ATOMISTIC CONGESTION TOLLS AT CONCENTRATED AIRPORTS? SEEKING A UNIFIED VIEW IN THE INTERNALIZATION DEBATE

JAN K. BRUECKNER
KURT VAN DENDER

CESIFO WORKING PAPER NO. 2033
CATEGORY 1: PUBLIC FINANCE
JUNE 2007

CESifo *Working Paper No. 2033*

ATOMISTIC CONGESTION TOLLS AT CONCENTRATED AIRPORTS? SEEKING A UNIFIED VIEW IN THE INTERNALIZATION DEBATE

Abstract

The goal of this paper is to bring some unity to the theoretical side of the debate on internalization of airport congestion by showing that all the literature's theoretical results can be derived within one simple and unified framework. The analysis starts by replicating the results of Brueckner (2002), who showed that, because airlines behaving in Cournot fashion internalize congestion, they should be charged low congestion tolls. The analysis then validates the findings of Daniel (1995), who argued that larger atomistic tolls are required in a model where a Stackelberg leader interacts with competitive fringe airlines. However, it is shown that this result only holds approximately when the carriers' outputs are imperfect substitutes.

JEL Code: H23, L50, L9.

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

Kurt Van Dender
Department of Economics
University of California, Irvine
3151 Social Science Plaza
Irvine, CA 92697
USA
kvandend@uci.edu

June 2007

Atomistic Congestion Tolls at Concentrated Airports? Seeking a Unified View in the Internalization Debate

by

Jan K. Brueckner and Kurt Van Dender

1. Introduction

Although the analysis of road congestion pricing has a long history, economists have only recently extended the theory to the case of airports, where congestion has become a major policy issue. The key difference between the two cases is that, while road users are atomistic, airlines are not. In other words, while road users have no incentive to take account of the congestion they impose on other drivers, an airline that schedules an extra flight at a crowded airport congests other airlines but also imposes congestion costs *on the other flights it operates*. Because some congestion is then internalized, airport congestion tolls apparently need not be as large as the atomistic tolls implied by road-pricing theory.

Relying on a simulation model, Daniel (1995) was the first to recognize the potential for internalization of congestion, and Brueckner (2002, 2005) explored the idea further using simple analytical models. Pels and Verhoef (2004), along with more recent contributions by Zhang and Zhang (2006) and Basso and Zhang (2006), provided further elaboration. Comprehensive empirical evidence in support of the internalization hypothesis was offered by Mayer and Sinai (2003), while Brueckner (2002) provided similar evidence using a much smaller data set. Both papers showed that flight delays are lower, other things equal, at highly concentrated airports, where the dominant carrier is likely to internalize much of the congestion it creates, thus limiting its extent.

Daniel (1995), despite identifying the potential for internalization of congestion, argued that an atomistic model, where carriers ignore the congestion impact of their scheduling decisions, is more empirically relevant. He argued that flight cutbacks by a dominant airline aimed at reducing congestion will be offset by the response of fringe carriers, who will schedule more flights so as to leave overall congestion unchanged. As a consequence, the dominant carrier will forego such flight cutbacks, in effect acting atomistically. Daniel (1995) showed that the intraday flight patterns at the Minneapolis-St. Paul airport exhibit too much intertemporal peaking to be consistent with internalization by the dominant carrier (Northwest), and that an atomistic model fits the data better. Daniel and Harback (2007) provided more extensive evidence of this type for a larger number of U.S. airports, while also offering a clearer exposition of model underlying the exercise.[1]

The correct congestion-pricing policy at hub airports depends critically on whether internalization occurs. If Daniel's contrary view is correct, then all carriers at the airport (the hub carrier as well as its competitors) should be charged atomistic congestion tolls. If internalization occurs, however, the hub carrier should be charged a low toll, reflecting internalization of most of the congestion it creates, while the smaller carriers (who internalize less congestion) should pay higher tolls. In the limit, as the hub carrier's flight share approaches 100 percent, congestion tolls effectively disappear, an outcome that can be used to argue against levying any tolls at all at dominated hubs.

Despite the stark contrast between these policy implications, little effort has been spent in attempting to reach a consensus view in the internalization debate. Several obstacles to such an outcome exist. First, the empirical findings in Daniel's work and those of Mayer and Sinai (2003) and Brueckner (2002) are based on very different methodologies. Daniel looks for excessive intraday traffic peaking as evidence of atomistic behavior, while the other work looks at variation in delays across airports, relating it to differences in concentration in an indirect test for internalization.

Consensus on the theoretical side of the debate is also stymied by the use of different types of models under the two approaches. While Daniel's papers are based on a complex stochastic queuing model, Brueckner (2002) and other authors use frameworks that are more elementary and transparent.

Although reconciling the divergent sets of empirical evidence on internalization poses a challenge, the goal of the present paper is to bring some unity to the theoretical side of the internalization debate. This goal is achieved by showing that all the existing theoretical results in the literature, including those of Daniel, can be derived within one simple and unified framework. With theoretical unity achieved, the difference between the opposing views is more easily grasped, perhaps paving the way toward some sort of resolution.[2]

Using a model with two symmetric carriers and some key simplifying assumptions, the analysis starts by characterizing the social optimum. Next, the equilibrium under Cournot behavior is analyzed, and results analogous to those of Brueckner (2002) are derived. In particular, it is shown that congestion tolls only half as large as the atomistic tolls from road-pricing theory are required to support the optimum. Then, the analysis turns to Stackelberg models like those favored by Daniel. In the first model, a Cournot follower interacts with a Stackelberg leader, a case that was briefly analyzed by Brueckner (2002). The analysis demonstrates that, while the follower continues to pay a Cournot congestion toll, the leader should be charged a toll that lies somewhere between the Cournot toll and the larger atomistic toll. In the second model, analogous to the one advanced by Daniel, a Stackelberg leader interacts with competitive fringe carriers. It is shown that, because flight adjustments by the leader are completely offset by the competitive fringe, the leader effectively behaves atomistically, not taking into account the effect of its choices on airport congestion. In this case, atomistic congestion tolls should be charged to both the leader and fringe in order to support the social optimum. The analysis establishing these results is carried out in section 2 of the paper.

While the results for the Stackelberg/fringe model match the conclusions of Daniel, section 3 explores their sensitivity to relaxation of several simplifying assumptions. The analysis first relaxes the assumption that demand is perfectly elastic, which is meant to remove carrier exploitation of market power from the model. It is shown, however, that atomistic tolls are still required when demand is imperfectly elastic. By contrast, relaxing Section 2's implicit assumption that carrier outputs are perfect substitutes partly overturns Daniel's conclusions. With imperfect substitutability between the leader and fringe outputs, the leader partially internalizes congestion, implying that he should be charged a toll that lies somewhere between the Cournot and atomistic tolls. The toll structure is thus not strictly atomistic, but it may nevertheless approximate the one prescribed by Daniel.

In light of these results, section 4 offers some observations on the state of the internalization debate and the implications for public policy.

2. Basic Analysis

2.1. The setup

Following Brueckner (2002), the analysis focuses on a single congested airport, without considering network issues, as in done by Pels and Verhoef (2004) and Brueckner (2005). In addition, the model collapses the peak and offpeak periods from Brueckner's (2002) analysis into a single period where congestion is always present, an approach that follows Pels and Verhoef (2004).

Given these assumptions, suppose that the congested airport is served by two airlines, denoted 1 and 2. The carriers choose their flight volumes to maximize profit, and the resulting equilibrium depends on the nature of interaction between them. In addition, the magnitude of congestion tolls, which are required to eliminate any divergence between the equilibrium and the social optimum, also depends on how the carriers interact.

The analysis considers three possible interaction scenarios. In the first scenario, the carriers behave in Cournot fashion, with each taking the other's flight volume as fixed in making choices. In the second scenario, carrier 2 behaves in Cournot fashion while carrier 1 is a Stackelberg leader, anticipating carrier 2's responses to its choices. In the third scenario, carrier 2 behaves in competitive fashion, viewing the level of congestion at the airport as unaffected by its flight volume, while carrier 1 is again a Stackelberg leader. In this last scenario, carrier 2 can be viewed not as a single airline but as collection of competitive fringe carriers.

In order to maintain the simplest possible focus on the congestion phenomenon, most of the analysis suppresses the market-power component found in many previous models. In these models, a reduction in a carrier's flight volume reduces the level of airport congestion while also raising fares through a standard market-power effect. As a result, airline choices involve both the exploitation of market power and the desire to limit congestion. To focus solely on the congestion issue, market power is eliminated from the basic analysis by assuming that carriers face a perfectly elastic demand for air travel. This assumption is relaxed in section 3 of the paper once the basic results are derived.

Accordingly, it is assumed that passengers are willing to pay a fixed "full price" p for travel in and out of the congested airport.[3] Since passengers dislike airport congestion, which imposes additional time costs, the actual fare that the airlines charge must be discounted below this full price.[4] To derive the discount, let f_1 and f_2 denote flight volumes for the two carriers, and let $t(f_1 + f_2)$ denote the extra time cost per passenger due to congestion and the resulting delays, a cost that depends on total flights at the congested airport. While passenger congestion cost is zero when total flights are low, the t function eventually becomes positive, with $t' > 0,\ t'' \geq 0$ holding over the positive range (so that the marginal congestion cost is constant or rising). Taking account of passenger congestion cost, the airlines are then able to charge a fare equal to $p - t(f_1 + f_2)$. When congestion cost is added to this fare, the resulting full price equals p.

It should be noted that the above framework involves the implicit assumption that the two airlines offer the same product, with their outputs selling at a common price. Stated differently, the outputs of the two carriers are assumed to be perfect substitutes. To judge the sensitivity of the ensuing results to this assumption, the analysis in section 3 relaxes it by allowing imperfect substitutability.

Letting s denoted the fixed seat capacity of an aircraft and assuming that all seats are filled, the total number of seats sold by carrier i is then $sf_i,\ i = 1, 2$. Multiplying by the fare, revenue is

$$[p - t(f_1 + f_2)]sf_i, \quad i = 1, 2. \tag{1}$$

In the absence of congestion, each airline incurs a cost per seat of τ, yielding operating costs of $\tau s f_1$ and $\tau s f_2$ for the two carriers. However, airport congestion raises operating costs, adding an extra cost of $g(f_1 + f_2)$ for each flight. Like $t(\cdot)$, the function g satisfies $g \geq 0$ and $g' > 0,\ g'' \geq 0$ when g is positive (g is zero when $f_1 + f_2$ is small). Total costs for the two airlines are then given by $[\tau s + g(f_1 + f_2)]f_1$ and $[\tau s + g(f_1 + f_2)]f_2$.

Combining the above information, airline profits can be written

$$\pi_i \;=\; [p - t(f_1 + f_2)]sf_i \;-\; [\tau s + g(f_1 + f_2)]f_i, \quad i = 1, 2, \tag{2}$$

and rewritten as

$$\pi_1 \;=\; (p - \tau)sf_1 \;-\; c(f_1 + f_2)f_1 \tag{3}$$

$$\pi_2 \;=\; (p-\tau)sf_2 \;-\; c(f_1+f_2)f_2, \tag{4}$$

where

$$c(f_1+f_2) \;\equiv\; st(f_1+f_2) \;+\; g(f_1+f_2) \tag{5}$$

gives passenger plus airline congestion cost per flight. Given the properties of the t and g functions, $c' > 0$ and $c'' \geq 0$ hold when c is positive.

2.2. Social optimum

The first step in the analysis is to characterize the social optimum. With perfectly elastic demand, consumer surplus is zero, which means that the social optimum simply maximizes the combined profits of the carriers. After adding the profit expressions in (3) and (4), differentiation with respect to f_1 and f_2 yields two identical first-order conditions, given by

$$p \;-\; \tau \;-\; [(f_1+f_2)c' + c]/s \;=\; 0 \tag{6}$$

The resulting solution is symmetric, with the common optimal flight volume for the two carriers found by substituting $f_1 = f_2 = f$ in (6).

Condition (6) says that a carrier's flight volume is optimal when the full price p equals the social marginal cost of a seat, which is given by τ plus the marginal social congestion cost per seat. This latter cost is computed taking into account the congestion cost imposed on both carriers when an extra flight is operated. In particular, when f_1 is increased, passenger plus airline congestion costs for airline 1 (given by cf_1) increase by $c + f_1c'$, while these costs for airline 2 (given by cf_2) increase by f_2c'. The sum of these two effects, divided by s, gives the marginal social congestion cost per seat.

2.3. Cournot behavior

The ensuing analysis characterizes equilibrium outcomes under different types of airline behavior and derives the congestion tolls needed to support the social optimum in each case. Cournot behavior is considered first.

With Cournot behavior, each airline maximizes profit viewing the other airline's flight volume as fixed. Thus, airline 1 chooses f_1 to maximize (3) taking f_2 as given, yielding the first-order condition

$$p \;-\; \tau \;-\; [f_1c' + c]/s \;\;=\;\; 0. \tag{7}$$

Airline 2 satisfies an analogous condition. The difference between this condition and (6) is the absence of f_2c' in the last term. This absence shows that, in scheduling an extra flight, airline 1 takes into account the additional congestion costs imposed on its own flights (f_1c'), ignoring the congestion imposed on airline 2 (f_2c'). Thus, while the airline internalizes some of the congestion from an extra flight, it ignores the impact on the other carrier. Airline 2 behaves in analogous fashion.

With both carriers ignoring a portion of the congestion effects they create, flight volumes are excessive, and a congestion toll is needed to reach the social optimum. The toll per flight is equal to that portion of the congestion damage from an extra flight not internalized by a carrier. The toll is thus equal to f_2c' for carrier 1 and f_1c' for carrier 2, and evaluating at the optimum, the toll is then given by $T = f^*c'(2f^*)$, where f^* is the common optimal flight volume for the two carriers. This toll can also be written as $\frac{1}{2}\text{MCD}^*$, where MCD^* is the marginal congestion damage from an extra flight evaluated at the optimum, equal to marginal congestion cost per flight, $c'(2f^*)$, times the number of flights affected, $2f^*$. Thus, each carrier is charged a toll equal to MCD^* times its airport flight share, which equals 1/2 in the symmetric equilibrium. By contrast, a mechanical application of road pricing rules would incorrectly imply an atomistic congestion toll equal to 100 percent of MCD^*. But this conclusion ignores the fact that each airline internalizes half of the congestion it creates, namely, that imposed on its own flights. With the congestion-toll liabilities of f_1T and f_2T subtracted from carrier profits, the resulting Cournot equilibrium coincides with the social optimum.

Following the approach of Brueckner (2005), this symmetric model can be generalized easily to cover the case of a hub airport, where the hub carrier has a high flight share and the other carrier a small share. Letting $\alpha > 1/2$ denote hub carrier's share, it must pay a relatively low toll of $(1-\alpha)\text{MCD}^*$, equal to the (small) portion of marginal congestion damage not internalized, while the nonhub airline pays a larger toll of αMCD^*, reflecting less internalization.

2.4. Stackelberg behavior with a Cournot follower

Now suppose that airline 1 is a Stackelberg leader, with 2 the follower. In Cournot fashion, carrier 2 chooses f_2 viewing f_1 as parametric, satisfying the condition

$$p \;-\; \tau \;-\; [f_2c' + c]/s \;\;=\;\; 0. \tag{8}$$

To find the response of f_2 to a change in f_1, needed in analyzing the leader's behavior, the condition (8) is totally differentiated, yielding

$$\frac{\partial f_2}{\partial f_1} \;\;=\;\; -\frac{f_2c'' + c'}{f_2c'' + 2c'} \;\;\equiv\;\; -\lambda \;\;<\;\; 0. \tag{9}$$

The negative sign of (9), which follows from the maintained assumptions on the c function, indicates that the airline 2 reduces its flights in response to an increase in airline 1's flights. Note that the λ expression in (9) is greater than 1/2 but less than unity, with the lower bound corresponding to the limiting case where $c'' \equiv 0$. Thus, airline 2 offsets somewhere between half and 100 percent of any increase in carrier 1's flights through a decrease in its own flights.

Knowing the response of carrier 2 to its f_1 choice, carrier 1 maximizes profit from (3), and the first-order condition is

$$p \;-\; \tau \;-\; \frac{1}{s}\left[f_1c'\left(1 + \frac{\partial f_2}{\partial f_1}\right) \;+\; c\right] \;\;=\;\; 0. \tag{10}$$

With $-1 < \partial f_2/\partial f_1 < -1/2$, the term involving c' in (10) is cut by more than half relative to its value in (7), indicating a reduction in carrier 1's internalization of congestion. The airline anticipates that a reduction in f_1 meant to reduce congestion will be mostly offset by an increase in f_2 by the follower airline. Therefore, carrier 1's incentive to limit its flights to restrain congestion is reduced relative to the Cournot case. Indeed, comparison of (9) and (10) shows that, for both conditions to be satisfied, $f_1 > f_2$ must hold, indicating that carrier 1 operates more flights than the follower.

Congestion tolls are once again needed to generate the optimum, but these tolls are now different across the carriers. Given its Cournot behavior, carrier 2's toll, denoted T_2, is the same as before, given by $\frac{1}{2}$MCD*. Carrier 1's toll, however, must eliminate the difference between the $f_1 c'(1 + \partial f_2/\partial f_1)$ term in (10) and the $(f_1 + f_2)c'$ term in the optimality condition (6), with this difference being evaluated at the optimum. Subtracting the first expression from the second, the toll per flight should equal

$$T_1 \;=\; \left(f_2 \;-\; f_1 \frac{\partial f_2}{\partial f_1}\right) c' \;=\; \frac{(f_2 + \lambda f_1)MCD}{f_1 + f_2} \tag{11}$$

evaluated at the optimum, where MCD $= (f_1 + f_2)c'$. Recalling that the optimum has symmetric flight volumes of f^*, the toll in (11) evaluated at the optimum reduces to

$$T_1 \;=\; \frac{1}{2}(1 + \lambda^*)\text{MCD}^* \tag{12}$$

where the asterisk indicates evaluation at the optimum. Thus, airline 1 pays a toll larger than MCD* times its flight share, reflecting its failure to internalize as much congestion as in the Cournot case.

A subtlety in (13) is that λ, the $-\partial f_2/\partial f_1$ derivative from (9), must be evaluated at $f_1 = f_2 = f^*$ (yielding λ^*) to compute carrier 1's toll. However, in the special case where the c function is linear, with $c'' = 0$, (9) shows that λ is a constant equal to 1/2. In this case, T_1 reduces simply to $\frac{3}{4}$MCD*. Thus, the appropriate toll for the Stackelberg leader lies exactly halfway between the Cournot toll of $\frac{1}{2}$MCD* and the atomistic toll of MCD*. While this simple statement does not apply when $c'' > 0$, the fact that $1/2 < \lambda^* < 1$ means the leader's toll nevertheless lies somewhere between the Cournot toll and the atomistic toll, as seen from inspection of (12).

Summarizing the preceding analysis yields

Proposition 1. *With a Stackelberg leader and a Cournot follower, the follower is charged a Cournot congestion toll, equal to* 1/2 *of the marginal congestion damage* *from an extra flight. The leader is charged a toll that lies between this Cournot value and the atomistic toll* (*which equals* 100 *percent of the marginal congestion damage from an extra flight*).

Thus, when the follower acts in Cournot fashion, the leader's Stackelberg behavior causes the toll structure to move toward an atomistic structure, without fully reaching it.

2.5. Stackelberg behavior with a competitive fringe

Now suppose that airline 1 continues to behave in Stackelberg fashion but that airline 2 behaves competitively. In the present context, competitive behavior means that carrier 2 ignores the effect of its own flight choice on airport congestion, thus viewing the value of the c function as parametric. The plausibility of such behavior can be heightened by viewing carrier 2 as collection of small competitive airlines rather than as a single entity, with f_2 giving the combined flights for this group. Note that the overall scale of operations of these fringe airlines need not itself be small. However, the airlines must have small flight shares *at the given airport*, justifying their parametric view of congestion. Thus, a fringe carrier could be a large airline that operates only a few flights into another carrier's hub.

Viewing c as parametric, the profit function (4) for carrier 2 is proportional to f_2, with the proportionality factor equal to $p - \tau - c/s$. If this expression is nonzero, then the airline will prefer either a zero or infinite flight volume, an outcome inconsistent with the achievement of equilibrium. To avoid this outcome, making carrier 2 indifferent to its flight volume, the equality

$$p \;-\; \tau \;-\; c/s \;\;=\;\; 0 \qquad (13)$$

must hold, a condition that implies zero profit for carrier 2. But observe that, even though c is viewed as parametric by carrier 2, c's actual value depends on the magnitude of $f_1 + f_2$. Therefore, writing c as $c(f_1 + f_2)$, condition (13) ends up determining an equilibrium value for f_2 conditional on f_1. In effect, for a given f_1 value, f_2 expands until $1/s$ times congestion cost per flight reaches $p - \tau$, yielding zero profit for airline 2.

The response of f_2 to an increase in f_1 can be derived immediately given that (13) implies a constant value for $f_1 + f_2$. The result is

$$\frac{\partial f_2}{\partial f_1} = -1, \tag{14}$$

indicating that an increase in f_1 is fully offset by a decrease in f_2.

To derive the leader's flight choice, (14) is substituted into the first-order condition for choice of f_1 (eq. (10)). Upon substitution, the term involving c' becomes zero, and (10) reduces to (13). Thus, the equilibrium conditions for airlines 1 and 2 are identical, and since they determine only the sum $f_1 + f_2$, the individual flight levels are indeterminate.[5]

By assumption, carrier 2 behaves atomistically, ignoring the impact of its flight choice on airport congestion. The key implication of the above analysis is that carrier 1, in effect, behaves atomistically as well, even though it plays the role of a Stackelberg leader. In particular, carrier 1's first-order condition, which mirrors the equilibrium condition for carrier 2, shows no accounting for the effect of its flight choice on airport congestion. This failure to internalize congestion arises because any change in f_1 is completely offset by adjustment of f_2, so that airport congestion is in fact insensitive to carrier 1's choice.

The absence of internalization implies that the equilibrium flight volumes are too large, and that the excess is greater than in the Cournot case from above, requiring a larger toll. Since the $(f_1 + f_2)c'/s$ term in the optimality condition (6) is absent in the equilibrium condition (13), an atomistic toll, given by $T = \mathrm{MCD}^*$, is required. Thus, behavior that is either atomistic by assumption (carrier 2) or effectively atomistic (carrier 1) necessitates the use of atomistic tolls. Summarizing yields

> **Proposition 2.** *With a Stackelberg leader and a competitive fringe, no congestion is internalized and all carriers should be charged atomistic congestion tolls, equal to* 100 *percent of the marginal congestion damage from an extra flight.*

It is important to note that Proposition 2 also applies in a model with multiple leader airlines, who interact in Cournot fashion with respect to one another while anticipating the response of competitive fringe carriers. This case may describe a number of congested airports where no carrier is dominant, such as New York-LaGuardia or Boston.

3. Do Weaker Assumptions Resurrect Internalization?

While the previous analysis shows that atomistic congestion tolls are required when a Stackelberg leader interacts with a competitive fringe, confirming Daniel's view, it is important to know whether this result continues to hold under weaker assumptions. To answer this question, the analysis in this section begins by relaxing the assumption that demand is perfectly elastic, allowing the exercise of market power by the leader. Then, while maintaining imperfectly elastic demand, the analysis allows the outputs of the leader and the fringe to be imperfect rather than perfect substitutes.

3.1. Imperfectly elastic demand

When demand is imperfectly elastic, the full price p depends on total traffic $s(f_1+f_2)$, with the demand function written $p = d[s(f_1+f_2)]$, where $d' < 0$. To understand the impact of this modification in the Stackelberg/fringe setting, it is helpful to first reconsider the Cournot model. In that model, carrier 1's first-order condition (7) is replaced by

$$d \;+\; sf_1d' \;-\; \tau \;-\; (f_1c' + c)/s \;\;=\;\; 0. \tag{15}$$

Since the social optimum is given by $d - \tau - [(f_1+f_2)c' + c]/s = 0$, where d replaces p in (6), two distortions are evident in (15).[6] Uninternalized congestion, reflected in the absence of the f_2c'/s term, tends to make f_1 too large, while exercise of market power, reflected in the presence of the negative term sf_1d', tends to make f_1 too small. Whether f in the symmetric equilibrium is then too large or too small relative to the social optimum is unclear.

Although a toll (or subsidy) per flight can remedy the net distortion in this simple setup, a more realistic network model generates a different prescription. In such a model, the congested airport serves a host of city-pair markets that may exhibit different degrees of competition and thus market-power distortions of varying severity. In this setting, achievement of the social optimum requires Cournot congestion tolls to be levied at the airport level while subsidies, designed to correct the market-power distortion, are paid at the level of the city-pair market. Thus, uninternalized congestion and exploitation of market power are remedied by different instruments. Brueckner (2005) provides a full analysis of such a network model.

Returning to the current non-network setting, consider again the case where a Stackelberg leader interacts with a competitive fringe. The equilibrium condition (13) for the fringe is replaced by

$$d \;-\; \tau \;-\; c/s \;\;=\;\; 0, \tag{16}$$

which once again determines f_2 conditional on f_1. Since both d and c depend on $f_1 + f_2$, (16) requires a constant value for this sum, again implying that a change in f_1 is fully offset by adjustment of f_2, as in (14).

The leader's first-order condition (10) is replaced by

$$d \;+\; sf_1 d' \left(1 + \frac{\partial f_2}{\partial f_1}\right) \;-\; \tau \;-\; \frac{1}{s}\left[f_1 c' \left(1 + \frac{\partial f_2}{\partial f_1}\right) \;+\; c\right] \;\;=\;\; 0. \tag{17}$$

After substituting (14), this condition reduces to (16), just as in the previous analysis. The leader again fails to internalize congestion, but in addition, exploitation of market power (seen in (15)) does not occur. In effect, the leader ends up behaving just like a competitive carrier when disciplined by the competitive fringe: it ignores congestion effects and does not exploit its market power. Only one distortion, uninternalized congestion, then needs correction, and this correction is achieved through atomistic tolls. Therefore, Proposition 2 remains relevant when demand is imperfectly elastic.

3.2. Imperfect substitutes: The case of full independence

To introduce imperfect substitutability, it is useful to first consider the polar case where the leader and the fringe offer independent goods, with cross-price elasiticities equal to zero. In this case, the full prices are given by $p_1 = d_1(sf_1)$ and $p_2 = d_2(sf_2)$, with $d_1', d_2' < 0$ (demands continue to be imperfectly elastic).[7]

The fringe equilibrium condition then becomes

$$d_2 \;-\; \tau \;-\; c/s \;\;=\;\; 0, \tag{18}$$

and since the argument of d_2 is now sf_2, not $s(f_1 + f_2)$, the fringe response to a change in f_1

no longer involves fully offsetting behavior. Differentiating (18) yields

$$\frac{\partial f_2}{\partial f_1} \;=\; -\,\frac{c'/s}{c'/s \;-\; sd_2'} \;\equiv\; -\mu. \tag{19}$$

Since $d_2' < 0$, it follows that μ lies between zero and 1, indicating that the fringe response does not fully offset a change in f_1.

The implication for congestion tolls can again be seen in the leader's first-order condition, which is now given by

$$d_1 \;+\; sf_1d_1' \;-\; \tau \;-\; \frac{1}{s}\left[f_1c'\left(1+\frac{\partial f_2}{\partial f_1}\right) \;+\; c\right] \;=\; 0. \tag{20}$$

Since $\partial f_2/\partial f_1$ lies between -1 and 0, (20) reflects partial internalization of congestion rather than its complete absence. To derive the corresponding toll, recall that, as in the Stackelberg/Cournot case, the toll must eliminate the difference between the first term in brackets in (20) and the $(f_1 + f_2)c'$ term in the optimality condition (6). The required toll is then given by

$$T_1 \;=\; \frac{1}{2}(1+\mu^*)\text{MCD}^* \tag{21}$$

where μ^* equals μ evaluated at the social optimum (compare (12)). Thus, because $0 < \mu^* < 1$, the leader's toll lies between the Cournot toll of $\frac{1}{2}$MCD* and the atomistic toll of MCD*, just as in the Stackelberg/Cournot case with perfectly elastic demand. The fringe, however, continues to pay an atomistic toll.

Note that the market-power term from the Cournot model (see (15)) remains in (20). The reason is that leader's full price is independent of f_2, which means that the offsetting behavior of the fringe does not restrain the exploitation of market power, in contrast to (17). The market-power distortion thus requires a separate correction, as discussed above.

3.3. Imperfect substitutes: The general case

While the previous results overturn Proposition 2, a more comprehensive statement requires consideration of the general case, where substitutability between the leader and fringe outputs lies between the extremes of independence and perfect substitutes. As might be expected, internalization of congestion in this intermediate case still occurs but is weaker than when the carriers' outputs are independent. Thus, the appropriate congestion toll for the leader moves closer to the atomistic toll.

To generate the inverse demand functions for the intermediate case, consider the primitive demand functions $sf_1 = D_1(p_1, p_2)$ and $sf_2 = D_2(p_1, p_2)$, which give traffic for the leader and the fringe as functions of the full prices. The own-price derivatives for these functions are assumed to be negative, while the cross-price derivatives are positive. Inverting the demand functions yields the inverse functions $p_1 = d_1(sf_1, sf_2)$ and $p_2 = d_2(sf_1, sf_2)$, whose properties depend on the properties of D_1 and D_2. Specifically,

$$d_1^1 = \frac{sD_2^2}{\Omega}, \qquad d_1^2 = -\frac{sD_1^2}{\Omega} \tag{22}$$

$$d_2^1 = -\frac{sD_2^1}{\Omega}, \qquad d_2^2 = \frac{sD_1^1}{\Omega} \tag{23}$$

where the superscripts indicate partial derivatives and where $\Omega = D_1^1 D_2^2 - D_2^1 D_1^2$. A natural assumption is that an increase in either full price reduces total traffic, so that $\partial s(f_1 + f_2)/\partial p_1 = D_1^1 + D_2^1 < 0$ and $\partial s(f_1 + f_2)/\partial p_2 = D_1^2 + D_2^2 < 0$. These inequalities imply $\Omega > 0$ and thus that all of the inverse-demand derivatives in (22) and (23) are negative. In addition, the inequalities yield

$$\left|d_1^1\right| > \left|d_1^2\right|, \qquad \left|d_2^2\right| > \left|d_2^1\right|, \tag{24}$$

indicating that a carrier's full price falls by more when its own flight volume increases than when volume for the other carrier rises.

To find the response of the fringe to an increase in f_1 in this new setting, the equilibrium condition (18) is differentiated, recognizing that d_2 now depends on both f_1 and f_2. The result is

$$\frac{\partial f_2}{\partial f_1} = -\frac{c'/s - sd_2^1}{c'/s - sd_2^2} \equiv -\eta. \tag{25}$$

Using (24), it follows that η in (25) satisfies $0 < \eta < 1$. Note also that (25) reduces to (19) in the independent case, where $d_2^1 = 0$, and that it equals -1 in the perfect substitutes case, where $d_2^1 = d_2^2$.

Holding the own-price effect constant, so that the magnitude of d_2^2 equals that of d_2' in the independent case, the presence of the extra term in the numerator means that η in (25) is closer to 1 than is μ in (19). As a result, relative to the independent case, the fringe more fully offsets an increase in f_1 with intermediate substitutability. Thus, less internalization occurs than in the independent case, which implies that the leader's toll moves closer to the atomistic toll. Repeating the previous steps, the toll in the intermediate substitutability case equals

$$T_1 \quad = \quad \frac{1}{2}(1 + \eta^*)\text{MCD}^*, \tag{26}$$

where the asterisk on η again indicates evaluation at the optimum. Since $\eta^* > \mu^*$ is satisfied holding the own-price price effect constant, the toll in (26) is larger than the one in (21).[8]

Summarizing all of the above results yields

> **Proposition 3.** *When a Stackelberg leader interacts with a competitive fringe, atomistic congestion tolls are still required when demand is imperfectly inelastic, as long as the carriers' outputs are perfect substitutes. But when substitutability is imperfect, congestion is partially internalized, and the leader's congestion toll lies between the Cournot toll and the atomistic toll.*

Judging the practical implications of Proposition 3 requires knowledge of the extent of substitutability among the outputs offered by different airlines. Given that most empirical work on the industry implicitly assumes that substitutability is perfect, evidence on this issue is scant. From one perspective, it could be argued that airline seats have a "commodity" status, with little fundamental product differentiation evident. While this view implies that substitutability should be close to perfect, airline frequent-flier programs constitute a powerful tool for inducing brand loyalty, thus limiting the extent of substitutability.[9] The upshot is that the airline outputs may not be perfect substitutes, but the extent of substitutability is probably large compared to other industries. If this view is correct, then the toll structure implied by Proposition 3 may be close to atomistic.

4. Discussion

A key focus of the internalization debate is on dominated hub airports, and proper congestion pricing for such airports depends critically on which of the above models is relevant, as noted above. If the Stackelberg model with a competitive fringe is relevant, then the hub carrier as well as the other airlines using the airport should be charged congestion tolls that are atomistic, or approximately so. On the other hand, if the Cournot model is relevant, then the hub carrier should be charged a low toll that approaches zero as its flight share nears 100 percent, a conclusion that could be used to argue against levying any tolls at all dominated hubs. Thus, the two models are diametrically opposed in their recommendations for congestion pricing at hub airports.

A similar, but less dramatic, contrast arises for other airports, such as New York-LaGuardia or Boston, that are congested but have no dominant carrier. At such airports, the Cournot model prescribes levying congestion tolls that are nonatomistic but still appreciable. For example, the Cournot toll would equal $\frac{2}{3}\text{MCD}^*$ at an airport with 3 symmetric carriers. But as mentioned above, as long as a competitive fringe operates at such an airport, with the major carriers anticipating its response, atomistic tolls equal to MCD^* (or tolls near this value) are again required, a 50 percent increase over the Cournot toll in the given example.

Where does the discussion in this paper leave us in appraising the internalization debate? The present analysis shows that Daniel's theoretical argument against the internalization hypothesis can be valid in a simple model like that used in the rest of the literature. Whether the implication of this argument (the need for atomistic, or approximately atomistic, tolls at every congested airport) should be embraced depends partly on a judgment regarding the realism of the competing models. In other words, is it realistic to assume that a congested airport's smallest carriers are small enough to view overall congestion as independent of their decisions but large enough for their responses to neutralize internalization by the airport's major carriers? The answer is unclear. Empirical evidence on internalization exists, of course, but it offers conflicting messages, as explained earlier. While the demonstrated empirical link between airport delays and concentration favors the internalization hypothesis, this evidence is indirect. By contrast, Daniel's contradictory evidence based on intraday flight patterns is suggestive but not very transparent.

In the face of these ambiguities, a possible stance in the internalization debate is suggested by the results of a recent paper by Morrison and Winston (2006). These authors show that, even if internalization is assumed to occur, the welfare loss from levying inappropriate atomistic congestion tolls is relatively small. These findings could provide support for a congestion-pricing policy that ignores potential internalization of congestion. This approach is the right one if Daniel is correct, and although it is wrong otherwise, the resulting welfare loss may be tolerable.

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.

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., 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., Harback, K.T., 2007. (When) Do hub airlines internalize their self-imposed congestion delays? *Journal of Urban Economics*, forthcoming.

Forbes, S.J., 2006. The effect of air traffic delays on airline prices. Unpublished paper, University of California, San Diego.

Mayer, C., Sinai, T., 2003. Network effects, congestion externalities, and air traffic delays: Or why all delays are not evil. *American Economic Review* 93, 1194-1215.

Morrison, S.A., Winston, C., 2006. Another look at airport congestion pricing. *American Economic Review*, forthcoming.

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

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

Footnotes

[*] We thank Joe Daniel and Leonardo Basso for helpful comments.

[1] See also Daniel (2001) and Daniel and Pahwa (2000).

[2] Reconciling the literature's empirical results is challenging because, even though the Daniel and Harback's (2007) results do not uniformly reject internalization, which is evident at a few of the 27 airports they study, Mayer and Sinai's (2003) findings show a significant (though modest) effect of airport-concentration on delays in a sample covering more than 250 airports. Their model uses airport fixed effects, along with a massive amount of data on delays for individual flights.

[3] The other endpoints connected to the airport are assumed to be uncongested. Note that in a more detailed model, p would vary across endpoints.

[4] See Forbes (2006) for empirical evidence that airport congestion reduces fares.

[5] Since both carriers effectively behave in competitive fashion, this indeterminacy is natural. Note also that, with (13) holding, profit for carrier 1 is zero.

[6] The social optimum now maximizes consumer surplus plus profit.

[7] When imperfect substitutability is added to the perfectly-elastic-demand case of section 2.5, full prices are fixed but potentially different across carriers. It is easy see, however, that $p_1 = p_2$ must hold for both the leader and the fringe to operate in equilibrium.

[8] It should be noted that imperfect substitutability has virtually no effect in the Cournot case. The only required modification is the substitution of d_1 and d_1^1 in place of d and d' in (15).

[9] Service quality differences, such as the availability of a direct flight when competitors only offer connecting flights, may also limit substitutability.

CESifo Working Paper Series

(for full list see www.cesifo-group.de)

1968 Michela Redoano, Does Centralization Affect the Number and Size of Lobbies?, April 2007

1969 Christian Gollier, Intergenerational Risk-Sharing and Risk-Taking of a Pension Fund, April 2007

1970 Swapan K. Bhattacharya and Biswa N. Bhattacharyay, Gains and Losses of India-China Trade Cooperation – a Gravity Model Impact Analysis, April 2007

1971 Gerhard Illing, Financial Stability and Monetary Policy – A Framework, April 2007

1972 Rainald Borck and Matthias Wrede, Commuting Subsidies with two Transport Modes, April 2007

1973 Frederick van der Ploeg, Prudent Budgetary Policy: Political Economy of Precautionary Taxation, April 2007

1974 Ben J. Heijdra and Ward E. Romp, Retirement, Pensions, and Ageing, April 2007

1975 Scott Alan Carson, Health during Industrialization: Evidence from the 19th Century Pennsylvania State Prison System, April 2007

1976 Andreas Haufler and Ian Wooton, Competition for Firms in an Oligopolistic Industry: Do Firms or Countries Have to Pay?, April 2007

1977 Eckhard Janeba, Exports, Unemployment and the Welfare State, April 2007

1978 Gernot Doppelhofer and Melvyn Weeks, Jointness of Growth Determinants, April 2007

1979 Edith Sand and Assaf Razin, The Role of Immigration in Sustaining the Social Security System: A Political Economy Approach, April 2007

1980 Marco Pagano and Giovanni Immordino, Optimal Regulation of Auditing, May 2007

1981 Ludger Woessmann, Fundamental Determinants of School Efficiency and Equity: German States as a Microcosm for OECD Countries, May 2007

1982 Bas Jacobs, Real Options and Human Capital Investment, May 2007

1983 Steinar Holden and Fredrik Wulfsberg, Are Real Wages Rigid Downwards?, May 2007

1984 Cheng Hsiao, M. Hashem Pesaran and Andreas Pick, Diagnostic Tests of Cross Section Independence for Nonlinear Panel Data Models, May 2007

1985 Luis Otávio Façanha and Marcelo Resende, Hierarchical Structure in Brazilian Industrial Firms: An Econometric Study, May 2007

1986 Ondřej Schneider, The EU Budget Dispute – A Blessing in Disguise?, May2007

1987 Sascha O. Becker and Ludger Woessmann, Was Weber Wrong? A Human Capital Theory of Protestant Economic History, May 2007

1988 Erkki Koskela and Rune Stenbacka, Equilibrium Unemployment with Outsourcing and Wage Solidarity under Labour Market Imperfections, May 2007

1989 Guglielmo Maria Caporale, Juncal Cunado and Luis A. Gil-Alana, Deterministic versus Stochastic Seasonal Fractional Integration and Structural Breaks, May 2007

1990 Cláudia Costa Storti and Paul De Grauwe, Globalization and the Price Decline of Illicit Drugs, May 2007

1991 Thomas Eichner and Ruediger Pethig, Pricing the Ecosystem and Taxing Ecosystem Services: A General Equilibrium Approach, May 2007

1992 Wladimir Raymond, Pierre Mohnen, Franz Palm and Sybrand Schim van der Loeff, The Behavior of the Maximum Likelihood Estimator of Dynamic Panel Data Sample Selection Models, May 2007

1993 Fahad Khalil, Jacques Lawarrée and Sungho Yun, Bribery vs. Extortion: Allowing the Lesser of two Evils, May 2007

1994 Thorvaldur Gylfason, The International Economics of Natural Resources and Growth, May 2007

1995 Catherine Roux and Thomas von Ungern-Sternberg, Leniency Programs in a Multimarket Setting: Amnesty Plus and Penalty Plus, May 2007

1996 J. Atsu Amegashie, Bazoumana Ouattara and Eric Strobl, Moral Hazard and the Composition of Transfers: Theory with an Application to Foreign Aid, May 2007

1997 Wolfgang Buchholz and Wolfgang Peters, Equal Sacrifice and Fair Burden Sharing in a Public Goods Economy, May 2007

1998 Robert S. Chirinko and Debdulal Mallick, The Fisher/Cobb-Douglas Paradox, Factor Shares, and Cointegration, May 2007

1999 Petra M. Geraats, Political Pressures and Monetary Mystique, May 2007

2000 Hartmut Egger and Udo Kreickemeier, Firm Heterogeneity and the Labour Market Effects of Trade Liberalisation, May 2007

2001 Andreas Freytag and Friedrich Schneider, Monetary Commitment, Institutional Constraints and Inflation: Empirical Evidence for OECD Countries since the 1970s, May 2007

2002 Niclas Berggren, Henrik Jordahl and Panu Poutvaara, The Looks of a Winner: Beauty, Gender, and Electoral Success, May 2007

2003 Tomer Blumkin, Yoram Margalioth and Efraim Sadka, Incorporating Affirmative Action into the Welfare State, May 2007

2004 Harrie A. A. Verbon, Migrating Football Players, Transfer Fees and Migration Controls, May 2007

2005 Helmuth Cremer, Jean-Marie Lozachmeur and Pierre Pestieau, Income Taxation of Couples and the Tax Unit Choice, May 2007

2006 Michele Moretto and Paolo M. Panteghini, Preemption, Start-Up Decisions and the Firms' Capital Structure, May 2007

2007 Andreas Schäfer and Thomas M. Steger, Macroeconomic Consequences of Distributional Conflicts, May 2007

2008 Mikael Priks, Judiciaries in Corrupt Societies, June 2007

2009 Steinar Holden and Fredrik Wulfsberg, Downward Nominal Wage Rigidity in the OECD, June 2007

2010 Emmanuel Dhyne, Catherine Fuss, Hashem Pesaran and Patrick Sevestre, Lumpy Price Adjustments: A Microeconometric Analysis, June 2007

2011 Paul Belleflamme and Eric Toulemonde, Negative Intra-Group Externalities in Two-Sided Markets, June 2007

2012 Carlos Alós-Ferrer, Georg Kirchsteiger and Markus Walzl, On the Evolution of Market Institutions: The Platform Design Paradox, June 2007

2013 Axel Dreher and Martin Gassebner, Greasing the Wheels of Entrepreneurship? The Impact of Regulations and Corruption on Firm Entry, June 2007

2014 Dominique Demougin and Claude Fluet, Rules of Proof, Courts, and Incentives, June 2007

2015 Stefan Lachenmaier and Horst Rottmann, Effects of Innovation on Employment: A Dynamic Panel Analysis, June 2007

2016 Torsten Persson and Guido Tabellini, The Growth Effect of Democracy: Is it Heterogenous and how can it be Estimated?, June 2007

2017 Lorenz Blume, Jens Müller, Stefan Voigt and Carsten Wolf, The Economic Effects of Constitutions: Replicating – and Extending – Persson and Tabellini, June 2007

2018 Hartmut Egger and Gabriel Felbermayr, Endogenous Skill Formation and the Source Country Effects of International Labor Market Integration, June 2007

2019 Bruno Frey, Overprotected Politicians, June 2007

2020 Jan Thomas Martini, Rainer Niemann and Dirk Simons, Transfer Pricing or Formula Apportionment? Tax-Induced Distortions of Multinationals' Investment and Production Decisions, June 2007

2021 Andreas Bühn, Alexander Karmann and Friedrich Schneider, Size and Development of the Shadow Economy and of Do-it-yourself Activities in Germany, June 2007

2022 Michael Rauscher and Edward B. Barbier, Biodiversity and Geography, June 2007

2023 Gunther Schnabl, Exchange Rate Volatility and Growth in Emerging Europe and East Asia, June 2007

2024 Erkki Koskela and Ronnie Schöb, Tax Progression under Collective Wage Bargaining and Individual Effort Determination, June 2007

2025 Jay Pil Choi and Marcel Thum, The Economics of Politically Connected Firms, June 2007

2026 Jukka Pirttilä and Roope Uusitalo, Leaky Bucket in the Real World: Estimating Inequality Aversion Using Survey Data, June 2007

2027 Ruslan Lukach, Peter M. Kort and Joseph Plasmans, Strategic R&D with Knowledge Spillovers and Endogenous Time to Complete, June 2007

2028 Jarko Fidrmuc, Neil Foster and Johann Scharler, Labour Market Rigidities, Financial Integration and International Risk Sharing in the OECD, June 2007

2029 Bernardina Algieri and Thierry Bracke, Patterns of Current Account Adjustment – Insights from Past Experience, June 2007

2030 Robert Dur and Hein Roelfsema, Social Exchange and Common Agency in Organizations, June 2007

2031 Alexander Libman and Lars P. Feld, Strategic Tax Collection and Fiscal Decentralisation: The Case of Russia, June 2007

2032 Øystein Foros, Hans Jarle Kind and Greg Shaffer, Resale Price Maintenance and Restrictions on Dominant Firm and Industry-Wide Adoption, June 2007

2033 Jan K. Brueckner and Kurt Van Dender, Atomistic Congestion Tolls at Concentrated Airports? Seeking a Unified View in the Internalization Debate, June 2007