Congestion Pricing, Slot Sales and Slot Trading in Aviation

Erik T. Verhoef

VU University Amsterdam, and Tinbergen Institute.

CONGESTION PRICING, SLOT SALES AND SLOT TRADING IN AVIATION

Erik T. Verhoef[*]
Department of Spatial Economics
VU University Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
Phone: +31-20-5986094
Fax: +31-20-5986004
Email: everhoef@feweb.vu.nl

Abstract

This paper studies the regulation of an airline duopoly on a congested airport. Regulation should then address two market failures: uninternalized congestion, and overpricing due to market power. We find that first-best charges are differentiated over airlines if asymmetric, and completely drive out the least efficient airline from the market. This is not generally the case for an undifferentiated charge, which is found to be a weighted average of first-best charge rules for the two airlines, and is less-than-optimally efficient because of its inability to differentiate between them. Tradeable slots may yield the first-best outcome if the congestion externality is relatively important and the market power distortion relatively unimportant, but may be less efficient than non-intervention when the reverse is true.

[*]Affiliated to the Tinbergen Institute, Roetersstraat 31, 1018 WB Amsterdam.

1.   Introduction

Congestion is a growing problem in aviation, and is attracting growing attention from transport economists. Daniel (1995), for example, applied Vickrey's (1969) bottleneck model of traffic congestion to the case of congested air traffic to study optimal pricing. Indeed, transport analysts have looked for inspiration in the road pricing literature for modelling congestion at airports, but important conceptual differences exist between the economics of congestion in these two modes. These are due to the fact that, in contrast to individual road users, there are typically at least one or some airlines with substantial market power active at an airport.

Bruecker (2002) pointed out that a conventional "Pigouvian" congestion charge, equal to the marginal external cost imposed upon all other flights at the airport, is excessive for an airline with market power, because the airline already internalizes the congestion that an additional flight would impose upon its own other services. The first-best congestion charge therefore requires a downward adjustment of the Pigouvian rule, multiplying it for airline i with a factor $(1-s_i)$, with s_i denoting the airline's market share. This implies that airlines with a higher market share should pay lower charges; probably not always an easy message from the political perspective.

Pels and Verhoef (2004) argued that if airlines with market power compete as Cournot oligopolists, a further downward adjustment in airport charges is in order, above that proposed by Brueckner, to reduce the oligopolists' incentive to apply a price mark-up – which decreases with demand elasticity – in setting ticket fares. They show that the second-best charge may then well be negative if demand is sufficiently inelastic, which follows Buchanan's (1969) findings for optimal environmental taxation of a monopolist. Basso and Zhang (2006) and Zhang and Zhang (2006) included the role of airports in modelling congestion between airlines at airports.

This paper looks into some alternative economic instruments for managing congestion at airports, notably slot sales and slot trading – which are frequently put forward as possibly promising instruments. Economic theory suggests that at least in idealized situations, tradable permits and Pigouvian charges should produce the same efficient outcome (*e.g.*, Dales, 1968; Montgomery, 1972). Verhoef, Nijkamp and Rietveld (1997) argued that tradeable permits may offer an interesting option for managing road transport externalities, although practical applicable may of course be hampered by the large number of potential traders present in a road network. At first sight, such problems seem less likely for aviation, where only a limited, manageable number of trading airlines would be active at most airports.

The paper contrasts two slot-based policies with 'classic' charges. The first is slot sales, where the regulator or airport (we assume a public, social welfare maximizing regulator controlling the airport) sells slots at a uniform price. The main shortcoming of this policy is that slot prices will be uniform over airlines, while efficiency may require the price for slots to be differentiated over airlines. This shortcoming would of course be relevant for any regulatory scheme that would lead to uniform slot prices. In fact, we will model the outcomes of this policy by treating it as a uniform charge; a policy that will not be first-best efficient

with asymmetric airlines for the reason mentioned above, but that may certainly be politically more acceptable. The second slot-based policy involves the case where slots are tradeable between airlines. Here we will see that the non-negativity of tradeable slot prices may cause another distortion, above that from uniformity of prices, compared to charges – which can be set below zero (*i.e.*, as a subsidy) when the market power effect dominates the congestion effect.

The plan of the paper is as follows. We will use a simple analytical model for the analysis, describing a congested airport with two Cournot duopolists. That model is presented in Section 2 below. Section 3 analyzes first-best charges; Section 4 uniform charges or slot sales, and Section 5 slot trading. Finally, Section 6 concludes.

2.   A simple model of Cournot duopolists

Consider two airlines serving a single market (*i.e.* a single origin-destination pair). The airport of interest is congested and publicly operated: the airport authority chooses social surplus as its objective. It is the sum of charge revenues, airline profits and consumer surplus for all consumers – which could reflect a situation where both the origin and destination are located in the same country and the airport of interest is operated by the national government. We thus ignore issues of tax competition that would arise if the airport authority would consider not all consumers' surpluses as relevant for its objective (Pels and Verhoef, 2004, consider this). Social surplus can also be defined as total consumer benefits minus the total cost of the airlines' operations. The other airport is not congested, has zero marginal cost, and is assumed not to charge an airport fee; so, we can ignore it altogether for welfare analyses (equivalently, we could assume that the cost functions defined below capture cost for both airports together, which are both run by the same public authority).

It is assumed that all demand and cost functions of interest are linear. This simplifies the analytics and facilitates the derivation of closed-form equilibrium expression, without compromising the results too much – at least when one is not primarily interested in second-order effects or the effects of non-linearities. The inverse demand for (return) trips is given by:

$$D = \delta^0 - \delta^1 \cdot \left(n_A + n_B\right) \tag{1}$$

where n_A and n_B reflect the number of return trips served by airlines *A* and *B*, respectively. (All parameters in the inverse demand and cost functions are non-negative.)

We assume that the number of flights offered is proportional to the number of passengers served. Absent congestion, the marginal cost (per passenger) would be constant and equal to γ^0. But congestion raises the cost per flight incurred by an airline. We assume that these congestion costs are symmetric in the sense that an extra flight by the competing airline raises an airline's average cost by the same amount as does an own flight. We allow for asymmetry between airlines by distinguishing airline-specific parameters in the resulting average cost functions, which – due to the assumption of constant passenger load factors – can be expressed directly in numbers of passengers:

$$c_i = \gamma_i^0 + \gamma_i^1 \cdot \left(n_A + n_B\right) \quad i = \{A, B\} \tag{2}$$

In order to limit the number of potential cases to be considered, we restrict attention to cases where either both airlines have identical cost parameters, or one airline (to be denoted A without loss of generality) is more cost-efficient at all output levels:

$$\text{sign}\{\gamma_B^0 - \gamma_A^0\} = \text{sign}\{\gamma_B^1 - \gamma_A^1\} \geq 0 \tag{3}$$

Congestion could be assumed to enter also passengers' generalized cost functions, but doing so does not affect the analysis in any fundamental way (see also Pels and Verhoef, 2004). We also refrain from formalizing the counteracting force of 'frequency benefits' that passengers might enjoy from higher aggregate output levels.

The airport of interest ("the airport", in what follows) may levy an airport charge τ. The most general case allows for differentiation of charges τ_i between airlines, in which case subscripts again denote airlines. As for cost functions, we express this charge as a charge per passenger, which is again allowed because of the assumption of exogenous load factors. The airport behaves like a Stackelberg leader: airlines take charges as given while the airport sets these optimally taking into account how airlines will respond.

As customary in the aviation literature, we assume that the duopolists engage in Cournot competition.[1] Furthermore, we will limit our attention to Nash games between the two airlines, with airline i taking the other airline's output, n_{-i}, as given when setting its own output n_i. Passengers consider the two airlines as pure substitutes, implying that a single ticket price prevails in equilibrium.

Under these assumptions, airline i's profit function can be written as:

$$\Pi_i = n_i \cdot \left(\delta^0 - \delta^1 \cdot (n_A + n_B)\right) - n_i \cdot \left(\gamma_i^0 + \gamma_i^1 \cdot (n_A + n_B) + \tau_i\right) \quad i = \{A, B\} \tag{4}$$

The first-order necessary condition for profit maximization is easily derived:

$$\frac{\partial \Pi_i}{\partial n_i} = \left(\delta^0 - \gamma_i^0 - \tau_i\right) - 2 \cdot \left(\delta^1 + \gamma_i^1\right) \cdot n_i - \left(\delta^1 + \gamma_i^1\right) \cdot n_{-i} = 0 \quad i = \{A, B\} \tag{5}$$

The Nash equilibrium outputs can therefore be written as:

$$n_i^N = \frac{\delta^0 - \gamma_i^0 - \tau_i}{2 \cdot \left(\delta^1 + \gamma_i^1\right)} - \tfrac{1}{2} \cdot n_{-i} \quad i = \{A, B\} \tag{6a}$$

or, in closed form, as:

$$n_i^N = \frac{2}{3} \cdot \frac{\delta^0 - \gamma_i^0 - \tau_i}{\delta^1 + \gamma_i^1} - \frac{1}{3} \cdot \frac{\delta^0 - \gamma_{-i}^0 - \tau_{-i}}{\delta^1 + \gamma_{-i}^1} \quad i = \{A, B\} \tag{6b}$$

Note that an airline's equilibrium output, not surprisingly, rises with δ^0 and falls with δ^1, and furthermore decreases with the airline's own cost parameters and increases with those of the competitor. Furthermore, for an interior equilibrium with positive outputs, the second term on the right-hand side of (6b) should not exceed, in absolute value, the first term, which implies that there are certain boundaries on the feasible parameter space.

---

[1] In an empirical analysis of Chicago-based airline routes, Oum *et al.* (1993) conclude that "the overall results indicate that the duopolists' conduct may be described as somewhere between Bertrand and Cournot behavior, but much closer to Cournot, in the majority of the sample observations". Also Brander and Zhang (1990), using similar data, find " strong evidence … against the highly competitive Bertrand hypothesis", and find Cournot behavior plausible for the markets under consideration.

Finally, in a fully symmetric setting, with equal cost parameters and charges, (6b) can be further simplified to:

$$n_i^N = \frac{1}{3} \cdot \frac{\delta^0 - \gamma^0 - \tau}{\delta^1 + \gamma^1} \quad i = \{A, B\} \tag{6c}$$

3.   First-best charges

First-best pricing rules, set so as to maximize social surplus, can be found after specifying social welfare as total benefits (determined as the relevant area below the inverse demand function) minus total costs:

$$\Omega = \delta^0 \cdot (n_A + n_B) - \tfrac{1}{2} \cdot \delta^1 \cdot (n_A + n_B)^2 - n_A \cdot (\gamma_A^0 + \gamma_A^1 \cdot (n_A + n_B)) - n_B \cdot (\gamma_B^0 + \gamma_B^1 \cdot (n_A + n_B)) \tag{7}$$

and determining the first-order necessary conditions:

$$\frac{\partial \Omega}{\partial n_i} = (\delta^0 - \gamma_i^0) - (\delta^1 + 2 \cdot \gamma_i^1) \cdot n_i - (\delta^1 + \gamma_i^1 + \gamma_{-i}^1) \cdot n_{-i} = 0 \quad i = \{A, B\} \tag{8}$$

These, in turn, imply the following optimal outputs:

$$n_i^* = \frac{\delta^0 - \gamma_i^0}{\delta^1 + 2 \cdot \gamma_i^1} - \frac{\delta^1 + \gamma_i^1 + \gamma_{-i}^1}{\delta^1 + 2 \cdot \gamma_i^1} \cdot n_{-i} \quad i = \{A, B\} \tag{9a}$$

As will become clear below, and will be intuitively clear in a set-up with homogeneous outputs, under asymmetric cost parameters only one airline will remain active in the optimum. The following closed-form optimum outputs then apply (anticipating that the more cost-efficient airline *A* will be the one remaining active):

$$\begin{cases} n_A^* = \dfrac{\delta^0 - \gamma_A^0}{\delta^1 + 2 \cdot \gamma_A^1} \\ n_B^* = 0 \end{cases} \tag{9b}$$

When the cost parameters are identical across airlines, optimum outputs as implied by (9a) are in fact only determined uniquely up to the aggregate level:

$$n_A^* + n_B^* = \frac{\delta^0 - \gamma^0}{\delta^1 + 2 \cdot \gamma^1} \tag{9c}$$

Due to the neutral scale economies (for a given level of aggregate output and hence congestion), the shares of both airlines in the aggregate optimal output level are immaterial for efficiency. When symmetry in outputs prevails, we would of course find:

$$n_i^* = \tfrac{1}{2} \cdot \frac{\delta^0 - \gamma^0}{\delta^1 + 2 \cdot \gamma^1} \quad i = \{A, B\} \tag{9d}$$

However, also unequal optimal firm-specific output levels can be supported by (unequal) first-best taxes when cost parameters are identical. To see why this is the case, first observe that the optimal taxes supporting the first-best equilibrium can be found by substituting (5) into (8), which yields:

$$\tau_i = \gamma_{-i}^1 \cdot n_{-i} - \delta^1 \cdot n_i \quad i = \{A, B\} \tag{10}$$

This tax rule consists of two components. The first, positive term is equal to the marginal external congestion costs imposed upon the other airline. Congestion costs imposed upon own

aircraft is internalized by the airline, and requires no corrective taxes (which would, in fact, be distortionary), as in Brueckner (2002). The second, negative term corrects for overpricing in Cournot equilibria. This term increases in the slope of the inverse demand function, reflecting that a steeper function induces a greater degree of overpricing, as in Pels and Verhoef (2004).[2] The optimal tax will be negative when the market power effect outweighs the congestion effect. Should subsidization be considered unwarranted for political reasons, then a zero tax is preferable on efficiency grounds to a positive congestion tax in such cases.

The optimal taxes in (10) are equal across identical airlines only in a fully symmetric optimum, *i.e.* when not only their cost parameters are the same but also their equilibrium outputs are equal. Equation (6c) confirms that with equal taxes and equal cost parameters, equilibrium outputs will indeed be equal. But output levels may also differ, despite the equal cost parameters, in the first-best optimum as implied by (9c). Equation (10) shows that the optimal charge an airline faces is increasing in the other airline's size (reflecting greater congestion externalities imposed) and decreasing in the airline's own size (reflecting a smaller degree of market power). But is (9c) exactly satisfied with unequal tolls set according to (10) for two airlines with identical cost parameters? This answer appears to be "yes": the implied different optimal tax levels would then still produce an optimal aggregate output as given by (9c).[3] As a consequence, there is no unique decentralized optimum in terms of individual airlines' outputs when the airlines have identical cost parameters.

This has the interesting implication that the regulator has scope for giving the one identical airline preferential treatment over the other, by differentiating charges, without creating losses in social surplus. In the context of our model, there would be no inherent reason to do so, and the fully symmetric, equal-charges outcome would be the outcome when the regulator's secondary objective would – quite realistically – involve fairness between competing airlines. It is worth noting that the non-uniqueness of firm-specific first-best taxes and outputs in this model is caused by the assumption of neutral scale economies at the firm level (for a given congestion level). Economies of scale at the firm level would cause the optimum to have only a single supplier even if the airlines have equal cost structures; diseconomies of scale at the firm level would require the two identical airlines to split the optimal output in equal portions.

---

[2] This term vanishes under perfectly elastic demand (with $\delta^1=0$). Brueckner and Van Dender (2008), who want to consider the congestion externality as the only source of market failure, consequently assume demand to be perfectly elastic – which can certainly be justified on grounds of analytical tractability. Nevertheless, market power in real aviation markets appears important enough to warrant explicit treatment in the context of congestion regulation, and it fortunately appears that the simple set-up in this paper allows such treatment in an analytical model without compromising tractability too much.

[3] This can be verified by substituting, for the case of equal cost parameters, (10) into (6b) and summing over both airlines. This yields, after some manipulation:

$$n_A + n_B = \frac{2}{3} \cdot \frac{\delta^0 - \gamma^0}{\delta^1 + \gamma^1} + \frac{1}{3} \cdot \frac{\left(\delta^1 - \gamma^1\right)}{\delta^1 + \gamma^1} \cdot \left(n_A + n_B\right)$$

which in turn can be solved to yield (9c) (the steps involve multiplication of both sides by 3, and a multiplication of the left-hand side by $(\delta^1+\gamma^1)/(\delta^1+\gamma^1)$. After taking the right-most term to the left-hand side, the result is straightforward to derive).

Let us now turn to asymmetry in cost parameters, with airline *A* more cost effective at all output levels as specified by equation (3). Under our assumptions, this will always result in optima with only one airline present in the market. This can be verified by noting that the two first-order conditions (8) can only be fulfilled simultaneously, as required for an interior solution with both airlines active, when the following equality holds:

$$
\gamma_A^0 + \left(\gamma_A^1 - \gamma_B^1\right) \cdot n_A = \gamma_B^0 + \left(\gamma_B^1 - \gamma_A^1\right) \cdot n_B
$$
$$
\Rightarrow n_B = \left(\gamma_A^0 - \gamma_B^0\right) + \frac{\gamma_A^1 - \gamma_B^1}{\gamma_B^1 - \gamma_A^1} \cdot n_A
\tag{11}
$$

The assumption made in (3) prevents this from being possible for positive supply levels under asymmetric cost parameters: n_B and n_A would have opposite signs. An interior optimum therefore does not exist, and the optimum is the corner solution in which only the most efficient airline (*A*, without loss of generality) will remain active – as might have been anticipated.[4] This means that under asymmetry in cost parameters, the optimal taxes in (10) will become:

$$
\begin{cases}
\tau_A = -\delta^1 \cdot n_A \\
\tau_B = \gamma_A^1 \cdot n_A
\end{cases}
\tag{12}
$$

The subsidy for the 'insider' (airline *A*) corrects for monopolistic pricing, while the positive tax for the 'outsider' (airline *B*), by internalizing the congestion costs that a first unit of output sold would create, effectively prevents unwarranted entry by an inefficient competitor.[5]

If the airline cost function, apart from the congestion costs, would be assumed to reflect decreasing rather than constant economies of scale, interior optima with both airlines active would be possible also under asymmetry. But airline cost functions may in fact be more

---

[4] Pels and Verhoef (2004) find a similar result in a numerical model (their Section 4.2), but do not provide an analytical proof.

[5] This can be verified by noting that airline *A* will set its first-order condition for profit maximization (5) equal to zero, with τ_A from (12) and $n_B = 0$ substituted:

$$
\delta^0 - \gamma_A^0 - \left(\delta^1 + 2 \cdot \gamma_A^1\right) \cdot n_A = 0
$$

Airline *B*'s first-order condition for profit maximization (5), with τ_B from (12) substituted, can be written as:

$$
\delta^0 - \gamma_B^0 - \left(\delta^1 + \gamma_A^1 + \gamma_B^1\right) \cdot n_A - 2 \cdot \left(\delta^1 + \gamma_B^1\right) \cdot n_B = 0
$$

After subtraction of airline *A*'s first-order condition we find:

$$
\left(\gamma_A^0 - \gamma_B^0\right) + \left(\gamma_A^1 - \gamma_B^1\right) \cdot n_A - 2 \cdot \left(\delta^1 + \gamma_B^1\right) \cdot n_B = 0
$$
$$
\Rightarrow n_B = \frac{\left(\gamma_A^0 - \gamma_B^0\right) + \left(\gamma_A^1 - \gamma_B^1\right) \cdot n_A}{2 \cdot \left(\delta^1 + \gamma_B^1\right)}
$$

It is then easily verified that, because $\gamma_B^0 > \gamma_A^0$ and $\gamma_B^1 > \gamma_A^1$ by construction, this condition can only be fulfilled for a negative n_B when $n_A > 0$, and that airline *B*'s profits are decreasing (or actually: losses are increasing) in n_B as long as it is non-negative. Therefore, τ_B in (12) effectively prevents entrance of the inefficient firm *B*.

likely to exhibit increasing rather than decreasing economies of scale,[6] in which case the result that one firm will be priced off the market with first-best taxes would be reinforced. The survival of only one airline in the first-best optimum is, therefore, certainly not just a peculiarity due to an overly simplified analytical representation. At the same time, it should be acknowledged that the model does not allow for various inefficiencies from market power other than through pricing behaviour, such as through X-inefficiency and rent-sharing. These might of course cause smaller, competing airlines to become still more cost-effective than their larger counterparts.

4. Uniform taxes and slot sales

A second-best alternative to the first-best charges just described would be a regime of uniform charges τ, undifferentiated by airline. This instrument could be chosen for reasons of (perceived) fairness, or because political constraints make differentiated tolls undesirable, or perhaps for yet different reasons. The instrument is interesting to study in itself, but also because the resulting equilibrium corresponds to the outcome of a system of 'slot sales', where the airport regulator sells landing rights at a uniform price τ.

We solve for the second-best optimal uniform charge level by solving the Lagrangian that has social surplus as the objective and the airlines' first-order conditions for profit maximization, from equation (5), as the constraints:

$$\Lambda = \delta^0 \cdot (n_A + n_B) - \tfrac{1}{2} \cdot \delta^1 \cdot (n_A + n_B)^2 - n_A \cdot (\gamma_A^0 + \gamma_A^1 \cdot (n_A + n_B)) - n_B \cdot (\gamma_B^0 + \gamma_B^1 \cdot (n_A + n_B))$$
$$+ \lambda_A \cdot [(\delta^0 - \gamma_i^0 - \tau) - 2 \cdot (\delta^1 + \gamma_A^1) \cdot n_A - (\delta^1 + \gamma_A^1) \cdot n_B] \tag{13}$$
$$+ \lambda_B \cdot [(\delta^0 - \gamma_i^0 - \tau) - (\delta^1 + \gamma_B^1) \cdot n_A - 2 \cdot (\delta^1 + \gamma_B^1) \cdot n_B]$$

The necessary first-order conditions are:

$$\frac{\partial \Lambda}{\partial n_i} = \delta^0 - \delta^1 \cdot (n_A + n_B) - \gamma_i^0 - 2 \cdot \gamma_i^1 \cdot n_i - (\gamma_i^1 + \gamma_{-i}^1) \cdot n_{-i}$$
$$- \lambda_i \cdot 2 \cdot (\delta^1 + \gamma_i^1) - \lambda_{-i} \cdot (\delta^1 + \gamma_{-i}^1) = 0 \quad i = \{A, B\} \tag{14a}$$

$$\frac{\partial \Lambda}{\partial \tau} = -\lambda_A - \lambda_B = 0 \tag{14b}$$

$$\frac{\partial \Lambda}{\partial \lambda_i} = (\delta^0 - \gamma_i^0 - \tau) - 2 \cdot (\delta^1 + \gamma_i^1) \cdot n_i - (\delta^1 + \gamma_i^1) \cdot n_{-i} = 0 \quad i = \{A, B\} \tag{14c}$$

Substitution of (14c) into (14a) yields, after using (14b), the following expressions for the Lagrangian multipliers:

$$\lambda_i = \frac{\tau + \delta^1 \cdot n_i - \gamma_{-i}^1 \cdot n_{-i}}{\delta^1 + 2 \cdot \gamma_i^1 - \gamma_{-i}^1} \quad i = \{A, B\} \tag{14d}$$

Summing over i and solving for τ then yields the following second-best optimal uniform charge:

---

[6] An important reason for economies of scale might be that larger aircraft can be used for larger passenger flows at lower average cost. Swan and Adler (2006) for example report scale elasticities of cost for direct flights with respect to seats of around 0.6 for a Cobb-Douglas cost function (their Section 4.1).

$$\tau = \frac{\dfrac{\gamma_B^1 \cdot n_B - \delta^1 \cdot n_A}{\delta^1 + 2 \cdot \gamma_A^1 - \gamma_B^1} + \dfrac{\gamma_A^1 \cdot n_A - \delta^1 \cdot n_B}{\delta^1 + 2 \cdot \gamma_B^1 - \gamma_A^1}}{\dfrac{1}{\delta^1 + 2 \cdot \gamma_A^1 - \gamma_B^1} + \dfrac{1}{\delta^1 + 2 \cdot \gamma_B^1 - \gamma_A^1}} \tag{15}$$

The charge is a weighted average of the first-best charge rules of (10) – but, of course, evaluated in the second-best optimum so that the quantities n_i will be different. Airline A's charge rule appears in the numerator of the first fraction in the expression's main numerator, and that for B in the second. An airline's weight increases when its equilibrium quantity offered is more sensitive to marginal toll changes than its competitor's quantity. More specifically, an airline's weight in the uniform charge of (15) amounts to:

$$w_i = \frac{1}{\delta^1 + 2 \cdot \gamma_i^1 - \gamma_{-i}^1} \tag{16}$$

The weights in (16) are in fact perfectly proportional with the airline's equilibrium demand sensitivities with respect to τ, defined as the derivative of (6b) with respect to a uniform toll $\tau = \tau_i = \tau_{-i}$:

$$-\frac{\partial n_i^N}{\partial \tau} = \frac{2}{3} \cdot \frac{1}{\delta^1 + \gamma_i^1} - \frac{1}{3} \cdot \frac{1}{\delta^1 + \gamma_{-i}^1} \quad i = \{A, B\} \tag{17a}$$

This implies the following ratio of the firms' equilibrium demand sensitivities:

$$\frac{\partial n_i^N / \partial \tau}{\partial n_{-i}^N / \partial \tau} = \frac{\dfrac{2}{\delta^1 + \gamma_i^1} - \dfrac{1}{\delta^1 + \gamma_{-i}^1}}{\dfrac{2}{\delta^1 + \gamma_{-i}^1} - \dfrac{1}{\delta^1 + \gamma_i^1}} = \frac{2 \cdot (\delta^1 + \gamma_{-i}^1) - (\delta^1 + \gamma_i^1)}{2 \cdot (\delta^1 + \gamma_i^1) - (\delta^1 + \gamma_{-i}^1)} = \frac{\delta^1 + 2 \cdot \gamma_{-i}^1 - \gamma_i^1}{\delta^1 + 2 \cdot \gamma_i^1 - \gamma_{-i}^1} \quad i = \{A, B\} \tag{17b}$$

The acclaimed perfect proportionality of weights w_i from (16) with demand responsiveness from (17a) is immediately visible from (17b). This proportionality reflects the notion, well known from the optimal taxation literature (*e.g.*, Sandmo, 1975, 1976), that deviations from first-best optimal prices are relatively more harmful for economic efficiency when equilibrium quantities are more sensitive to price changes.

The charge rule in (15) bears great similarity with the second-best undifferentiated congestion toll that Verhoef, Nijkamp and Rietveld (1995; equation (5)) derive for road traffic with group specific marginal congestion effects (*e.g.*, with small cars and large cars), although the pricing rules are of course not identical between the two models. Specifically, the numerator's numerators for the atomistic car drivers in Verhoef *et al.* (1995) differ from the corresponding terms in (15) above in that they have the traditional 'full' marginal external cost term (because there is no internalization of self-imposed congestion), and in that they lack the demand-related term (because individual road users have no market power and do not engage in monopolistic pricing). The weight for the second-best uniform road toll in Verhoef *et al.* (1995) has the (absolute value of the) slope of the inverse demand function, as in (15), and adds to this the slope of the average user cost function. These two terms together determine by how much a group's equilibrium demand changes after a marginal toll in the uniform road toll – so the function of the weights is exactly the same as in the current context, although the analytical expression is different.

As in Verhoef *et al*. (1995), and in other studies on second-best pricing reviewed by Small and Verhoef (2007), non-zero equilibrium value of the Lagrangian multipliers λ_i reflect that social surplus could be improved further when moving from a constrained second-best equilibrium towards the first-best outcome. Note that these multipliers would be zero if differentiated charges were possible, or if the first-best charge rules happen to produce the same preferred charge level in the second-best optimum, or if the airline's equilibrium demand happens to be completely insensitive with respect to marginal charge changes. These are the situations where the lack of charge differentiation has no consequences for the social surplus that can be achieved with second-best charges, as compared to first-best pricing.

The uniform charge τ in (15) converges to first-best undifferentiated charges under symmetry, and that the multipliers λ_i will then both approach zero. Indeed, it is to be expected that the relative efficiency of this uniform charge increases when the degree of asymmetry reduces. This is of course the case if the cost parameters γ for the two firms are closer, and if (as a natural consequence) the equilibrium outputs n are closer.

As an aside, note that the weight in (16) will certainly be positive for airline *B*, but may appear to possibly take on a negative value for airline *A*, which would lead to difficulties in interpretation. However, a negative weight for airline *A* could only occur outside the feasible parameter space that can be identified on the basis of equation (6b) – that is, it corresponds with a negative equilibrium value for airline *B*'s output – so that we can ignore this possible complication.[7]

When there is asymmetry, the common charge will not generally drive out the less efficient airline – unlike first-best differentiated charges do. And furthermore, because of the weighted-average structure of the second-best charge, the more efficient airlines faces a charge that is above, or a subsidy that is below, the charge or subsidy implied by the first-best charge rule evaluated in the second-best optimum. The reverse is true for the less efficient airline. It should be emphasized that the comparison between equilibrium levels of first-best and second-best charges or subsidies is not directly possible on the basis of tax rules (12) and (15) because also equilibrium passenger flows n_i will differ between the first-best and second-best optimum. But one may expect the more efficient airline to face an above-optimal charge or to receive a below-optimal subsidy, while the opposite is true for the less efficient airline.

In conclusion, firm-specific outputs will generally differ between first-best and second-best equilibria, and social surplus for the latter will therefore typically be lower. For the aggregate output, we may expect a usually larger volume under first-best pricing than under second-best pricing, because marginal cost is higher in the latter case.

This conclusion was confirmed by a closely related analysis by Brueckner (2008), who applied the model of Brueckner and Van Dender (2008) to study the efficiency of uniform slot prices. That model ignores the market power distortion by assuming perfectly elastic demand,

---

[7] A necessary condition for the output for the less efficient airline *B* in (6b) to be positive under a uniform tax is $\delta^1 + \gamma_A^1 > \frac{1}{2} \cdot \left(\delta^1 + \gamma_B^1 \right)$. Whenever $\gamma_A^0 < \gamma_B^0$, an even stricter inequality applies (*i.e.*, the factor $\frac{1}{2}$ on the right-hand side should be increased), so the inequality it is a necessary, not a sufficient condition. This necessary condition can be rewritten as $\delta^1 + 2 \cdot \gamma_A^1 > \gamma_B^1$. Under the latter inequality, w_A in (16) cannot be negative.

which requires airline cost functions to exhibit decreasing scale economies for an interior equilibrium to exist. The two airlines use the same airport, as in this paper, but serve different markets so that congestion is the only type of interaction between them. Again, first-best congestion charges decrease in a firm's market share of total airport operations, and a uniform slot price therefore cannot produce the first-best outcome.

5.   Slot trading

We have treated slot sales as equivalent to uniform charging in the above. This seems justified at least as long as slot trading between airlines is prohibited. But slot trading is of course a third possible instrument, which has gained increasing interest over the past years. An important similarity between slot trading and slot sales or uniform charges as discussed above is that the same shadow price for slots applies for the two airlines, at least if there is a marginal slot traded between them. One important difference, however, between uniform charges and slot trading is that the former instrument allows for a negative shadow price – *i.e.*, a subsidy. Indeed, in absence of congestion, both the first-best tax for the active airline *A* in (12) and the uniform charge in (15) will become negative. For slot trading, a negative market price for the associated permits appears impossible, at least when the owner of a slot has no 'service obligation' as we will call it; that is, the owner of the permit is not obliged to actually use it. This means that in cases where the market power distortion outweighs the congestion distortion and optimal second-best charges would be negative, slot trading would become a useless mechanism: the number of permits that the regulator would like to see traded and especially used simply exceeds the airlines' aggregate demand for permits as it would be at a zero market price for permits.

The consideration of slot trading is therefore only meaningful when the congestion distortion outweighs the market power distortion, and that is the assumption we will be making in this section. Specifically, we assume that the initial equilibrium is the Nash equilibrium from equation (6) with taxes equal to zero, and that the regulator finds it attractive to reduce aggregate equilibrium outputs to a level *N* (to be optimized at a later stage) for which:

$$N < n_A^N + n_B^N \tag{18}$$

There are basically two ways of introducing slot trading into the market. The first would be to sell slots at a predetermined price, and allowing slot trading afterwards. The second would be to distribute the permits between the airlines for free ('grandfathering'), in which case it appears natural that both airlines should receive fewer permits initially (N_i^{ini}), than the initial outputs:

$$N_A^{ini} < n_A^N$$
$$N_B^{ini} < n_B^N \tag{19}$$

We consider the case where permits are grandfathered in a way that satisfies (19). Both airlines would then in principle have a positive marginal willingness to pay for slots. This can be seen by consulting the 'pre-slot-payment' profit functions:

$$\Pi_i = n_i \cdot \left(\delta^0 - \delta^1 \cdot (n_A + n_B) \right) - n_i \cdot \left(\gamma_i^0 + \gamma_i^1 \cdot (n_A + n_B) \right) \quad i = \{A, B\} \tag{20}$$

Two derivatives are of interest now. The first is with respect to the own output, keeping the other airline's output fixed. In line with equation (5), we find:

$$\frac{\partial \Pi_i}{\partial n_i} = \left(\delta^0 - \gamma_i^0\right) - 2 \cdot \left(\delta^1 + \gamma_i^1\right) \cdot n_i - \left(\delta^1 + \gamma_i^1\right) \cdot n_{-i} \quad i = \{A, B\} \tag{21a}$$

which implies that the 'desired output' $n_i^{\#}$, for a given n_{-i} and disregarding a purchase price for slots, amounts to:

$$n_i^{\#} = \frac{\left(\delta^0 - \gamma_i^0\right)}{2 \cdot \left(\delta^1 + \gamma_i^1\right)} - \tfrac{1}{2} \cdot n_{-i} \quad i = \{A, B\} \tag{21b}$$

Both (21a) and (21b) imply that given an initial endowment of permits according to (19), both airlines would use all slots. When doing so, the derivative in (21a) is still positive, implying a desire to expand output further, which is confirmed by comparing (21b) with the Nash output in (6a) for $n_{-i} < n_{-i}^N$. This implies that:

$$n_i^{\#} > n_i^N > N_i^{ini} \quad \text{if} \quad n_{-i} = N_i^{ini} < n_{-i}^N \quad i = \{A, B\} \tag{21c}$$

Both airlines would therefore like to expand output in principle, but whether this will lead to slot trading depends on the question of whether the one airline is willing to pay more for an additional slot or a set of slots than the other airline requires for giving it up. We first look at the question of whether there is an incentive to trade a single, marginal slot, in the case where the slot was initially used by the selling airline and will be used by the buying airline.

The answer depends on the second type of derivative of (20), namely the one that considers simultaneous but opposite changes in n_i and n_{-i}, so that $n_A + n_B$ remains constant (the case where the buying airline decides not to fully use the slot purchased is considered below). Using this derivative rather than the one in (21a) acknowledges that the trading airlines realize that the maximum total number of passengers is given and determined by *N*. The airlines therefore do not, naïvely, expect the other airline's output to be constant after buying or selling a used slot. We denote this as a derivative with respect to $T_{-i,i}$, where *–i* is the selling airline and *i* the buying airline, and where *T* stands for slots traded. Suppose airline *A* buys and *B* sells. The derivatives of the profit functions with respect to $T_{-i,i}$, for a negotiated permit price *p*, amount to:

$$\begin{cases} \left.\dfrac{\partial \Pi_A}{\partial T_{B,A}}\right|_{dn_A = -dn_B} = \left(\delta^0 - \delta^1 \cdot \left(n_A + n_B\right)\right) - \left(\gamma_A^0 + \gamma_A^1 \cdot \left(n_A + n_B\right)\right) - p \\[4mm] \left.\dfrac{\partial \Pi_B}{\partial T_{B,A}}\right|_{dn_A = -dn_B} = -\left(\delta^0 - \delta^1 \cdot \left(n_A + n_B\right)\right) + \left(\gamma_B^0 + \gamma_B^1 \cdot \left(n_A + n_B\right)\right) + p \end{cases} \tag{22}$$

If there exists a *p* for which both expressions become positive, one can expect a mutually beneficial trade to occur. Such a *p* always exists when airline *A* is more efficient than airline *B*. Slot trading can then be expected to continue until the output of airline *B* is completely eliminated (which is why our choice of considering a sale from *B* to *A* is consistent with the assumption that if the airline differs, *A* is more efficient). The result is that airline *A* obtains a monopoly position. And this, of course, raises the question of whether the assumption that

airline A fully utilizes every slot purchased is warranted. The answer depends on the number of slots available, N, relative to the profit-maximizing monopoly output for A, n_A^M. The latter follows as a special case of (21b):

$$n_A^M = \frac{\left(\delta^0 - \gamma_A^0\right)}{2 \cdot \left(\delta^1 + \gamma_A^1\right)} \tag{23}$$

If $N \le n_A^M$, the derivative of airline A's profits with respect to its output is still positive even when owning all slots, and full utilization is the (constrained) profit maximizing strategy. If $N > n_A^M$, there will be a certain level of permits traded beyond which airline A does not find it profitable to fully utilize each additional permit purchased from B. According to (21b), this occurs as soon as n_B has become so low that $n_A^\# < N - n_B$. More specifically, (21b) shows that airline A, in these circumstances, prefers to use half of each additional slot purchased from B. Does this possibly eliminate the existence of a mutually attractive permit price p? No, it does not. Equation (22) shows that both airlines can always benefit from an additional permit traded if airline A is more efficient and fully utilizes the next slot. If it decides, for profit considerations, not to do so, it will profit even more than when utilizing it completely, so the same p that would support a trade under full utilization for A would also support half utilization. But also airline B can only be better off when making the trade and airline A does not fully utilize the additional slot. The ticket price for B's remaining output increases, and the congestion suffered decreases compared to full utilization by A. Therefore, A's decision not to fully utilize the slot can never cause B to refuse.[8]

As a first set of conclusions, we find that if airlines differ, a system of slot trading will cause the inefficient airline to be eliminated from the market. If airline A has no service obligation, it will offer the minimum of N and n_A^M in the resulting equilibrium. If it does have a service obligation, it will offer N. But a service obligation for a monopolistic airline appears a rather theoretical construct, especially if the monopolist is free to charge any ticket price it prefers.

Given these results, we can now determine the optimal number of permits N, if satisfying (19): it can take on any value between n_A^M and $n_A^N + n_B^N$. Any value in this range will lead to the end result that the aggregate output will become equal to n_A^M. A lower output is socially less desirable, since n_A^M in (23) is generally below the socially optimal output n_A^* in (9b) when only firm A is active (assuming finite demand elasticity). N should therefore not be below n_A^M.

Tradeable slots thus produce the profit-maximizing monopolistic equilibrium with only firm A active. A positive consequence is that the less efficient airline B is no longer active, as in the first-best equilibrium. A negative consequence is that the lack of competition

---

[8] We ignored one possibility in the reasoning above, and that is the situation where after some trading from airline B to A, B's output is still positive but below the number of slots it then holds. If such a situation would occur, A could simply bid a price p for the amount of B's unused slots plus 1, and otherwise the same reasoning applies as in the main text: whether or not A will in fact fully uses all slots purchased, it can make a mutually profitable bid even if constrained to fully replace the single unit of output that B gives up.

induces monopolistic price setting by airline A. It is not possible to decide whether the overall welfare effects of these two consequences is always positive. That is, it may be desirable not to have tradeable permits in place and having the free-market Nash outcome $\{ n_A^N, n_B^N \}$ rather than the monopolistic outcome n_A^M. But the reverse may also be true.

For example, if demand is perfectly elastic so that δ^1 equals zero, the monopolistic output n_A^M in (23) coincides with the socially optimal output n_A^* in (9b); both being equal to $\frac{1}{2} \cdot (\delta^0 - \gamma_A^0) / \gamma_A^1$. The issuing of n_A^M tradeable slots then leads to a first-best welfare level. The free-market Nash aggregate output implied by (6c), in contrast, would, if the airlines become nearly identical, approximate $\frac{2}{3} \cdot (\delta^0 - \gamma_A^0) / \gamma_A^1$, which is socially excessive because the two airlines do not perfectly internalize the congestion externality.

However, if congestion is absent so that γ^1 is zero for both airlines, we find as the socially optimal output in (9b) $n_A^* = (\delta^0 - \gamma_A^0) / \delta^1$; for the monopolistic output in (23) $n_A^M = \frac{1}{2} \cdot (\delta^0 - \gamma_A^0) / \delta^1$; and an aggregate duopoly output implied by (6c) that would, if the airlines become nearly identical, approximate $\frac{2}{3} \cdot (\delta^0 - \gamma_A^0) / \delta^1$. The duopoly output is now closer to the socially optimal output, because overpricing due to market power has reduced.

Therefore, depending on whether the congestion externality or the market power distortion dominates, tradeable slots and the associated monopolistic outcome may or may not be more attractive than the free-market Nash equilibrium.

Finally, for perfectly symmetric airlines, equation (22) might be taken to suggest that no slot trading will take place. But it is important to recall that airlines need not trade permits one by one. If N is not below the monopoly output n^M (which now is equal for both airlines), freely tradable slots in fact give the duopolists the incentive and possibility to collude and maximize joint profits. By definition, the airline that out-buys the other airline will enjoy profits from monopolist operation that exceed the initial sum of duopoly profits, and can therefore make a mutually attractive bid for the full set of the other airline's slots. When N is smaller than n^M, the incentive for such collusion is in fact eliminated, and it is immaterial for total industry profits and for social surplus which of the identical airlines serves how many passengers.

6. Conclusion

Our simple model of airline duopoly competition on a congested airport produced the following conclusions.

First, consistent with Brueckner (2002) we found that first-best charges should be differentiated between asymmetric airlines. Following Pels and Verhoef (2004), the first-best charge comprises two elements: the marginal external cost imposed upon the competitor, and a demand-related subsidy to counter overpricing due to market power. We found that, with neutral or positive scale economies at the firm level (with the degree of scale economies defined for a given level of congestion), first-best charges will lead to the elimination of the least efficient airline when asymmetry exists.

Next, undifferentiated charges, which we treat as an instrument similar to the selling of slots, do not generally drive out the least efficient airline. The optimal uniform charge is a weighted average of first-best charge rules, and is likely to be less efficient as airlines become more different in their cost-efficiency, so that first-best charges would be more strongly differentiated.

The third instrument, slot trading, is an ineffective policy option when the market power distortion outweighs the congestion distortion so that the optimum requires an increase in output compared to the free market outcome. While charges may be negative, and hence become subsidies, a negative tradeable slot price is hard to imagine.

If the number of tradeable slots is set below the free-market Nash levels, so that a positive shadow price results, slot trading will lead to the monopolistic outcome for the more efficient airline – provided the total number of permits is, as second-best optimality requires, not below the monopolistic output level for the most efficient airline. The policy drives out the less efficient airline, like first-best pricing does. In fact, also with symmetric airlines, the monopolistic outcome will result. The reason is that the opportunity to trade slots gives the airlines the possibility to maximize joint profits, without having to collude tacitly. Whether this monopolistic outcome is preferable to the free-market Nash equilibrium depends on the circumstances. If the congestion externality is relatively important and the market power distortion relatively unimportant because demand is relatively elastic, this monopolistic outcome becomes relatively more attractive from the social perspective. If the opposite occurs, the free-market Nash equilibrium may be more efficient than the monopolistic outcome that follows from tradeable permits.

An unambiguous welfare ranking between the two second-best instruments cannot be given. The second-best undifferentiated charge will generally succeed in improving welfare beyond the level in the free-market Nash equilibrium, whereas tradeable permits may vary between being as efficient as first-best pricing and being completely ineffective, Numerical exercises, with realistically calibrated models, should give more insight into this matter. Some other aspects of the problem that seem to deserve further study include, in no particular order of priority, the consideration of multiple markets and more complex networks, interactions between congested airports, and private versus public airports. However, because the ambiguity in welfare ranking is already found in the most simple set-up considered here, it is unlikely to disappear for certain in more elaborate settings.

References

Basso, L.J. and A. Zhang (2006) "Congestible facility rivalry in vertical structures" *Journal of Urban Economics* **61** 218-237.

Brander, J.A. and A. Zhang (1990) "Market conduct in the airline industry: an empirical investigation" *Rand Journal of Economics* **21** (4), 567-583.

Brueckner, J.K. (2001) "Airport congestion pricing when carriers have market power" *American Economic Review* **92** 1357-1375.

Brueckner, J.K. (2008) "Inefficiency of slot markets when carriers internalize airport congestion" Unpublished Manuscript, University of California at Irvine.

Brueckner, J.K. and K. Van Dender (2008) "Atomistic congestion tolls at concentrated airports: seeking a unified view in the internalization debate" *Journal of Urban Economics* forthcoming.

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

Dales, J.H. (1968) "Land, water, and ownership" *Canadian Journal of Economics* **1** 791-804.

Daniel, J.I. (1995) "Congestion pricing and capacity of large hub airports: a bottleneck model with stochastic queues" *Econometrica* **63** 103-130.

Montgomery, W.D. (1972) "Markets in licenses and efficient pollution control programs" *Journal of Economic Theory* **5** 395-418.

Oum, T.H. and Y. Zhang (1990) "Airport pricing, congestion tolls, lumpy investment and cost recovery" *Journal of Public Economics* **43** 353-374.

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

Sandmo, A. (1975) "Optimal taxation in the presence of externalities" *Swedish Journal of Economics* **77** 86-98.

Sandmo, A. (1976) "Optimal taxation: an introduction to the literature" *Journal of Public Economics* **6** 37-54.

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

Swan, W.M. and N. Adler (2006) "Aircraft trip cost parameters: a function of stage length and seat capacity" *Transportation Research* **42E** 105-115.

Verhoef, E.T., P. Nijkamp and P. Rietveld (1995) "Second-best regulation of road transport externalities" *Journal of Transport Economics and Policy* **29** (2) 147-167.

Verhoef, E.T., P. Nijkamp and P. Rietveld (1997) "Tradeable permits: their potential in the regulation of road transport externalities" *Environment and Planning B: Planning and Design* **24** (4) 527-548.

Vickrey, W.S. (1969) "Congestion theory and transport investment" *American Economic Review (Papers and Proceedings)* **59** 251-260.

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