

# 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*

January 2009, revised September 2009

## **Abstract**

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 into account 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 derives alternate, manipulable toll rules, which are designed to achieve the social optimum when agents anticipate the full impact of their actions on 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. An alternative to this “classical” Pigouvian approach, which is useful when the regulator knows the damage function but lacks all the information needed to compute the optimum, is to charge a toll equal to marginal damage evaluated at the expected equilibrium activity level, and to revise the tax if this expectation is not fulfilled. In equilibrium, the expectation is confirmed, and the socially optimal activity level again emerges (Kraus and Mohring, 1975).

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 this rule at a particular activity level.

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. Another reason why agents could be aware of the rule generating the Pigouvian tax is that the principles under which the tax level is determined may need to be transparent and publicly accessible for legal reasons. Agents can then observe the rule directly. In either case, with agents possessing this knowledge, the regulator cannot credibly pretend that the tax is parametric. 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 they pay. Such behavior undermines the Pigouvian 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.

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. For concreteness, the paper's analysis is developed in the airport context. Although airport congestion pricing according to economic principles does not yet exist in practice, such an approach is likely to be considered as a serious policy option in the near future, as capacity limits of airports are approached and expansion is hindered for reasons of space, environmental concerns, noise, or other considerations. Moreover, various studies of airline pricing behavior have shown that airlines typically behave like Cournot oligopolists, implying that manipulation of market prices is not uncommon in aviation. This behavior makes it more likely that Pigouvian tolls, when implemented to manage congestion, will also be manipulated by airlines when they see sufficient scope for doing so. This likelihood motivates our focus on the aviation 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.<sup>1</sup> 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 (1995) 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.<sup>2</sup>

These studies looked into different aspects of airport congestion management, such as the relation with airport capacity choices, different ways to model congestion, and Stackelberg leadership of one of the airlines. But as our discussion indicates, all of this recent analysis

of airport congestion presumes that, when congestion tolls are in place, airlines treat them as parametric. This approach implies a remarkable inconsistency, given that the assumed non-atomistic nature and strategic behavior of airlines suggests the potential for manipulation of such tolls. The present paper is meant to redress this omission. By analyzing the nature of 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, specifically when agents are large relative to the total size of the externality in the relevant market. This condition is, of course, more likely to be met in case of local externalities, such as pollution, noise, etc.<sup>3</sup>

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. Section 3.1 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. Then, section 3.2 explores the incentives for toll manipulation. Manipulable tolls, required in response to manipulative behavior, are analyzed next. The analysis in section 3.3 derives the most general form for such tolls, and section 3.4 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. Section 3.5 considers the issue of airport cost recovery under the different toll regimes. 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. Sections 4.1 and 4.2 respectively consider general and restricted toll schedules for the general case, with the latter involving a common toll per flight across the now-asymmetric carriers. Section 5 considers application of the lessons of the analysis to other externality contexts, and section 6 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, such as Daniel (1995) and Brueckner (2002), 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 a Cournot-Nash 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 would include the airfare and the value of lost passenger time due to airport congestion. However, without loss of generality, these time costs are suppressed, so that  $D(f_1 + f_2)$  equals the fare. With time costs absent, the sole effect of airport congestion is then to raise airline operating costs. This congestion effect is allowed to differ across carriers, with congestion cost per flight for carrier  $i$  given by  $c_i(f_1 + f_2)$ ,  $i = 1, 2$ , functions of the total flight volume. These functions are assumed to be increasing and convex over the relevant range (they may be zero at low flight volumes). 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  $\tau_i$  is operating cost per flight,  $i = 1, 2$ .

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

$$\pi_i = D(f_1 + f_2)f_i - \tau_i f_i - c_i(f_1 + f_2)f_i, \quad i = 1, 2. \quad (1)$$

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

$$p - \tau - c(f_1 + f_2) - (f_1 + f_2)c'(f_1 + f_2) = 0. \quad (2)$$

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  $(f_1 + f_2)c'(f_1 + f_2) \equiv \text{MCD}$ . 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-Nash 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

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

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  $\hat{f}$ , are too large, satisfying  $\hat{f} > f^*$ . 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 decentralized optimum is modeled as the outcome of a game in which the regulator sets the toll acting as a Stackelberg leader, taking into account how a parametric toll affects the airlines' quantity choices. The airlines take the toll as given, and behave as Cournot-Nash duopolists in choosing their flight volumes, as they were assumed to do in the laissez-faire equilibrium described above. 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 classical Pigouvian toll is then given by<sup>4</sup>

$$z = f^* c'(2f^*). \quad (4)$$

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

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

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  $\text{MCD}^* = 2f^* c'$ . But internalization of self-imposed congestion means that a toll this large is not required, with  $z$  instead equal to  $\text{MCD}^*/2$ .

### *3.2. Incentives for manipulation*

While the “classical” Pigouvian approach embodied in (4) assumes that the regulator has sufficient information to compute the optimum, an actual application of the Pigouvian method might be based on incomplete information. The regulator, for example, may lack information on the cost parameter  $\tau$  or the full price  $p$ . The regulator is then unable to predict the optimum beforehand, and therefore unable to determine the optimal value of the parametric Pigouvian toll before implementing it. Assuming, however, that the regulator

knows the congestion cost function  $c(\cdot)$ , a variant of the classical approach can be envisioned. It generates the same optimal outcome as the classical approach when carriers would treat the tolls as parametric, as assumed by the regulator. But because the regulator will compute and impose a sequence of parametric tolls, leading to successive equilibria, the approach may in fact invite toll manipulation, so that this assumption of “toll-taking behavior” is not justified.

Under this alternate approach, the regulator computes the congestion toll per flight using the Pigouvian rule  $f_{-i}c'(f_1 + f_2)$ . But rather than evaluating the rule at the social optimum, which the regulator is unable to compute beforehand, the rule is evaluated at the flight volumes that the regulator expects to emerge in the equilibrium with tolling. This expected equilibrium is, therefore, the regulator’s expectation of the optimum. These expectations must subsequently be confirmed, however, by carrier choices. Given their perception of the toll (as parametric or, alternatively, subject to manipulation), the flight-volume choices of the carriers must match the regulator’s expectations regarding these volumes. If not, the regulator would adjust the parametric tolls, just as described in Kraus and Mohring (1975), until expectations are confirmed. The equilibrium is then achieved, and the regulator believes it must be the decentralized optimum, since it is the equilibrium consistent with imposition of optimal Pigouvian tolls.

In the case where the carriers view the toll as parametric, the equilibrium requirement for this decentralized optimum is straightforward. In particular, the Pigouvian toll per flight  $f_{-i}c'(f_1 + f_2)$ , which is viewed as parametric, is substituted in place of  $f^*c'(2f^*)$  in (5). Evaluating the Pigouvian rule at the flight volumes that satisfy (5) yields a toll that, when perceived as parametric by the carriers, elicits these same flight volumes as choices. Thus, the regulator’s expectations in computing the toll are confirmed. Since the modified version of (5) coincides with the optimality condition (2), the resulting equilibrium is efficient, just like the one based on the classical Pigouvian toll. With perfectly elastic demand, positively sloped convex congestion costs, and a  $p$  that exceeds marginal costs at a zero a volume, there will be one and only one symmetric equilibrium in which the regulator’s expectations are met (the same conclusion also applies when demand is imperfectly elastic).<sup>5</sup>

Now consider the case where the airlines know the toll rule, perhaps because the parametric

toll has been adjusted a number of times in response to changes in the equilibrium (due to variation in  $p$  or  $\tau$ , for example) or because the toll rule is public. In addition, suppose that the carriers use this knowledge to their advantage, so that they no longer behave as the Stackelberg-followers who take the parametric toll as given when playing their Cournot-Nash game. Instead, they take the toll *rule* as given. Specifically, carrier  $i$  would view profit as being given by

$$(p - \tau)f_i - c(f_1 + f_2)f_i - f_{-i}c'(f_1 + f_2)f_i, \quad (6)$$

but the toll per flight expression ( $f_{-i}c'(f_1 + f_2)$ ) appearing in the last term *would no longer be viewed as a parametric value*. Carriers would then optimize taking account of the effect of their decisions on the toll paid per flight. Computing the relevant first-order condition and evaluating at the symmetric equilibrium, the resulting common flight volume, denoted  $\tilde{f}$ , satisfies

$$p - \tau - c(2\tilde{f}) - 2\tilde{f}c'(2\tilde{f}) - \tilde{f}^2c''(2\tilde{f}) = 0. \quad (7)$$

The regulator expects symmetric flight volumes of  $\tilde{f}$  and thus charges a toll per flight of  $\tilde{f}c'(2\tilde{f})$ . But (7) ensures that carrier choices, which now involve manipulation of the toll, generate a common flight volume of  $\tilde{f}$ , confirming these expectations. The key observation from (7), however, is that since  $c'' > 0$ , the equation is satisfied at an  $\tilde{f}$  value smaller than  $f^*$ . Thus, manipulation of the (non-classical) Pigouvian congestion toll leads to an inefficiently low flight volume.<sup>6</sup>

### 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, intentionally, become “manipulable.” The structure of the game is then such that the regulator still acts as a Stackelberg leader, but now having

the toll *rule* rather than the parametric toll *level* as its instrument. The carriers are followers who take this toll rule as given, and whose quantity choices represent the equilibrium of the Cournot-Nash game that they play among themselves.

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

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

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

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

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 equal to the level of the classical Pigouvian toll  $z$ , which itself represents a marginal toll given the toll liability of  $z f_i$ . 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 classical 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

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

where  $K_i$  is a constant of integration. Thus  $T_1(f_1, f_2) = f_2 c(f_1 + f_2) + K_1$  and  $T_2(f_1, f_2) = f_1 c(f_1 + f_2) + K_2$ , 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, a natural choice is to set  $K_i = -f_{-i}c(f_{-i})$ , so that

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

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  $i$* .

To see the difference in average tolls between the manipulable and Pigouvian cases, observe that average manipulable toll from (10) is  $[f_{-i}c(f_1 + f_2) + K_i]/f_i$ , while the average Pigouvian toll is  $f_{-i}c'(f_1 + f_2)$ .<sup>7</sup> To compare the total toll liabilities in the two cases, suppose the manipulable toll takes the form in (11). Then the toll liability can be rewritten as  $\int_0^{f_i} f_{-i}c'(x + f_{-i})dx$ , while under Pigouvian tolling the liability equals  $f_i f_{-i}c'(f_i + f_{-i})$ . 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 (evaluated at the social optimum). However, given the freedom to adjust  $K_i$ , a different choice can make the manipulable toll's liability exceed that of the Pigouvian toll.

Because  $T_i(f_1, f_2)$  gives the total toll liability as a function of both flight volumes, the question arises whether the choices of volumes, given the announced toll liability schedules, will be mutually consistent between airlines. What if they are not? The chosen game theoretic set-up, with one-shot Nash equilibria, does not allow us to describe dynamic adjustment toward the decentralized optimum in case these choices are inconsistent. Indeed, the same question could in fact be posed for the laissez-faire equilibrium, which also assumes that the two flight volumes are mutually consistent, with each carrier taking the other carrier's output as given. The decentralized optimum with the manipulable toll rule (11) in place is an equilibrium in the sense that the regulator does not have an incentive to change the toll rule, taking into account how carriers respond to that rule, and the carriers have no incentives to change volumes, taken the toll rule and the other carrier's output as given.

Finally, it should be noted that, since it charges carrier  $i$  for the congestion imposed on the

other carrier, the manipulable toll in (11) bears an intriguing resemblance to Vickrey-Clarke-Groves-Ledyard mechanisms (Vickrey, 1961; Clarke, 1971; Groves and Ledyard, 1976). But in our problem, the regulator is not applying the mechanism to elicit private information from the airlines, as is the case in these revelation mechanisms. The similarity between the different schemes, then, appears to reflect a more general principle, namely, that individual optimizing behavior can be aligned with social surplus maximization only under a mechanism where the individual faces incentives that reflect the full social costs and benefits of her or his actions.

#### 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

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

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

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

This condition is a linear, first-order differential equation in the unknown function  $t(\cdot)$ . Multiplying through by  $2F$ , the left-hand side becomes  $2Ft(F) + F^2 t'(F)$ , an expression whose

antiderivative is  $F^2 t(F)$ . Integrating the modified right-hand side of (13) (equal to  $F^2 c'(F)$ ) and rearranging, the  $t(\cdot)$  function satisfying (13) is then given by

$$t(F) = \frac{1}{F^2} \int F^2 c'(F) dF + \frac{H}{F^2}, \quad (14)$$

where  $H$  is a constant. To see what (14) implies in a familiar case, suppose that  $c(F) \equiv \alpha F^\theta$ , where  $\theta \geq 1$ . Evaluating the integral in (14) and setting  $H = 0$ ,

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

Thus, 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 the restricted manipulable function is smaller than in the Pigouvian case. Under the assumed form of  $c(\cdot)$ , the marginal Pigouvian toll per flight (either in the classical case or the non-classical equilibrium) 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. 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, 2009)).

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), cost recovery again fails in this case. The deficit under both types of manipulable tolls must be covered by additional lump-sum charges. Obviously, the constants  $K_i$  could be used to boost the degree of self-financing of airport capacity.

#### 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

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

and the social optimum requires

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

where the subscript  $-i$  on  $c(\cdot)$  denotes the function belonging to carrier  $i$ 's competitor. Note that (17) differs from (2) in the lack of symmetry and the appearance of  $D(\cdot)$  in place of  $p$ .

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

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

Note that the second term in (18) arises from the exploitation of market power, which allows the carriers to raise the fare by limiting flights. When a classical Pigouvian toll is used to

eliminate the difference between (17) and (18), carrier  $i$ 's toll equals

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

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 the carriers. This correction vanishes as carriers become infinitesimally small and  $f_i$  approaches zero. The downward adjustment of Pigouvian taxes 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

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

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

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

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

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

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 varies perfectly in parallel with social surplus. In hindsight, it would have been surprising if this would *not* have been true.

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 (22) may be positive (a net tax) or negative (a net subsidy). The latter outcome emerges when consumer surplus is large relative to the congestion externality (requiring sufficiently inelastic demand). 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, for example, 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 (22).

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 consumer's 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

$$t(f_1 + f_2) = \frac{f_1^2}{f_1 - f_2} c'_1(f_1, f_2) - \frac{f_2^2}{f_1 - f_2} c'_2(f_1, f_2) \quad (23)$$

$$t'(f_1 + f_2) = -\frac{f_1}{f_1 - f_2} c'_1(f_1, f_2) + \frac{f_2}{f_1 - f_2} c'_2(f_1, f_2) + D'(f_1, f_2). \quad (24)$$

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 right-hand side of (24) is *not equal* to the derivative of the right-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 classical Pigouvian toll in (19), 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

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

yields a manipulable toll rule that generates the social optimum, where the slope and intercept terms  $\delta$  and  $\beta$  are derived using (23) and (24) and themselves depend on  $f_1^*$  and  $f_2^*$ .<sup>8</sup>

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 airport authority 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. Manipulable Taxes in Other Contexts

The principles of the preceding analysis can be applied in other contexts where manipulation of Pigouvian taxes may occur. Consider, for example, the pollution context, where two firms generate pollution as a by-product of production, with the externality damage a function of the sum of these two levels (this damage is suffered by other agents, not the firms themselves). A firm’s pollution level depends on its output along with the level of a costly abatement activity, and the demand for each firm’s output is assumed to be perfectly elastic.<sup>9</sup>

In this case, the social optimum requires equality between the marginal profit from an extra unit of output and the marginal pollution damage from that unit (equal to the additional pollution times marginal pollution damage). In addition, the marginal benefit from abatement (equal to abatement productivity times marginal pollution damage) must equal the marginal

abatement cost. Together, these two conditions determine the firms' optimal output and abatement levels. Provided that the firms view any pollution tax as parametric, the optimum can be generated by a Pigouvian tax set equal to marginal pollution damage, evaluated at the socially optimal pollution level under the classical approach or at the expected equilibrium level under the alternate approach.

When firms manipulate the Pigouvian tax, the optimum fails to emerge, and a manipulable tax is then required. The form of this tax can be predicted given the previous analysis. In particular, each firm should pay a tax liability *equal to total pollution damage*, adjusted by some constant. When such a tax, whose structure is understood by the firms, is subtracted from profit, optimal output and abatement levels are then chosen. It should be noted that, since pollution damage is suffered by agents other than the firms themselves, the manipulable tax structure does not share a key feature of the previous congestion tolls: charging for damage to the other firm. Instead each firm is charged for damage done to (outside) pollution victims, adjusted by constant.

A manipulable Pigouvian tax would have a similar form in other contexts. These contexts would share the key airport feature of having actors that are large relative to the total externality in the relevant market. This condition is more likely to hold when the externality is local in character, with examples including noise nuisance from (for example) rail transport, bad odors from specific factories, or public risks from particular productive activities. Furthermore, the problem of manipulable Pigouvian subsidies for positive externalities will be conceptually identical to the case considered here. But because agents may have a greater incentive to internalize the external benefits from their actions than the external costs, the practical relevance of this case appears to be more limited.

## 6. 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 into account 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 alternative, manipulable toll rules, which are designed to achieve the social optimum when agents anticipate the full impact of their actions on 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. The approach used in the paper can be extended to derive manipulable Pigouvian taxes in other contexts where large agents generate externalities and understand the impact of their actions on corrective tax liabilities.

## 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., 2009. Price vs. quantity-based approaches to airport congestion management. *Journal of Public Economics* 93, 681-690.
- 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.
- CLARKE, E.H., 1971. Multipart pricing of public goods. *Public Choice* 11, 17-33.
- 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., 2001. Distributional consequences of airport congestion pricing. *Journal of Urban Economics* 50, 230-258.
- 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.
- DANIEL, J.I., PAHWA, M., 2000. Comparison of three empirical models of airport congestion pricing. *Journal of Urban Economics* 47, 1-38.
- GROVES, T, LEDYARD, J., 1977. Optimal allocation of public goods: A solution to the free

- rider problem. *Econometrica* 45, 783-809.
- 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.
- KRAUS, M., MOHRING, H., 1975. The role of pollutee taxes in externality problems. *Economica* 42, 171-176.
- 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.T., 2008. Congestion pricing, slot sales and slot trading in aviation. *Transportation Research Part B*, forthcoming.
- VICKREY, W., 1961. Counterspeculation, auctions, and competitive sealed auctions. *Journal of Finance* 16, 8-37.
- ZHANG, A., ZHANG, Y., 2006. Airport capacity and congestion when carriers have market power. *Journal of Urban Economics* 60, 229-247.

## Footnotes

\*We thank Anming Zhang, Ken Small, David Myatt, Will Strange and two referees for helpful comments. Any shortcomings in the paper, however, are our responsibility.

<sup>1</sup>Note that it is therefore not the lumpiness of aircraft but the market power of airlines that makes the atomistic approach inapplicable.

<sup>2</sup>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).

<sup>3</sup>A 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).

<sup>4</sup>Note that the adjective “classical” is used to denote a parametric toll that is set under the correct prediction of the optimum. The term is therefore not used to distinguish tolls with and without partial internalization of congestion.

<sup>5</sup>It should be noted that a classical Pigouvian toll is not subject to the kind of manipulation described in this section. Since a classical toll comes from evaluating the Pigouvian rule at the social optimum, the only way carriers could affect the toll level is by somehow changing the regulator’s perception of the optimum. Such behavior is conceivable, for example, in a setting where the regulator relies on the carriers for information on the cost parameter  $\tau$ , used in computing  $f^*$ . Carriers could misrepresent their costs, thus influencing the regulator’s computed  $f^*$  and ultimately the level of the toll they pay. Exploration of such an alternative setting, however, is beyond the scope of the paper.

<sup>6</sup>Note that the regulator’s expectations are also confirmed in the classical Pigouvian case. The regulator uses the socially flight volume  $f^*$  in computing the toll, and this flight volume is indeed chosen by the carriers.

<sup>7</sup>Evaluated at the optimum, the average manipulable toll equals  $c(2f^*) + K_i/f^*$ .

<sup>8</sup>In (25),  $\beta$  equals right-hand side of (24) evaluated at  $f_1^*$  and  $f_2^*$ , while  $\delta$  equals the right-hand

side of (23) evaluated at the optimum minus  $\beta(f_1^* + f_2^*)$ .

<sup>9</sup>If the firms produce the same good, the common market demand is assumed to be perfectly elastic.