(z^f(\overline{e}^L)\right)\right\}$$

This confirms our earlier intuition. If transit *X* is small, the right hand side is negative and the inequality is automatically satisfied at *S=0*. No financial transfer is needed to induce honest reporting. If there is a lot of transit, financial transfers may be needed to induce honesty by the city.

Proposition 7. If there is uncertainty about the level of the locally generated externality and through traffic is important, then implementing the socially optimal externality-reducing investment policy requires financial incentives to induce honest reporting of externalities by the city. No such transfers are needed if through traffic is unimportant.

6.3. Should the federal government limit externality-reducing investment by cities?

So far we ignored implementation costs of imposing optimal policies. Of course, such costs may be so high that it may not be feasible to implement the federal optimum. This may be the case even with perfect information; alternatively, it may be the consequence of asymmetric information, requiring financial transfers that may be prohibitively high.

To keep things simple, let us ignore asymmetric information but assume that the transaction cost to implement the federal optimum is prohibitively high. Is it then better not to allow the city to invest in externality-reduction at all, or is it better simply not to intervene? To find out, we compare federal welfare when the city implements its optimal policy and federal welfare when the instrument is not used at all. This welfare difference is defined as

$$W^f(z = z^u) - W^f(z = 0) =$$

$$\int_0^{Y^u} P^Y(y)dy + \int_0^{X^u} P^X(x)dx - \left[c + \frac{1}{2}d(z^u)^2 + \tau_C\right](V^u) + \tau_C(V^u) - \Psi\left\{e\left(1 - \alpha(z^u)\right)(V^u)\right\} - \frac{b}{2}(z^u)^2$$

$$- \left\{\int_0^{Y^0} P^Y(y)dy + \int_0^{X^0} P^X(x)dx - (c + \tau_C)(V^0) + \tau_C(V^0) - \Psi\left\{e(V^0)\right\}\right\}$$

Here W^f captures federal welfare. The superscripts '*u*' and '*0*' refer to values at the urban government's optimal choice of z and at *0*, respectively. Reformulating previous expression yields:

$$W^f(z = z^u) - W^f(z = 0) = \left(NCS^u - NCS^0\right) + \tau_C\left(V^u - V^0\right) - \frac{1}{2}b(z^u)^2 \tag{32}$$

$$+ \left[eV^0 - e\left(1 - \alpha z^u\right)V^u\right]$$

As before, $V^i = Y^i + X^i \quad i = u, 0$, and

$$NCS^u = \left[\int_0^{Y^u} P^Y(y)dy + \int_0^{X^u} P^X(x)dx - \left(c + \frac{1}{2}d(z^u)^2 + \tau_C\right)V^u\right]$$

$$NCS^0 = \left[\int_0^{Y^0} P^Y(y)dy + \int_0^{X^0} P^X(x)dx - \left(c + \tau_C\right)V^0\right]$$

represents the net (of generalized cost) consumer surplus in the two cases.

Assuming linear demands, the change in net surplus is just the change in the generalized price times the average demand between $z = z^u$ and $z=0$:

$$NCS^u - NCS^0 = -\frac{1}{2}d(z^u)^2\left(\frac{V^u + V^0}{2}\right)$$

It follows that (32) reduces to

$$W^f(z = z^u) - W^f(z = 0) = -\frac{1}{2}(z^u)^2\left[b + d\left(\frac{V^u + V^0}{2}\right)\right] + (\tau_C - e)(V^u - V^0) + e\alpha V^u z^u \tag{33}$$

The first and the third term on the right hand side of (33) are negative and positive, respectively; the second term depends on the price elasticity of demand and on the toll level in place. It will be positive and large if the toll is much smaller than the unit externality e and demand is quite price sensitive. It then follows that it is better for the federal government simply to restrict z to zero and not let the urban government implement its optimal policy, unless the unit externality is large

and under-taxed, demand is very price sensitive, and using the externality-reducing investment z is very effective at reducing the external cost (large α).

As a special case, suppose demand is totally inelastic. This implies $V^u = V^0 = V$; moreover, using the urban government's optimal investment in externality-reduction

$$z^u = \frac{\alpha eV}{b + dY}$$

we can then show that (33) reduces to

$$W^f(z = z^u) - W^f(z = 0) = \frac{1}{2}[b + d(Y - X)](z^u)^2$$

This expression illustrates the important role of through traffic. It says that it is better to let the city have its way as long as *X* is not too large (in fact, it should be sufficiently larger than local demand *Y*). If transit is much more important than local traffic it is better to forbid the use of the instrument. Note that one can show that this result continues to hold, conditional on a given price sensitivity.

Proposition 8. If the implementation cost to the government prevents implementation of the federal optimum, then it is better to leave the city government unregulated if there is not too much transit, the unit externality is large and under-taxed, and using the externality-reducing investment *z* is very effective at reducing the external cost. If the opposite conditions hold, it is better not to allow the city to use the instrument at all, and restrict *z* to 0.

7. Summary and conclusions

In this paper we studied a series of policy instruments that governments can use to reduce pollution-type externalities in inner cities, and we specifically focused on the distortions that local government decisions may give as compared to federal solutions. The measures considered included tolls, emission standards, low emission zones, bypass capacity to guide traffic around the city center, and a variety of specific externality-reducing investments (traffic lights,

increasing the red phase of traffic lights, road bumps, imposing speed limits, building noise walls, etc.).

We obtained some remarkable results. For example, we found that the city government will over-invest in externality-reducing infrastructure whenever this infrastructure increases the generalized cost of transit traffic. This implies, for example, that it will over-invest in general in 'capacity reduction' (through road bumps, taking away road capacity for biking paths or walkways, red light phasing, etc.), but not in noise reduction via quiet asphalt, etc. We also showed that, when implementing low emission zones, the urban government will set both the fee for non-compliance and the standard at a higher level than the federal government; moreover, at sufficiently high transit levels the urban government will prefer imposing a toll instead of implementing a low emission zone, even when the latter would be the optimal federal policy to pursue. We further find that the city will underinvest in bypass capacity compared to the social optimum. More importantly, it will only invest in bypass capacity when it is allowed to earn toll revenues on the bypass that exceed the investment cost. Finally, we considered the need for federal intervention, taking into account asymmetric information and implementation costs. If there is uncertainty about the level of the locally generated externality and through traffic is important, implementing the optimal federal policy requires financial incentives to induce honest reporting of externalities by the city; no such transfers are needed if through traffic is unimportant. Furthermore, it may in fact be better to leave the city government unregulated than not to allow the city to use externality-reducing instrument at all. This will be the case if there is not too much transit, the unit externality is large and under-taxed, and the externality-reducing investment is very effective at reducing the external cost.

Of course, the model used in this paper was quite simple. Extensions might include a more detailed analysis of the role of congestion in the discussion about bypass capacity, a more realistic description of low emission zones -- allowing for individual cost differences in compliance costs --, an analysis of optimal division of authority over policy instruments, and considering more sophisticated strategic behavior by governments.

References

Annema, J. and S.-L. Proag, 2011. Ex post evaluation of the Dutch car scrappage scheme and heavy good vehicles Low Emission Zones

Arnott, de Palma and Lindsey, 1993. A structural model of peak period congestion: a traffic bottleneck with elastic demand, American Economic Review 83, 161-179.

Bento A., Goulder L., Jacobsen M. and R. von Haefen, 2009. Distributional and Efficiency Impacts of Increased US Gasoline Taxes, American Economic Review, 99 (3), 1-37.

Brueckner, J.K., 2004. Network structure and airline scheduling, Journal of Industrial Economics LII (2), 291-312.

De Borger, B. and S. Proost (eds), 2001. Reforming transport pricing in the European Union, Edward Elgar.

De Borger, B. and S. Proost, 2011, A political economy model of urban road pricing, Journal of Urban Economics,71(1),79-92

De Borger, B., Dunkerley, F. and S. Proost, 2007. Strategic investment and pricing decisions

in a congested transport corridor, Journal of Urban Economics 62, 294–316.

De Palma, A. and R. Lindsey, 2007, Transport user charges and cost recovery, in De Palma, Lindsey and Proost (Eds.), Investment and the use of tax and toll revenues in the transport sector, Research in Transportation Economics Volume 19, Elsevier, Amsterdam.

Fischer C, Harrington W., Parry I.W.H., 2007. Should Corporate Average Fuel Economy (CAFE) Standards be Tightened?, Energy Journal, 28, 1-29.

Fullerton D., West S., 2002. Can Taxes on Cars and on Gasoline Mimic an Unavailable Tax on emissions?, Journal of Environmental Economics and Management, 43, 135- 157.

Glaeser E.L. (2008), Cities, Agglomeration and spatial equilibrium, Oxford University Press, USA, 288 p

Mayeres, I., Proost, S., Miltz, D. (1993). The Geneva hydrocarbon protocol: Economic insights from a Belgian perspective. Environmental & Resource Economics, 3, 107-127.

Mohrung, H. and M. Harwitz, 1962. Highway benefits: an analytical framework, Northwersten University Press, Evanston.

Parry, I. and A. Bento, 2001. Revenue recycling and the welfare effects of road pricing, Scandinavian Journal of Economics 103(4), 645-671

Parry, I and K. Small, 2005. Does Britain of the United States have the right gasoline tax, American Economic Review 95, 1276-1289.

Parry, I.W.H., Walls M., Harrington W., 2007. Automobile Externalities and Policies, Journal of Economic Literature XLV, 374- 400.

Proost, S., 2011, Theory of external costs, chapter 14 in A Handbook in Transport Economics, de Palma A., R. Lindsey, E. Quinet and R. Vickerman (eds). Edward Elgar

Spence, M.A., 1975. Monopoly, quality and regulation, Bell Journal of Economics and Management Science 6 (2), 417-429.

Van Der Loo S, Proost S. (2011), The European road pricing game: how to enforce optimal pricing in high-transit countries under asymmetric information, Discussion Paper CES-KULeuven http://www.econ.kuleuven.be/eng/ew/discussionpapers/Dps11/DPS1119.pdf

Verhoef, E., Nijkamp, P. and P. Rietveld, 1996. Second-best congestion pricing: The case of an untolled alternative, Journal of Urban Economics 40, 279–302.

Ubbels, B. and E. Verhoef, 2008. Governmental competition in road charging and capacity choice, Regional Science and Urban Economics.

Van Dender, K., 2003. Duopoly prices under congested access, Journal of Regional Science 45, 343–362.

Wolff, H. and L. Perry, 2010. Trends in Clean Air Legislation in Europe: Particulate Matter and Low Emission Zones, Review of Environmental Economics and Policy 4 (2), 293–308.

Appendix 1: Details on the city's optimal choice of z

Using equality between generalized price and generalized cost in equilibrium (i.e., $P^Y(Y) = G(z, \tau_C) = c + \frac{1}{2} dz^2 + \tau_C$) and noting the constraint $z \leq \frac{1}{\alpha}$, we can write the first-order conditions for $z \geq 0$ associated with problem (4) in general as follows:

$$\left\{ \left[\Psi' \alpha e(Y+X) + (\tau_C - \Psi' e(1-\alpha z)) \left(\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P} \right) dz - (b + dY)z \right] - \lambda \right\} \leq 0$$

$$z \left\{ \left[\Psi' \alpha e(Y+X) + (\tau_C - \Psi' e(1-\alpha z)) \left(\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P} \right) dz - (b + dY)z \right] - \lambda \right\} = 0 \qquad \text{(A1.1)}$$

$$z \leq \frac{1}{\alpha}; \quad \lambda \left(z - \frac{1}{\alpha} \right) = 0$$

In these expressions, $\lambda \geq 0$ is the multiplier associated with the inequality constraint.

Assuming an internal solution exists (hence $\lambda = 0, 0 < z < \frac{1}{\alpha}$), the first-order conditions (A1.1) boil down to

$$\Psi' \alpha e(Y+X) + \{\tau_C - \Psi' e(1-\alpha z)\} \left[\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P} \right] dz = bz + dzY$$

In principle, of course, corner solutions where either the measure is not used at all ($z=0$) or where it is used up to its full potential ($z = \frac{1}{\alpha}$) are possible. However, inspection of system (A1.1) implies that $z=0$ will never be optimal as long as $\alpha > 0$. To see this, if $z=0$ the three expressions in (A1.1) would jointly imply $\Psi' \alpha e(Y+X) < 0$. As the externality cost is increasing in the externality level ($\Psi' > 0$), this gives a contradiction, hence $z>0$. System (A1.1) further also implies that the maximum possible investment will be optimal if, evaluated at $z = \frac{1}{\alpha}$, the marginal benefit of investment still exceeds the marginal cost. Indeed, $z = \frac{1}{\alpha}$ implies $\lambda > 0$, so that the second equation of (A1.1) gives:

$$\Psi' \alpha e(Y+X) + \tau_C \left(\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P} \right) dz > (b + dY)z$$

Appendix 2: Low emission zones with identical compliance costs

If all users have the same compliance costs, depending on the cost of complying with the standard relative to the fee, either all users will comply or no one will comply. Hence, the government's problem boils down to choosing between an optimal standard (and a prohibitively high fee, so that all users comply) and an optimal fee (plus a prohibitively strict standard so that no one complies).

Consider first the federal government. It sets the standard so as to

$$
\underset{z_s^f}{Max} \quad W_s^f = \left\{ \int_0^{Y_s^f} p^Y(y)dy - \left[c + \frac{1}{2}d\left(z_s^f\right)^2 \right] Y_s^f \right\}
$$
$$
+ \left\{ \int_0^{X_s^f} p^X(x)dx - \left[c + \frac{1}{2}d\left(z_s^f\right)^2 \right] X_s^f \right\} - e\left(1 - \alpha z_s^f\right)\left(Y_s^f + X_s^f\right)
$$

The subscript 's' denotes a standard, the superscript 'f' refers to the values at the optimum from the federal perspective. After rearrangement, the first-order condition can be written as follows:

$$
(\frac{\alpha e}{d} - z_s^f)(Y_s^f + X_s^f) = e(1 - \alpha z_s^f)\left[(\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P})z_s^f \right]
$$

Selecting the optimal fee is the same as setting an optimal toll. Denoting the fee as F we have that the optimal fee solves

$$
\underset{F^f}{Max} \quad W_F^f = \left\{ \int_0^{Y_F^f} p^Y(y)dy - \left[c + \tau^f \right] Y_F^f \right\}
$$
$$
+ \left\{ \int_0^{X_F^f} p^X(x)dx - \left[c + \tau^f \right] X_F^f \right\} + \tau^f \left(Y_F^f + X_F^f \right) - e\left(Y_F^f + X_F^f \right)
$$

The notation Y_F^f, X_F^f refers to the values of Y and X when the federal government uses an optimal fee. Given the equivalence of a toll and a fee in this case we obviously find that the federal fee equals marginal external cost $F^f = e$.

To study the welfare difference between the use of an optimal standard and an optimal toll, we denote federal welfare with the optimal standard and optimal fee as W_s^f, W_F^f respectively. We then have

$$
W_s^f - W_F^f = NCS_{Y,s}^f - NCS_{Y,F}^f + NCS_{X,s}^f - NCS_{X,F}^f - e(1 - \alpha z_s^f)(X_s^f + Y_s^f)
$$

Here the different terms reflecting net consumer surplus (NCS) are defined as

$$NCS^f_{Y,s} = \left\{ \int_0^{Y^f_s} p^Y(y)dy - \left[c + \frac{1}{2}d\left(z^f_s\right)^2 \right]Y^f_s \right\}; \quad NCS^f_{X,s} = \left\{ \int_0^{X^f_s} p^X(x)dx - \left[c + \frac{1}{2}d\left(z^f_s\right)^2 \right]X^f_s \right\}$$

$$NCS^f_{Y,F} = \left\{ \int_0^{Y^f_F} p^Y(y)dy - \left[c + \tau^f \right]Y^f_F \right\}; \quad NCS^f_{X,F} = \left\{ \int_0^{X^f_F} p^X(x)dx - \left[c + \tau^f \right]X^f_F \right\}$$

Assuming linear demands we can, using straightforward algebra, rewrite the welfare difference as

$$W^f_s - W^f_F = \frac{1}{2}\alpha e z^f_s \left(Y^f_s + X^f_s\right)$$

$$+ \frac{1}{2}\left(\frac{\partial Y}{\partial P} + \frac{\partial X}{\partial P}\right)\left\{\left[e - \frac{1}{2}d\left(z^f_s\right)^2\right]^2 + e(1-\alpha z^f_s)d\left(z^f_s\right)^2\right\}$$

We proceed in a similar way for the urban government. It would set the standard according to

$$\underset{z^u_s}{Max} \quad W^u = \left\{ \int_0^{Y^u_s} p^Y(y)dy - \left[c + \frac{1}{2}d\left(z^u_s\right)^2 \right]Y^u_s \right\} - e\left(1-\alpha z^u_s\right)\left(Y^u_s + X^u_s\right)$$

The subscript 's' refers, as before, to the standard, the superscript 'u' of course refers to the values at the optimum from the urban perspective. Rearranging the first-order condition, one obtains:

$$\left(\frac{\alpha e}{d} - z^f_s\right)(Y^u_s) = e(1-\alpha z^u_s)\left[\left(\frac{\partial Y}{\partial P}\right)z^u_s\right] - \frac{\alpha e}{d}X^u_s + e(1-\alpha z^u_s)\left[\left(\frac{\partial X}{\partial P}\right)z^u_s\right]$$

If a fee were the only instrument it easily follows that the optimal fee would solve

$$\underset{F^u}{Max} \quad W^u = \left\{ \int_0^{Y^u_F} p^Y(y)dy - \left[c + F^u \right]Y^u_F \right\} + F^u\left(Y^u_F + X^u_F\right) - e\left(Y^u_F + X^u_F\right)$$

The notation Y^u_F, X^u_F refers to the values of Y and X when the urban government uses an optimal fee. We find the optimal fee

$$F^u = e - \frac{X^u_F}{\dfrac{\partial X}{\partial F^u} + \dfrac{\partial Y}{\partial F^u}}$$

Denoting urban welfare with the optimal standard and optimal fee as W_s^u, W_F^u respectively, we have

$$W_s^u - W_F^u = NCS_s^u - NCS_F^u - e(1-\alpha z_s^u)(X_s^u + Y_s^u) - (F^u - e)(X_F^u + Y_F^u)$$

Here *NCS* stands for net consumer surplus, defined as

$$NCS_s^u = \left\{ \int_0^{Y_s^u} p^Y(y)dy - \left[c + \frac{1}{2}d\left(z_s^u\right)^2 \right] Y_s^u \right\}$$

$$NCS_F^u = \left\{ \int_0^{Y_F^u} p^Y(y)dy - \left[c + F^u \right] Y_F^u \right\}$$

Assuming linear demands, one easily shows that

$$NCS_s^u - NCS_F^u = \left[F^u - \frac{1}{2}d\left(z_s^u\right)^2 \right]\left[Y_s^u + \frac{Y_F^u - Y_s^u}{2} \right]$$

Again using the assumption of linear demands, we find after long but simple algebra

$$W_s^u - W_F^u = \left[\alpha e z_s^u - d\left(z_s^u\right)^2 \right] Y_s^u + \frac{1}{2}\frac{\partial Y}{\partial P}\left[F^u - \frac{1}{2}d\left(z_s^u\right)^2 \right]^2 + eY_F^u$$
$$- e(1-\alpha z_s^u)X_s^u - (F^u - e)(X_F^u - Y_s^u)$$

Appendix 3. Details on low emission zones

Consider an internal solution for the federal government in which locals obey the standard whereas transit users prefer to pay the fine. This internal optimum solves

$$\underset{z_s^f, F^f}{Max} \quad W^u = \left\{ \int_0^{Y^f} p^Y(y)dy - \left[c + \frac{1}{2}d_Y\left(z_s^f\right)^2 \right] Y^f \right\} + \left\{ \int_0^{X^f} p^X(x)dx - \left[c + F^f \right] X^f \right\}$$
$$+ F^f X^f - e\left(1-\alpha z_s^f\right)\left(Y^f\right) - eX^f$$

There are fee revenues on transit only. Moreover, transit produces more external cost per kilometer than local demand because it does not comply with the standard. It is straightforward to show, using equality between generalized price and generalized cost, that the optimal values satisfy

$$\alpha e Y^f - e(1-\alpha z_s^f)\left[\frac{\partial Y}{\partial P} \right]d_Y z_s^f - d_Y z_s^f Y_s^f = 0$$

$$F^f = e$$

Similarly, for an internal optimum (in which locals obey the standard whereas transit users prefer to pay the fine) the urban government solves:

$$\underset{z_s^u, F^u}{Max} \quad W^u = \left\{ \int_0^{Y^u} p^Y(y) dy - \left[c + \frac{1}{2} d_Y \left(z_s^u \right)^2 \right] Y^u \right\} + F^u X^u - e \left(1 - \alpha z_s^u \right) \left(Y^u \right) - e X^u$$

The first order conditions are

$$\alpha e Y^u - e(1 - \alpha z_s^u) \left[\frac{\partial Y}{\partial P} \right] d_Y z_s^u - d_Y z_s^u Y^u = 0$$

$$F^u = e - \frac{X^u}{\frac{\partial X}{\partial F}}$$

Appendix 4. Proof of expression (27)

First, note that by the optimality of $\tilde{\tau}_C, \tilde{z}$ it must be the case that

$$W^u(\tilde{z}, \tilde{\tau}_C, 0, 0) > W^u(z^*, \tau_C^*, 0, 0) \tag{A4.1}$$

Moreover, it is easy to show that

$$W^u(z, \tau_C, \tau_B, k; \tau_B X - \frac{f}{2} k^2 = 0) = W^u(z^*, \tau_C^*, 0, 0) \tag{A4.2}$$

where $W^u(z, \tau_C, \tau_B, k; \tau_B X - \frac{f}{2} k^2 = 0)$ is optimal city welfare under a break even constraint for the bypass: toll revenues equal bypass investment costs. To see why (A4.2) holds, straightforward algebra shows that the solution to

$$\underset{z, \tau_C, k, \tau_B}{Max} \quad \int_0^Y P^Y(y) dy - \left[c + \frac{1}{2} dz^2 + \tau_C \right] Y + \tau_C Y + \tau_B X - e(1 - \alpha z) Y - \frac{b}{2} z^2 - \frac{f}{2} k^2$$

$$s.t. \quad \tau_B X - \frac{f}{2} k^2 = 0$$

yields z^*, τ_C^* as optimal values for externality reducing investment and for the city toll, respectively. Moreover, using $\tau_B X - \frac{f}{2} k^2 = 0$ and noting that all through traffic uses the bypass, optimal welfare for this case satisfies

$$W^u(z, \tau_C, \tau_B, k; \tau_B X - \frac{f}{2}k^2 = 0)$$

$$= \int_0^{Y*} P^Y(y)dy - \left[c + \frac{1}{2}d(z^*)^2 + \tau_C^* \right] Y^* + \tau_C^* Y^* - e(1 - \alpha z^*)Y^* - \frac{b}{2}(z^*)^2$$

$$= W^u(z^*, \tau_C^*, 0, 0)$$

Comparing (A4.1) and (A4.2) it then follows

$$W^u(\tilde{z}, \tilde{\tau}_C, 0, 0) > W^u(z, \tau_C, \tau_B, k; \tau_B X - \frac{f}{2}k^2 = 0) \tag{A4.3}$$

This shows that not investing in bypass capacity is always better than investing in the bypass when the bypass has to break even.

Appendix 5. A general model with congestion

In Section 5 of the paper, we studied the urban government's incentives for investing in bypass capacity to guide traffic around the city, and we compared the behavior of the urban government with that of the federal government. To keep the analysis transparent and to focus on pollution-type externalities, the model we used made abstraction of congestion on both the urban road and the bypass. In this appendix, we extend the model to explicitly take account of potential congestion on both roads. If there is no congestion, either all through traffic uses the urban road or all through traffic uses the bypass, depending on which road has the lowest generalized cost. However, in the presence of congestion, through traffic will be distributed over the two alternative routes. The question then is to what extent this affects the results summarized in Proposition 5.

As before, demand Y by locals for using the urban road is described by the inverse demand function $P^Y(Y)$. Similarly, demand for through traffic X is described by inverse demand $P^X(X)$. However, this through traffic may pass through the city (using the urban road) or it may use the bypass; the respective traffic volumes are given by X_c, X_b, where $X = X_c + X_b$. We assume that through traffic uses the road with the lowest generalized cost so that, if both routes are used by through traffic, Wardrop equilibrium implies equality of generalized prices.

The generalized cost of using the urban road is specified as follows:

$$c + \frac{1}{2}dz^2 + \left[n(Y + X_c) \right] + \tau_c$$

As before, it depends on externality-reducing investment z and it captures the potential toll on the urban road. Moreover, it increases in the total traffic flow using the road: more traffic raises the cost, reflecting speed reductions due to congestion. The impact of traffic on the money plus time cost is specified linearly for simplicity; the parameter n determines how strongly an increase in the traffic flow raises the generalized cost.

Similarly, the generalized cost of using the bypass is specified as:

$$l + g\left(\frac{X_b}{k}\right) + \tau_b.$$

The parameter l is the money plus time cost at free flowing traffic if there were no user toll. The generalized cost is assumed to rise linearly in the volume-capacity ratio (X_b / k). Finally, the toll on the bypass is denoted τ_b.

There are three types of traffic and two roads. Equality of generalized prices and generalized costs implies the following equilibrium conditions:

$$P^Y(Y) = c + \frac{1}{2}dz^2 + \left[n(Y + X_c)\right] + \tau_c$$

$$P^X(X_c + X_b) = c + \frac{1}{2}dz^2 + \left[n(Y + X_c)\right] + \tau_c \qquad (A5.1)$$

$$P^X(X_c + X_b) = l + g\frac{X_b}{k} + \tau_b$$

Following the methodology described in detail in De Borger, Proost and Van Dender (2005) and De Borger, Dunkerley and Proost (2007), the system of three equations (A5.1) generates 'reduced-form' demand functions for the three types of traffic (Y, X_c, X_b) as functions of all policy variables considered, viz. tolls on both roads, externality-reducing investment and bypass capacity: τ_c, τ_b, z, k. To arrive at these reduced-form demand functions, the procedure is to totally differentiate (A5.1) and to solve the resulting system by Cramer's rule[15]. It is then easy to show that:

$$\frac{\partial Y}{\partial \tau_c} < 0; \quad \frac{\partial Y}{\partial \tau_b} < 0; \quad \frac{\partial Y}{\partial z} < 0; \quad \frac{\partial Y}{\partial k} > 0$$

$$\frac{\partial X_c}{\partial \tau_c} < 0; \quad \frac{\partial X_c}{\partial \tau_b} > 0; \quad \frac{\partial X_c}{\partial z} < 0; \quad \frac{\partial X_c}{\partial k} < 0$$

---

[15] The derivations are straightforward but quite long. They are available from the authors.

$$\frac{\partial X_b}{\partial \tau_c} > 0; \quad \frac{\partial X_b}{\partial \tau_b} < 0; \quad \frac{\partial X_b}{\partial z} > 0; \quad \frac{\partial X_b}{\partial k} > 0$$

These effects are all quite plausible. First, both a higher toll on the urban road and on the bypass (because this attracts more transit through the city) reduce local demand. Investing in externality reduction reduces local demand; more bypass capacity raises it, as it diverts some through traffic from the urban road to the bypass. Second, a higher toll on the urban road reduces through traffic via the city, a bypass toll raises it. Both investing in externality reduction in the city and investing in bypass capacity reduce through traffic via the urban road. Finally, for bypass traffic the signs are also as expected.

The objective function for the urban government is adapted to capture urban road congestion; moreover, it allows for tolls on both the urban and the bypass road, and captures the possibility to invest in z and k. It can be written as

$$\int_0^Y P^Y(y)dy - \left\{ c + \frac{1}{2}dz^2 + [n(Y + X_c)] + \tau_c \right\} Y$$

$$+ \tau_c(Y + X_c) + \tau_b X_b - e(1 - \alpha z)(Y + X_c) - \frac{b}{2}z^2 - \frac{f}{2}k^2 \qquad (A5.2)$$

Similarly, the objective function of the federal government reads

$$\int_0^Y P^Y(y)dy + \int_0^X P^X(x)dx - \left\{ c + \frac{1}{2}dz^2 + [n(Y + X_c)] + \tau_c \right\}(Y + X_c) - \left[l + g\frac{X_b}{k} + \tau_b \right] X_b$$

$$+ \tau_c(Y + X_c) + \tau_b X_b - e(1 - \alpha z)(Y + X_c) - e_B X_B - \frac{b}{2}z^2 - \frac{f}{2}k^2$$

$$(A5.3)$$

In this expression, the non-congestion external cost per kilometer on the bypass is denoted as e_B.

With the above setup, we reconsider the analysis of investing in bypass capacity reported in Section 5 of the paper. We start by looking at bypass capacity k as the only policy variable and try to find out to what extent Propositions 5a and 5b have to be adapted when allowing for congestion. Next we deal with the case where governments can use all instruments to see whether Propositions 5c and 5d still hold in this more general setting.

Let bypass capacity k be the only policy variable. Maximize (A5.2) with respect to k, use the equality between generalized price and generalized cost, and rearrange to obtain

$$[\tau_c - nY - e(1 - \alpha z)]\frac{\partial(Y + X_c)}{\partial k} + \tau_b \frac{\partial X_b}{\partial k} = fk \qquad (A5.4)$$

Noting from before that $\dfrac{\partial(Y+X_c)}{\partial k}<0$, the urban government's optimal bypass investment rises in the external cost e and decreases in capacity cost f. Moreover, it decreases in the toll on the urban road, and it rises in the toll on the bypass. Similarly, maximizing (A5.3) yields, for the federal government, the first-order condition

$$\left[\tau_c - n(Y+X_c) - e(1-\alpha z)\right]\frac{\partial(Y+X_c)}{\partial k} + (\tau_b - g\frac{X_b}{k} - e_b)\frac{\partial X_b}{\partial k} = fk - g\left(\frac{X_b}{k}\right)^2 \quad (A5.5)$$

Optimal investment in bypass capacity depends on the exogenous toll levels. If the tolls reflect marginal external cost, both on the urban road and the bypass, we have

$$\left[\tau_c - n(Y+X_c) - e(1-\alpha z)\right] = 0; \quad (\tau_b - g\frac{X_b}{k} - e_b) = 0$$

In that case it follows from (A5.5) that optimal investment in bypass capacity is governed by the rule

$$fk = g\left(\frac{X_b}{k}\right)^2$$

This rule implies equality between the marginal cost of extra capacity (left-hand-side) and the direct marginal benefit of more capacity (right-hand-side); the latter reflects the savings in time costs on the bypass for all users. If the exogenous tolls are not set optimally, (A5.5) further implies that federal investment in bypass capacity is decreasing in the urban toll and rising in the toll on the bypass.

Straightforward algebra allows us to rewrite (A5.4)-(A5.5) as

$$k^u = \frac{\left[\tau_c - nY - e(1-\alpha z)\right]\dfrac{\partial(Y+X_c)}{\partial k} + \tau_b\dfrac{\partial X_b}{\partial k}}{f}$$

$$k^f = \frac{\left[\tau_c - nY - e(1-\alpha z)\right]\dfrac{\partial(Y+X_c)}{\partial k} + \tau_b\dfrac{\partial X_b}{\partial k} - nX_c\dfrac{\partial(Y+X_c)}{\partial k} - e_b\dfrac{\partial X_b}{\partial k} + g\left(\dfrac{X_b}{k}\right)^2\left(1-\varepsilon_{b,k}\right)}{f}$$

In this last expression

$$\varepsilon_{b,k} = \frac{\partial X_b}{\partial k}\frac{k}{X_b}$$

44

is the elasticity of bypass demand with respect to capacity. Plausibly, this elasticity is smaller than one[16].

Of course, the expressions just derived are no closed-form solutions. However, they are helpful is finding out how Propositions 5a and 5b have to be adapted for the case of congestion. First, assuming $\varepsilon_{b,k} < 1$ and noting that $\frac{\partial (Y + X_c)}{\partial k} < 0, \frac{\partial X_b}{\partial k} > 0$, the expressions imply that the urban government will invest less in bypass capacity than the federal government, unless the external cost of pollution on the bypass itself is very high (e_b large). If pollution due to extra traffic on the bypass is severe then the federal government may in fact invest less in bypass capacity than the urban authority. Second, the expression for k^u suggests that optimal bypass investment by the urban government increases in the externality e and declines in the investment cost f.

Second, let us reconsider the incentives for the urban government to invest in bypass capacity for the case where all instruments are available, and see to what extent Propositions 5c and 5d continue to go through. Differentiating (A5.2) with respect to the two tolls τ_c, τ_b and with respect to investments in externality reduction z and in bypass capacity k yields, after standard algebra, the following set of first-order conditions[17]:

$$\left[\tau_c - nY - e(1-\alpha z) \right] \frac{\partial (Y + X_c)}{\partial \tau_C} + \tau_b \frac{\partial X_b}{\partial \tau_C} + X_C = 0 \tag{A5.6}$$

$$\left[\tau_c - nY - e(1-\alpha z) \right] \frac{\partial (Y + X_c)}{\partial \tau_b} + \tau_b \frac{\partial X_b}{\partial \tau_b} + X_b = 0 \tag{A5.7}$$

$$\left[\tau_c - nY - e(1-\alpha z) \right] \frac{\partial (Y + X_c)}{\partial z} + \tau_b \frac{\partial X_b}{\partial z} - (b + dY)z + \alpha e(Y + X_C) = 0 \tag{A5.8}$$

$$\left[\tau_c - nY - e(1-\alpha z) \right] \frac{\partial (Y + X_c)}{\partial k} + \tau_b \frac{\partial X_b}{\partial k} = fk \tag{A5.9}$$

---

[16] It is easy to show that it equals minus the generalized price elasticity of demand for bypass use times a fraction which by definition is smaller than one, viz. the ratio of the congestion component of the generalized price (gX_b / k) and the full generalized price $(l + (gX_b / k) + \tau_b)$.

[17] Derivations are again straightforward; they are available from the authors.

To see the implications of these conditions, consider expressions (A5.6) and (A5.9). First, if the city is not allowed to set a toll on the bypass (hence let $\tau_b = 0$) but does have the authority to decide on its optimal toll on the urban road, condition (A5.6) implies that the optimal urban toll will exceed the local marginal external cost: we have $[\tau_c - nY - e(1-\alpha z)] > 0$. However, in that case, (A5.9) implies that the optimal investment in bypass capacity is negative. In practice, we have a corner solution $k=0$ where no extra bypass capacity is offered. This means that Proposition 5c continues to hold: if the urban government can charge the optimal urban toll but cannot toll the bypass, then it is not interested in investing in bypass capacity at all. Second, however, the presence of congestion and the distribution of through traffic over the urban and the bypass road implies Proposition 5d. does not hold in exactly the same strong form. It does follow from (A5.9) that optimal bypass capacity is an increasing function of the bypass toll: the urban government will be much more interested in investing if it can charge high tolls to collect revenue.