Manipulable Congestion Tolls

by

Jan K. Brueckner
Department of Economics
University of California, Irvine
3151 Social Science Plaza
Irvine, CA 92697
e-mail: jkbrueck@uci.edu

and

Erik T. Verhoef
Department of Spatial Economics
VU University
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
e-mail: everhoef@feweb.vu.nl

December 2008

Abstract

The recent literature on congestion pricing contains a remarkable inconsistency: agents are large enough to recognize self-imposed congestion and exert market-power over prices, but they do not take account of the impact of their own actions on the magnitude of congestion tolls. When large agents are confronted with tolls derived under this parametric assumption but understand the rule used to generate them, the toll system will no longer guide the market to the social optimum. To address this problem, this paper derives an alternative, manipulable toll rule, which is designed to achieve the social optimum when agents do take into account the full impacts of their own actions on total toll liabilities.

Manipulable Congestion Tolls

by

Jan K. Brueckner and Erik T. Verhoef*

Pigouvian taxes are standard instruments for dealing with negative externalities. The problem, of course, is that the activity generating the negative external effect occurs at too high a level. The Pigouvian tax is designed to raise the price for the activity, reducing its overall level to one that is socially optimal. To compute the tax, the marginal externality damage from the activity is evaluated at the social optimum, with the tax on each unit of the activity then set at the resulting value. Faced with the tax, the offending agents restrict their individual activity levels, leading to an optimal overall level.

In a typical application, the marginal externality damage is itself a function of the overall activity level, usually an increasing one. However, the computed tax is just a scalar value, generated by evaluating the marginal damage function at the socially optimal activity level. Thus, while the Pigouvian tax rule involves the level of the activity, the per-unit tax ultimately charged is just a number generated by evaluating the rule at the social optimum.

The Pigouvian approach assumes that agents treat the resulting tax as parametric, independent of their chosen activity levels. This view is appropriate when the externality is jointly generated by many agents, each of whom makes a small contribution to the overall level of the offending activity. However, when the externality is generated by just a few large agents, a parametric view of the Pigouvian tax may be less plausible.

To understand this point, imagine that the socially optimal activity level varies from period to period as a result of changes in the economy's parameters (shifts in demand or cost curves, for example). Then, if the marginal damage function is increasing, the Pigouvian tax will vary as well, exhibiting positive correlation with both the overall activity level and individual activity levels within the small set of agents. This correlation may reveal to the agents the nature of the rule used to compute the Pigouvian tax, showing that the tax per unit depends on the overall activity level and thus, in the case where agents are large, on their individual activity levels. Another reason why agents could be aware of the rule underlying the calculation of Pigouvian taxes is that the principles according to which the tax level is determined may have to be made transparent and publicly accessible for legal reasons. Agents can then observe the rule directly. In either case, with this knowledge, agents may then attempt to manipulate the tax regime, further restricting their activity levels in order to depress the magnitude of the per-unit tax that they pay. Such behavior, however, undermines the Pigouvian tax regime, making it incapable of achieving the optimum.

The purpose of the present paper is to explore an alternative approach to corrective taxation that recognizes the potential existence of such manipulative behavior. Our approach confronts manipulation head on by replacing the conventional Pigouvian tax, designed to be treated parametrically by individuals, with a tax rule that is *designed to be manipulated*. In other words, the planner announces a rule that gives each agent's tax liability as a function of his or her own activity level and the levels of other agents. Each agent then optimizes with full knowledge of the rule used to compute the tax liability, with the optimization being carried out conditional on the choices of other agents. Thus, instead of using a Pigouvian rule that operates behind the scenes to generate a tax value meant to be viewed as parametric, our approach presents a transparent, manipulable rule that is directly exploited by the agents in choosing their activity levels. The goal of the analysis is to derive the form of such manipulable tax rules, providing a comparison to the Pigouvian case.

For concreteness, the analysis focuses on congestion externalities, where Pigouvian taxes take the form of congestion tolls. Traditional analyses of road congestion pricing (see Small and Verhoef (2007)) are mostly immune to the manipulation critique from above, a consequence of the fact that road users are typically atomistic and thus unable to manipulate Pigouvian congestion tolls. However, a recent literature focuses on a case where the agents generating congestion are non-atomistic, raising concerns about manipulation. This is the case of airport congestion, which arises from usage of a capacity-constrained airport by a relatively small number of airlines, some of which may account for an appreciable share of the total flights. The paper's analysis is developed in the airport context.

Departing from previous studies, a recent literature on airport congestion recognizes the non-atomistic nature of airlines while exploring a particular consequence of this alteration of the standard road-oriented model: internalization of congestion. Internalization occurs because a non-atomistic carrier, in scheduling an extra flight, takes into account the additional congestion costs imposed on the other flights it operates. As a result, a Pigouvian congestion toll need only charge an airline for the congestion imposed on other carriers, excluding the congestion the airline imposes on itself. One implication of this rule is that, when carriers are asymmetric, they should pay different tolls. A carrier with a large flight share should pay a low toll given that it internalizes most of the congestion from its operation of an extra flight, while a small carrier, which internalizes little of the congestion it creates, should pay a high toll. Using a simulation model, Daniel (2005) was the first to recognize the potential for internalization of congestion, while Brueckner (2002, 2005) and Pels and Verhoef (2004) explored the implications of internalization using simple analytical models. A burgeoning literature has followed these initial studies.[1]

As this discussion indicates, the recent analysis of airport congestion presumes the use of Pigouvian congestion tolls, even though the non-atomistic nature of airlines suggests the potential for manipulation of such tolls. The present paper is meant to redress this omission. By analyzing the nature of the manipulable congestion tolls in an airport context, it adds a missing component to the new theory of congestion pricing in the presence of non-atomistic agents. The results, however, apply more generally to the theory of corrective taxation, showing that manipulable taxes may need to replace Pigouvian charges in other contexts.[2]

The plan of the paper is as follows. Section 2 presents the model, which takes the simplest possible form. Two airlines serve a single travel market, with a congested airport at one endpoint. While the general model has an elastic demand for travel and cost functions that potentially differ across carriers, section 3 begins the discussion by considering a base case where demand is perfectly elastic and costs are symmetric. The analysis first derives the social optimum for this case, and then computes the Pigouvian congestion tolls required to support it, assuming that the tolls are viewed as parametric by the carriers. Manipulable tolls, required in the absence of this parametric view, are analyzed next. The analysis first derives the most general form for such tolls, and then considers their form under a plausible restriction, which requires a carrier's toll liability to equal its own flight volume times a toll per flight that is common across carriers. The issue of airport cost recovery under the different toll regimes is then considered. Section 4 focuses on the general model, in which carriers are asymmetric and demand is no longer perfectly elastic, leading to inefficient pricing mark-ups as carriers exercise market power. The analysis again considers both general and restricted toll schedules, with the latter involving a common toll per flight across the now-asymmetric carriers. Section 5 offers conclusions.

2. The Model

The analysis focuses on a single travel market with a congested airport at one endpoint. In contrast to some earlier papers, the model does not distinguish between peak and off-peak periods, so that congestion is always present. The market is served by two airlines, denoted 1 and 2, which interact in Cournot fashion. Let f_i denote the number of flights operated by carrier i, and let the number of passengers per flight be constant and normalized to unity, so that $f_1 + f_2$ represents both the total flight volume and the total number of passengers.

The demand for flights is given by the inverse demand function $D(f_1 + f_2)$, which gives the marginal willingness to pay for travel. Passenger volume is determined by equating this willingness to pay to the "full price" of travel, which includes the airfare and the value of lost passenger time due to airport congestion. With congestion depending on total flights at the airport, time cost per passenger is given by $h(f_1 + f_2)$, a function that is assumed to be increasing and convex over the relevant range of flight volumes (it may be zero at low volumes). Since the airfare plus time cost equals the full price, it follows that the fare is given by $D(f_1 + f_2) - h(f_1 + f_2)$. Thus, congestion generates a fare discount, as verified in the empirical study of Forbes (2008). Carrier i's revenue is then equal to $[D(f_1+f_2)-h(f_1+f_2)]f_i$, $i = 1, 2$.

In addition to raising passenger time costs, airport congestion increases airline operating costs, with the effect allowed to differ across carriers. Congestion cost per flight for carrier i is given by $g_i(f_1 + f_2)$, $i = 1, 2$, with these functions again assumed to be increasing and convex over the relevant range. An airline also incurs operating costs that are unrelated to airport congestion, costs that again may differ across carriers. Assuming constant returns, these costs are given by $\tau_i f_i$, where τ_i is operating cost per flight, $i = 1, 2$.

Combining the above elements, airline i's profit is given by

$$\begin{aligned}
\pi_i \;&=\; [D(f_1+f_2) - h(f_1+f_2)]f_i \;-\; \tau_i f_i \;-\; g_i(f_1+f_2)f_i \\
&=\; D(f_1+f_2)f_i \;-\; \tau_i f_i \;-\; c_i(f_1+f_2)f_i, \quad i = 1,2
\end{aligned} \tag{1}$$

where $c_i(f_1 + f_2) \equiv h(f_1 + f_2) + g_i(f_1 + f_2)$ gives passenger plus airline congestion cost. Note that the two types of congestion costs enter the profit function symmetrically because an increase in passenger time costs implies an equally large decrease in the fare that can be charged for a given output level. Firm profits are therefore equally sensitive to $h(f_1 + f_2)$ and $g_i(f_1 + f_2)$. Social welfare is measured by total profit, $\pi_1 + \pi_2$, plus consumer surplus, which is given by $\int_0^{f_1+f_2} D(x)dx - (f_1 + f_2)D(f_1 + f_2)$. Use of this welfare function requires the absence of income effects on demand, in which case consumer surplus is an exact measure.

3. Base Case: Symmetric Carriers and Perfectly Elastic Demand

It is useful to begin by considering a base case where further simplifications are imposed on the model. Accordingly, suppose that demand is perfectly elastic, so that the D function from above is equal to a constant, denoted p. In addition, let costs be symmetric across carriers, so that operating and congestion costs in (1) lose their i subscripts. After deriving results for this base case, the analysis returns to the general case in section 4.

3.1. Social optimum, laissez-faire equilibrium, and Pigouvian tolls

With perfectly elastic demand, consumer surplus is zero and social welfare equals total profit, given by $[p - \tau - c(f_1 + f_2)](f_1 + f_2)$. Maximizing this expression by choice of f_1 and f_2 yields two identical first-order conditions, given by

$$p \;-\; \tau \;-\; c(f_1+f_2) \;-\; (f_1+f_2)c'(f_1+f_2) \;\;=\;\; 0. \tag{2}$$

This condition determines a symmetric optimal flight volume, given by $f_1 = f_2 = f^*$. Note that (2) says that extra flights should be operated up to the point where the full price minus operating and congestion costs per flight equals the marginal congestion damage generated by an extra flight, given by $(f_1 + f_2)c'(f_1 + f_2) \equiv \mathrm{MCD}$. This expression equals marginal congestion cost per flight, given by the slope of the congestion cost function, times the number of flights experiencing the additional congestion.

The laissez-faire equilibrium is generated by profit maximization under Cournot behavior. Carrier i's profit equals $[p - \tau - c(f_1 + f_2)]f_i$, and the first-order condition for maximization of this expression is

$$p \ - \ \tau \ - \ c(f_1 + f_2) \ - \ f_i c'(f_1 + f_2) \ = \ 0, \quad i = 1, 2. \tag{3}$$

In (3), f_i rather than $f_1 + f_2$ multiplies c', indicating that carrier i does not take into account the congestion damage imposed on the other carrier when it schedules an extra flight. Therefore, the symmetric equilibrium flight volumes, equal to $\widehat{f}$, are too large, satisfying $\widehat{f} > f^*$. Note that, even though it ignores the impact on the other carrier, carrier i does internalize the congestion it imposes on itself, viewing its own congestion damage ($f_i c'$) as part of the cost of operating an extra flight.

A Pigouvian toll, if viewed as parametric by the carriers, can remedy this inefficiency. The toll charges each carrier for the congestion damage that it does not take into account, equaling $f_{-i}c'(f_1 + f_2)$ per flight for carrier i, where $-i$ denotes the other carrier. Evaluating this expression at the symmetric social optimum, the Pigouvian toll is then given by

$$z \ = \ f^* c'(2f^*). \tag{4}$$

When zf_i is subtracted from carrier i's profit, the new first-order conditions are

$$p \ - \ \tau \ - \ c(f_1 + f_2) \ - \ f_i c'(f_1 + f_2) \ - \ f^* c'(2f^*) \ = \ 0, \quad i = 1, 2. \tag{5}$$

which yield the solution $f_1 = f_2 = f^*$.

In the road context, where internalization of congestion does not occur, the Pigouvian toll would equal the full marginal congestion damage evaluated at the social optimum, given here by $\mathrm{MCD}^* = 2f^*c'$. But internalization means that a toll this large is not required, with z instead equal to $\mathrm{MCD}^*/2$.

3.2. Incentives for manipulation

As p varies over the business cycle, the socially optimal flight volume f^* varies in step, as does the magnitude of Pigouvian toll per flight, $f^*c'(2f^*)$. Since individual flight volumes are both equal to f^* in equilibrium, carriers would observe the toll varying with the current flight volume f according to the function $m(f) \equiv fc'(2f)$. Carrier i might then form an expectation as to how the toll would vary with unilateral changes in *its own* flight volume f_i. Given the observed pattern, a reasonable view might be that this variation would occur according to the function $m(\cdot)$, with carrier i expecting the toll to equal $m(f_i)$ when it chooses flight volume f_i.

Under this belief, carrier i would view its profit as equal to

$$(p-\tau)f_i \;-\; c(f_1+f_2)f_i \;-\; m(f_i)f_i \;\;=\;\; (p-\tau)f_i \;-\; c(f_1+f_2)f_i \;-\; f_i^2c'(2f_i). \tag{6}$$

In choosing f_i to maximize (6), the carrier engages in manipulation of the Pigouvian toll, exploiting its knowledge of the toll rule rather than taking the toll as parametric. Note also that, despite its belief about the toll, the carrier continues to behave in Cournot fashion with respect to congestion costs, treating f_{-i} in the $c(f_1+f_2)f_i$ term in (6) as parametric in choosing f_i. Differentiating with respect to f_i and imposing symmetry, the common individual flight volume (denoted $\widetilde{f}$) then satisfies

$$p \;-\; \tau \;-\; c(2\widetilde{f}) \;-\; 3\widetilde{f}c'(2\widetilde{f}) \;-\; 2\widetilde{f}^2c''(2\widetilde{f}) \;\;=\;\; 0. \tag{7}$$

Since the analogous condition from (2) determining the socially optimal flight volume is $p - \tau - c(2f^*) - 2f^*c'(2f^*) = 0$, and since (7) has a smaller left-hand side at a common f value, it follows that $\widetilde{f} < f^*$. Thus, under the assumed beliefs about toll liabilities, manipulation of the Pigouvian congestion toll leads to an inefficiently low flight volume.

While a carrier's assumed beliefs about its toll impact may be plausible, a different approach would assume that each carrier understands the rule used in setting the Pigouvian toll. In particular, even though the common toll z per flight ends up being charged, carrier i would know that the toll it pays is generated by the rule $f_{-i}c'(f_1+f_2)$. After multiplying this expression by f_i to get carrier i's toll liability and substituting in place of the last term in the profit expressison (6), a new version of (7) emerges (Cournot behavior with respect to f_{-i} still prevails). The coefficients on the fourth and fifth terms in (7) fall to 2 and 1 respectively, and while $\widetilde{f} < f^*$ continues to hold, the flight deficiency is not as large as before.

Manipulation can be generated in perhaps a more convincing fashion if congestion tolls are set using an iterative procedure, adjusting toward the optimum over a sequence of periods. Such an iterative approach, which might be used in any real world implementation of congestion tolls, would give ample opportunity for carriers to learn the toll rule as iterations proceed. For a discussion of toll manipulation under an iterative procedure, see the appendix.

3.3. A manipulable congestion toll

To circumvent manipulative behavior, the planner could abandon the Pigouvian toll and instead announce to the carriers a complete rule that determines their toll liabilities as a function of flight volumes. This toll rule is designed to be manipulated in the sense that carriers are given full information about the connection between toll liabilities and their flight-volume choices, which can be exploited in decision making. Although the toll rule is given for the carrier, the toll level has thus become "manipulable."

Let $T_i(f_1, f_2)$ denote the manipulable toll rule for carrier i, which gives its total toll liability as a function of both flight volumes (carrier i's toll per flight, t_i, is therefore equal to $T_i(f_1, f_2)/f_i$). Faced with this function, carrier i's first-order condition for choice of f_i is

$$p \; - \; \tau \; - \; c(f_1+f_2) \; - \; f_i c'(f_1+f_2) \; - \; \frac{\partial T_i(f_1,f_2)}{\partial f_i} \;\; = \;\; 0, \quad i=1,2. \tag{8}$$

The goal is to choose the T_i functions so that the solutions to the two conditions in (8) coincide with the social optimum. This goal can be achieved if the manipulable toll rule is chosen so that (8) is the same as the social optimality condition in (2). Inspection of the two conditions shows that this coincidence requires

$$\frac{\partial T_i(f_1,f_2)}{\partial f_i} \;\; = \;\; f_{-i}c'(f_1+f_2). \tag{9}$$

The left-hand side of (9) gives the marginal toll as perceived by a toll-manipulating carrier: the additional toll liability resulting from adding an extra flight. Note that the right-hand side of (9), when evaluated at the social optimum, is simply the Pigouvian toll z, which itself represents a marginal toll given the toll liability of zf_i. Thus, (9) implies that the marginal toll from the manipulable case (when evaluated at the social optimum) is equal to the marginal Pigouvian toll. Despite this coincidence of marginal tolls in the two cases, it will become clear below that the average tolls (z in the Pigouvian case vs. $T_i(f_1, f_2)/f_i$ in the manipulable case) will not be equal, nor will the total toll liabilities.

Integrating both sides of (9) with respect to f_i directly yields the manipulable toll rule, which is given by

$$T_i(f_1, f_2) \;\; = \;\; f_{-i}c(f_1 + f_2) \; + \; K_i. \tag{10}$$

where K_i is a constant of integration. Thus $T_1(f_1, f_2) = f_2c(f_1 + f_2) + K_1$ and $T_2(f_1, f_2) = f_1c(f_1 + f_2) + K_2$, so that a carrier's toll liability equals *the other carrier's total congestion cost* plus a constant. While K_i could be set at a fixed numerical value, it need only be independent of f_i and thus could be a function of f_{-i}, the other carrier's flight volume. Given this possibility, an natural choice is to set $K_i = -f_{-i}c(f_{-i})$, so that

$$T_i(f_1, f_2) \;\; = \;\; f_{-i}c(f_1 + f_2) \; - \; f_{-i}c(f_{-i}). \tag{11}$$

Then, carrier i's manipulable toll is zero when $f_i = 0$ and equals *the increase in the other carrier's congestion cost due to the flights operated by carrier i.*

It is interesting to compare the toll liabilities in the manipulable and Pigouvian cases. When the manipulable toll takes the form in (11), the toll liability equals $\int_0^{f_i} f_{-i}c'(x + f_{-i})dx$, while under Pigouvian tolling it equals $f_i f_{-i}c'(f_i + f_{-i})$, with both expressions evaluated at the social optimum to generate equilibrium values. If $c(\cdot)$ is strictly convex, the second expression is larger for any values of f_i and f_{-i}, implying that the toll liability under Pigouvian tolling exceeds that under manipulable tolling. However, given the freedom to adjust K_i, a different choice can make the manipulable toll's liability exceed that of the Pigouvian toll.

3.4. A restricted manipulable toll

The manipulable toll in (10) has a form that might be viewed as unappealing in a practical sense, given that it embodies a charge that depends on the other carrier's congestion cost. To eliminate this drawback, a different toll rule can be derived, subject to a more natural restriction on the rule's form. In particular, suppose that a carrier's toll liability is required to equal its own flight volume times a function, common to both carriers, that depends on the total flight volume at the airport. Stated differently, the requirement is that the *average toll* paid by a carrier, equal to its toll liability divided by its own flight volume, be given by a common function that depends on the flight total. Letting this average toll function be written $t(f_1 + f_2)$, carrier i's toll liability is then given by $f_i t(f_1 + f_2)$. Thus, the restriction implies $T_1(f_1, f_2) \equiv f_1 t(f_1 + f_2)$ and $T_2(f_1, f_2) \equiv f_2 t(f_1 + f_2)$.

Under this restriction, $\partial T_i(f_1, f_2)/\partial f_i = t(f_1 + f_2) + f_i t'(f_1 + f_2)$ holds, so that (9) becomes

$$t(f_1 + f_2) \;+\; f_i t'(f_1 + f_2) \;\;=\;\; f_{-i} c'(f_1 + f_2). \tag{12}$$

But recognizing that the flight volumes are symmetric in equilibrium, $f_1 = f_2 = \frac{F}{2}$ holds, where F is the total flight volume. Making this substitution, (12) can be rewritten as

$$t(F) \;+\; \frac{F}{2} t'(F) \;\;=\;\; \frac{F}{2} c'(F). \tag{13}$$

This condition is a linear, first-order differential equation in the unknown function $t(F)$. Given the presence of the 1/2 factor, the left-hand side expression cannot be integrated, apparently preventing derivation of a general solution. However, a solution can be derived under a fairly general functional form for $c(\cdot)$. In particular, if $c(F) \equiv \alpha F^\theta$, where $\theta \geq 1$, then a $t(F)$ function with same exponent but a different multiplicative factor can satisfy (13). Letting $t(F) = \beta F^\theta$, (13) reduces, after differentiation and substitution, to the following requirement:

$$\beta F^\theta \;+\; \frac{F}{2} \beta \theta F^{\theta - 1} \;\;=\;\; \frac{F}{2} \alpha \theta F^{\theta - 1}, \tag{14}$$

which is satisfied when $\beta = \alpha\theta/(2+\theta)$. Thus, the manipulable average toll function for this special case is given by

$$t(F) \quad = \quad \frac{\alpha\theta}{2+\theta}F^{\theta} \tag{15}$$

When carrier i faces a toll liability of $f_i t(f_1+f_2) = f_i \frac{\alpha\theta}{(2+\theta)}(f_1+f_2)^{\theta}$, the resulting profit-maximizing flight volumes are socially optimal.

The equilibrium toll payment under this restricted manipulable function is smaller than in the Pigouvian case. Under the assumed form of $c(\cdot)$, the Pigouvian toll per flight is $(F^*/2)c'(F^*) = (\alpha\theta/2)F^{*\theta}$, which exceeds the average manipulable toll in equilibrium, given by (15) evaluated at F^* (the socially optimal total flight volume). This relationship also holds in general, as can be seen from inspection of (13). Since the right-hand side evaluated at F^* is the Pigouvian toll per flight, it follows that $t(F^*)$ must be smaller than this magnitude as long as $t'(\cdot)$ is positive. Intuitively, the average manipulable toll is less than the Pigouvian toll when the average toll is increasing since this effect provides an additional deterrent in limiting a carrier's flight volume.

3.5. Informational advantages of manipulable tolls

Computation of the Pigouvian toll in (4) requires substantial information on the part of the planner. The planner must know the congestion cost function $c(\cdot)$ to make use of the Pigouvian rule, but since the socially optimal flight volume f^* is also required to compute the toll, the planner must also know the full price p and the airline cost parameter τ, which help determine f^*.

By contrast, the requirements for implementation of a manipulable toll, either in its general or restricted version, are much more modest. In particular, since the manipulable rules do not depend on f^*, the planner need only know the congestion cost function, on which the rules depend. In addition to their superior incentive properties, manipulable toll rules thus have a clear informational advantage over the Pigouvian rule.

3.6. Airport cost recovery

The well-known self-financing theorem from road pricing theory says that, when roads are built with neutral scale economies and an additional zero-degree homogeneity assumption holds for the congestion function (*i.e.*, a doubling of traffic flow and capacity would leave the travel time constant), the cost of the optimal-size road is exactly covered by Pigouvian congestion toll revenue (Mohring and Harwitz, 1962). This result fails to hold, however, in the airport context due to internalization of congestion, which generates smaller tolls than in the road setting (for a formal derivation, see Brueckner (2002, 2008)).

Given the conclusions derived above, the revenue shortfall is even larger under a restricted manipulable toll rule. In other words, since the average manipulable toll is smaller than the Pigouvian toll in equilibrium, and since the latter toll itself already fails to cover airport cost, a larger deficit emerges in the manipulable case. Since the general manipulable toll also generates a smaller liability than the Pigouvian toll when it takes the form in (11) and $c(\cdot)$ is convex, cost recovery again fails in this case. The deficit under both types of manipulable tolls must be covered by additional lump-sum charges.

4. The General Model: Asymmetric Carriers and Imperfectly Elastic Demand

4.1. The unrestricted manipulable toll

Consider now the more general case where carriers need not be symmetric and demand is not necessarily perfectly elastic. Social surplus is now defined by

$$W \;=\; \int_0^{f_1+f_2} D(x)dx \;-\; \sum_{j=1}^{2} f_j(\tau_j \;+\; c_j(f_1+f_2)), \tag{16}$$

and the social optimum requires

$$D(f_1+f_2) \;-\; \tau_i \;-\; c_i(f_1+f_2) \;-\; f_i{c_i}'(f_1+f_2) \;-\; f_{-i}{c_{-i}}'(f_1+f_2) \;=\; 0, \quad i=1,2, \tag{17}$$

where the subscript $-i$ on $c(\cdot)$ denotes the function belonging to carrier i's competitor.

Using (1), the first-order condition for profit maximization in the absence of tolls is

$$D(f_1+f_2) \;+\; f_iD'(f_1+f_2) \;-\; \tau_i \;-\; c_i(f_1+f_2) \;-\; f_i{c_i}'(f_1+f_2) \;=\; 0, \quad i=1,2. \tag{18}$$

When a Pigouvian toll is used to eliminate the difference between (17) and (18), carrier i's toll equals

$$z_i \;=\; f_i^* D'(f_1^* + f_2^*) \;+\; f_{-i}^* c_{-i}{}'(f_1^* + f_2^*), \tag{19}$$

as seen in Pels and Verhoef (2004). The asterisks again denote socially optimal values, which are now asymmetric, as are the Pigouvian tolls themselves. Note that the first term in (19) is negative when demand is imperfectly elastic, indicating that the toll is adjusted downward to mitigate over-pricing by carriers with market power. This correction vanishes as carriers become infinitesimally small and f_i approaches zero. The downward adjustment of Pigouvian taxes under under market power was originally derived by Buchanan (1969), and it further erodes the scope for self-financing of optimal airport capacity. Observe that the market-power adjustment in (19) is larger for bigger carriers, offsetting their greater incentive to restrict output in order to raise the fare.

As before, the manipulable toll rule $T_i(f_1, f_2)$ for carrier i should be set so that the marginal toll coincides with the Pigouvian toll rule. Using (19), this requirement yields

$$\frac{\partial T_i(f_1, f_2)}{\partial f_i} \;=\; f_i D'(f_1 + f_2) \;+\; f_{-i} c_{-i}{}'(f_1 + f_2), \quad i = 1, 2. \tag{20}$$

The manipulable toll rule is found by integrating (20), which yields

$$T_i(f_1, f_2) \;=\; \int_0^{f_i} x D'(x + f_{-i}) dx \;+\; f_{-i} c_{-i}(f_1 + f_2) \;+\; K_i, \quad i = 1, 2. \tag{21}$$

Integrating the first terms by parts and imposing the previous requirement $K_i = -f_{-i} c_{-i}(f_{-i})$ on the constants of integration, (21) reduces to

$$T_i(f_1, f_2) = -\left[\int_0^{f_i} D(x + f_{-i}) dx - f_i D(f_1 + f_2)\right] \;+\; f_{-i}(c_{-i}(f_1 + f_2) - c_{-i}(f_{-i})), \; i = 1, 2. \tag{22}$$

Note that (22) equals the increase in the other carrier's congestion cost due to carrier i's operations minus the addition to consumer surplus from these operations. With toll liability rule (22) in place, airline i's profit function becomes effectively equal to social surplus above the level achieved if only f_{-i} were supplied. This fact can be verified by substitution of (22) into the expression giving profit net of the toll and comparing with (16). The intuition is simple: if a toll manipulator is to be seduced to behave so as to maximize social surplus, the toll liability should be defined so that profit inclusive of this liability should vary perfectly in parallel with social surplus. In hindsight, it would have been surprising if this would *not* have been true.[3]

The distributional impact of the toll schedule, rewarding the carrier with the full increase in social surplus it creates given the competitor's output, may of course be considered undesirable. The K_i terms may then be used to adjust this impact by setting them at values different from $K_i = -f_{-i}c_{-i}(f_{-i})$. But what the analysis shows is that post-tax profits should vary perfectly in parallel with social surplus for the manipulable toll to be optimal, a property that is independent of the choice of K_i. This requirement leaves the K_i terms as instruments to address distributional concerns.

As in the case of the Pigouvian toll from (20), the total toll liability in (24) may be positive (a net tax) or negative (a net subsidy). The latter outcome emerges if the integral of the inverse demand function is large relative to the congestion externality plus the carrier's revenues. This ambiguity reflects the use of a single toll instrument to address two opposing distortions: a market-power distortion that inefficiently limits flight volumes, and a congestion externality that inefficiently inflates them. As suggested by Brueckner (2005), however, these two distortions could be addressed by different instruments in a situation where the distortions can be separated. Such a situation arises when the carriers serve multiple markets, with each providing service between a common set of n uncongested endpoints and the same congested airport (which might be a hub). Since they involve different endpoint cities, such markets may have different demand functions, leading to market-power distortions of different magnitudes. These distortions can then be addressed by market-level subsidies, while congestion is addressed by separate airport-level tolls, either of Pigouvian or manipulable form. With two asymmetric carriers, the required subsidies would be both carrier and market-specific, yielding $2n$ different subsidies in the various markets. In addition, two different, carrier-specific tolls would be levied at the congested airport. The subsidies and tolls would separately represent the market-power and congestion terms, suitably generalized to multiple markets, from either the Pigouvian toll formula (20) or the manipulable formula (23).

A final point regarding the correction of the market-power distortion concerns first-degree price discrimination. Under such behavior, every ticket would be sold at a price equal to the consumers maximum willingness to pay, and the market-power distortion would vanish. Although price discrimination in fare setting is standard practice in the airline industry, it does not take on this ideal form, so that some downward adjustment of congestion tolls (or the payment of separate market-level subsidies) will remain necessary to compensate for excessive fares.

4.2. Feasibility of a manipulable rule with a common average toll

As seen in section 3.2, a restricted manipulable toll, where both carriers are charged the same average toll, is feasible in the basic model. A natural question is whether such a toll is feasible in the general setting under consideration. One might expect that carrier asymmetry makes a restricted toll like that derived above infeasible, and the ensuing analysis shows that this conjecture is indeed correct. Even though this result may come as no surprise, the analysis further illuminates the issues involved in deriving manipulable toll rules.

As before, when carriers pay a common average toll $t(\cdot)$, the marginal manipulable toll is given by $\partial T_i(f_1, f_2)/\partial f_i = t(f_1+f_2)+f_i t'(f_1+f_2)$, so that the marginal-toll condition (20) must hold with the latter expression replacing the left-hand side expression. In the previous analysis, the analog to (20) (namely, eq. (10)) reduced to the single condition (11) given symmetry of the carriers. However, with asymmetry, (20) remains as two separate equations involving the level and derivative of the desired average toll function $t(\cdot)$. These equations can be solved for these unknown quantities, yielding

$$t(f_1 + f_2) \;=\; \frac{{f_1}^2}{f_1 - f_2} c_1'(f_1, f_2) \;-\; \frac{{f_2}^2}{f_1 - f_2} c_2'(f_1, f_2) \tag{23}$$

$$t'(f_1 + f_2) \;=\; -\frac{f_1}{f_1 - f_2} c_1'(f_1, f_2) \;+\; \frac{f_2}{f_1 - f_2} c_2'(f_1, f_2) \;+\; D'(f_1, f_2). \tag{24}$$

Although a restricted toll function could be derived in the basic model by finding the solution to a differential equation, a similar outcome is not feasible here. To see the difficulty, observe that the left-hand of (24) is *not equal* to the derivative of left-hand side of (23), which means that a function $t(\cdot)$ satisfying these two conditions for all values of f_1 and f_2 does not exist. While this conclusion means that a restricted manipulable toll rule like that in section 3 is infeasible, it is possible to generate a "fake" manipulable rule, one that is computed using knowledge of the social optimum. In this sense, the toll resembles the Pigouvian toll in (20), whose computation requires all the information needed to derive the socially optimal values f_1^* and f_2^*. In particular, it can be verified that the linear average toll function given by

$$\widetilde{t}(f_1 + f_2) \quad = \quad \delta + \beta(f_1 + f_2) \tag{25}$$

yields a manipulable toll rule that generates the social optimum, where the slope and intercept terms δ and β are derived using (23) and (24) and themselves depend on f_1^* and f_2^*.[4]

A toll rule with a common average toll per flight may be appealing on equity grounds, but despite the negative conclusion just derived, equity could be achieved in a different fashion. Specifically, since the regulator can adjust toll liabilities under the general manipulable toll by varying the K_i terms in (21), any distributional goal can be met even though the desired type of restricted toll function is not available.

5. Conclusion

The recent literature on congestion pricing with large agents contains a remarkable inconsistency: though agents are large enough to recognize self-imposed congestion and exert market power over prices, they do not take account of the impact of their own actions on the magnitude of congestion tolls. When large agents are confronted with tolls derived under this parametric assumption but understand the rule used to generate them, the toll system will no longer guide the market to the social optimum. To address this problem, the present paper has derived an alternative, manipulable toll rule, which is designed to achieve the social optimum when agents do take into account the full impacts of their own actions on total toll liabilities. The analysis shows that, although the marginal tolls do not differ between the conventional

Pigouvian and manipulable cases, the average and total toll liabilities generally will be different. In addition, revenues from manipulable tolling are lower than under parametric Pigouvian tolling, further reducing the scope for self-financing of capacity. But manipulable tolls have an informational advantage over the Pigouvian variety, enhancing their practicality.

Appendix

This appendix analyzes manipulative behavior when tolls are set according to an iterative procedure. To implement this procedure, the airport authority need only know the congestion-cost function $c(\cdot)$, with knowledge of p and τ (and thus f^*) not required. To develop the procedure, let the above static model be modified by the introduction of distinct time periods, indexed by $k = 1, 2, \ldots$, and suppose that tolls are initially viewed as parametric. In period k, the airport authority levies a congestion toll computed using the Pigouvian formula *evaluated at the flight volume from the previous period*, $k-1$. For example, if the laissez-faire equilibrium obtains in period 0 and tolling starts in period 1, the toll per flight in period 1 would equal $\widehat{f}c'(2\widehat{f})$, the Pigouvian rule evaluated at period 0's (laissez-faire) flight volume. Recognizing that flight volumes in each period will be symmetric across carriers, and letting f^k denote this common volume for period k, f^k will satisfy

$$p \;-\; \tau \;-\; c(2f^k) \;-\; f^k c'(2f^k) \;-\; f^{k-1}c'(2f^{k-1}) \;\;=\;\; 0, \qquad (a1)$$

where $f^{k-1}c'(2f^{k-1})$ gives the (parametric) toll per flight in period k. Assuming convergence of the iterative procedure, its limit can be found by setting f^k and f^{k-1} equal to a common value $\widetilde{f}$, which must satisfy

$$p \;-\; \tau \;-\; c(2\widetilde{f}) \;-\; 2\widetilde{f}c'(2\widetilde{f}) \;\;=\;\; 0. \qquad (a2)$$

This condition, however, is the same as the social optimality condition, implying $\widetilde{f} = f^*$ and thus that the limit of the iterative procedure is efficient.

However, the assumed parametric view of tolls may be implausible. As in the case where p varies, changes in the toll from period to period as the interative procedure converges could allow carriers to learn the form of the toll rule and to exploit this knowledge in their decisions. Carriers observe that the toll per flight paid in period $k+1$ is equal to $m(f^k) \equiv f^k c'(2f^k)$, but as the analysis in section 3.2, carrier i might again believe that the toll it pays is actually determined by $m(f_i^k)$, the function $m(\cdot)$ evaluated at its own flight volume. Then, in choosing f_i^k, carrier i would expect the choice to affect both profit in period k and the toll (and hence profit) in period $k+1$. Discounting the latter profit by the factor $\delta \leq 1$, carrier i's maximand would be

$$(p - \tau)f_i^k - c(f_1^k + f_2^k)f_i^k - [f^{k-1}c'(2f^{k-1})]f_i^k - \delta[f_i^k c'(2f_i^k))]f_i^{k+1}, \qquad (a3)$$

where the third term is the congestion toll liability in period k (dependent on period $k-1$'s common flight volume) and the fourth term is the discounted toll liability expected in period $k+1$, the only part of that period's profit that depends on f_i^k. Differentiating $(a3)$ with respect to f_i^k and imposing symmetry, the result is a condition like (4) with k superscripts added and the additional term $\delta(c'(2f_i^k) + 2f_i^k c''(2f_i^k)]f^{k+1}$ subtracted. Assuming convergence, the new $\widetilde{f}$ value then satisfies

$$\begin{aligned} p - \tau - c(2\widetilde{f}) - 2\widetilde{f}c'(2\widetilde{f}) - \delta(c'(2\widetilde{f}) + 2\widetilde{f}c''(2\widetilde{f})]\widetilde{f} &= \\ p - \tau - c(2\widetilde{f}) - (2+\delta)\widetilde{f}c'(2\widetilde{f}) - 2\delta\widetilde{f}^2 c''(2\widetilde{f}) &= 0, \qquad (a4) \end{aligned}$$

a condition that reduces to (6) if $\delta = 1$. Since the left-hand side expression in $(a4)$ is smaller than the expression in (2) at a common f value, it again follows that $\widetilde{f} < f^*$. Thus, manipulation of the iterative Pigouvian toll again leads to an inefficiently low flight volume.

References

Basso, L.J., Zhang, A., 2007. Congestible facility rivalry in vertical structures, *Journal of Urban Economics* 61, 218-237.

Brueckner, J.K., 2002. Airport congestion when carriers have market power, *American Economic Review* 92, 1357-1375.

Brueckner, J.K., 2005. Internalization of airport congestion: A network analysis. *International Journal of Industrial Organization* 23, 599-614.

Brueckner, J.K., 2008. Price vs. quantity-based approaches to airport congestion management. Unpublished paper, UC Irvine.

Brueckner, J.K., Van Dender, K., 2008. Atomistic congestion tolls at concentrated airports? Seeking a unified view in the internalization debate. *Journal of Urban Economics* 64, 288-295.

Buchanan, J.M., 1969. External diseconomies, external taxes, and market structure. *American Economic Review* 59, 174-177.

Daniel, J.I., Pahwa, M., 2000. Comparison of three empirical models of airport congestion pricing. *Journal of Urban Economics* 47, 1-38.

Daniel, J.I., 2001. Distributional consequences of airport congestion pricing. *Journal of Urban Economics* 50, 230-258.

Daniel, J.I., 1995. Congestion pricing and capacity of large hub airports: A bottleneck model with stochastic queues, *Econometrica* 63, 327-370.

Daniel, J.I., Harback, K.T., 2008a. (When) Do hub airlines internalize their self-imposed congestion delays? *Journal of Urban Economics* , 63, 583-612.

Daniel, J.I., Harback, K.T., 2008b. Pricing the major hub airports. *Journal of Urban Economics*, forthcoming.

Daniel, J.I., Harback, K.T., 2008c. Pricing Canadian airports. Unpublished paper, University of Delaware.

Forbes, S.J., 2008. The effect of air traffic delays on airline prices. *International Journal of Industrial Organization* 26, 1218-1232.

Hahn, R.W., 1984. Market power and transferable property rights. *Quarterly Journal of Economics* 99, 753-765.

Johnson, T., Savage, I., 2006. Departure delays, the pricing of congestion, and expansion proposals at Chicago's O'Hare airport. *Journal of Air Transport Management* 12, 182-190.

Mohring, H., Harwitz, M., 1962. *Highway Benefits: An Analytical Framework.* Northwestern University Press, Evanston, IL.

Morrison, S.A., Winston, C., 2007. Another look at airport congestion pricing. *American Economic Review* 97, 1970-1977.

Pels, E., Verhoef, E., 2004. The economics of airport congestion pricing. *Journal of Urban Economics* 55, 257-277.

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

Verhoef, E., 2008. Congestion pricing, slot sales and slot trading in aviation. Unpublished paper, Free University of Amsterdam.

Zhang, A., Zhang, Y., 2006. Airport capacity and congestion when carriers have market power. *Journal of Urban Economics* 60, 229-247.

Footnotes

[1]See Brueckner (2008), Brueckner and Van Dender (2008), Verhoef (2008), Daniel and Pahwa (2000), Daniel (2001), Daniel and Harback (2008a,b,c), Zhang and Zhang (2006), Basso and Zhang (2007), Johnson and Savage (2006), and Morrison and Winston (2007).

[2]A related concern about manipulation arises in markets for pollution permits, where large polluters could manipulate the permit price to their advantage. See Hahn (1984) for an early study investigating this issue. Note that a similar phenomenon could emerge in the market for airport slots, which represent an alternate instrument (analogous to pollution rights) for dealing with airport congestion (see Brueckner (2008) and Verhoef (2008) for analyses of slot markets).

[3]Note that, although computation of the manipulable toll rule now requires knowledge of both the demand and congestion-cost functions, one information advantage remains relative to the Pigouvian toll: the planner need not know the τ_i parameters.

[4]In (25), β equals left-hand side of (24) evaluated at f_1^* and f_2^*, while δ equals the left-hand side of (23) evaluated at the optimum minus $\beta(f_1^* + f_2^*)$.

[5]Note that, since the only required information is the $c(\cdot)$ function, the iterative procedure can be implemented without knowing the magnitude of f^*, in contrast to the usual Pigouvian toll.