



UNIVERSITAT  
ROVIRA I VIRGILI

## WORKING PAPERS

Col·lecció “DOCUMENTS DE TREBALL DEL  
DEPARTAMENT D’ECONOMIA”

“Congested Hubs”

Ricardo Flores-Fillol

Document de treball n° -4- 2009

**DEPARTAMENT D’ECONOMIA**  
**Facultat de Ciències Econòmiques i Empresariales**



UNIVERSITAT  
ROVIRA I VIRGILI

*Edita:*

Departament d'Economia

[http://www.fcee.urv.es/departaments/economia/public\\_html/index.html](http://www.fcee.urv.es/departaments/economia/public_html/index.html)

Universitat Rovira i Virgili

Facultat de Ciències Econòmiques i Empresariales

Avgda. de la Universitat, 1

432004 Reus

Tel. +34 977 759 811

Fax +34 977 300 661

*Dirigir comentaris al Departament d'Economia.*

Dipòsit Legal: T - 610 - 2009

ISSN 1988 - 0812

**DEPARTAMENT D'ECONOMIA**  
**Facultat de Ciències Econòmiques i Empresariales**

# CONGESTED HUBS\*

RICARDO FLORES-FILLOL<sup>†</sup>

March 2009

## Abstract

Hub congestion is a major problem and a relevant policy issue because it causes delays and many organizational problems at airports that end up implying unpleasant consequences both for air travelers and airlines. In a competitive framework in which carriers choose aircraft size, this paper suggests that airlines schedule too many flights using overly small aircraft, which constitutes a major contributor to congestion. *Two-part congestion tolls*, accounting for the congestion imposed on other carriers and the congestion imposed on all passengers, are needed to recover efficiency. Finally, we analyze the validity of the results by studying the effects of network size, airport capacity, competition in layover time, and the formation of airline alliances.

*Keywords:* congestion; hub-and-spoke networks; overprovision of frequency; congestion internalization; congestion tolls

*JEL Classification Numbers:* L13; L2; L93

---

\*I am indebted to Jan K. Brueckner for extremely helpful suggestions and to Jordi Jaumandreu and Pedro Cantos-Sánchez for additional comments. I acknowledge financial support from the Spanish Ministry of Education and Science (SEJ2006-00538 and SEJ2007-67891) and Generalitat de Catalunya (2005SGR00949). I am also grateful to all the comments received during the congresses *Government Restructuring: Governments and Markets* at Real Colegio Complutense - Harvard University (US), *XXIII Jornadas de Economía Industrial* at Reus (Spain) and *XXXIII Simposio de Análisis Económico* at Zaragoza (Spain); and during my seminars at Universitat de Valencia (Spain), Universitat Rovira i Virgili at Reus (Spain) and University of Macedonia at Thessaloniki (Greece).

<sup>†</sup>Departament d'Economia, Universitat Rovira i Virgili, Avinguda de la Universitat 1, 43204 Reus, Spain. Tel.: +34977759850; fax: +34977759810; email: ricardo.flores@urv.cat.

# 1 Introduction

Airport congestion is a growing problem which is especially worrisome in hub-and-spoke network structures.<sup>1</sup> With hubs distributing connecting traffic from different origins, local problems may have a general impact because they are transmitted throughout the network, and problems affecting hubs become especially serious. In this framework, hub congestion is a major concern and a relevant policy issue because it causes flight delays, cancellations, missed connections and many organizational breakdowns at airports that end up implying unpleasant consequences both for air travelers and airlines.

Congestion problems are mainly a consequence of a growing demand for air travel in a framework characterized by the emergence of low-cost carriers and the competitive response by legacy carriers. Currently, the National Airspace System (NAS) handles 750 million passengers each year and the Federal Aviation Administration (FAA) expects this number to reach one billion by 2015.<sup>2</sup> To meet this growing demand, building additional runway capacity is not always possible due to physical constraints (e.g., New York-LaGuardia) and environmental constraints (e.g., Long Beach-Daugherty Field). In addition, even in airports where runway expansion is possible, the lead-time to bring a planned improvement from concept to commissioning may be substantial (between 10 and 15 years), and the cost may be really high. Thus, leaving out the possibility of adding new airport capacity, airlines can increase flight frequency, aircraft size or load factor. Since high load factors are a prerequisite for profitable operations, they do so relatively well with an industry average of around 75%.<sup>3</sup> This implies that airlines are left with two instruments: flight frequency and aircraft size, as suggested in Givoni and Rietveld (2006).

In this competitive environment, congestion problems are exacerbated by airlines' behavior that may prefer to increase flight frequency (to reduce passengers' schedule delay) using smaller capacity aircraft (regional jets or even turboprops).<sup>4</sup> As a consequence, it simply takes more operations to move the same number of people from and to the airport and the congestion problem is aggravated.

Previous analyses of airport congestion are incapable of addressing the interplay between flight frequency and aircraft size because they just focus on total traffic at the airport. Thus, this paper fills this gap by adding airport congestion in the frequency-choice model of Brueckner and Flores-Fillol (2007) and the results show that, in equilibrium, airlines

schedule too many flights using overly small aircraft, which constitutes a major contributor to congestion.

Given the rising importance of congestion, a literature on airport congestion has recently emerged. On the one hand, Brueckner (2002, 2005) and Mayer and Sinai (2003) point out that, differently to road users, airlines internalize their own congestion and therefore airport congestion tolls should be lower than atomistic tolls. On the other hand, Daniel (1995) and Daniel and Harback (2007), despite recognizing the potential for internalization, support the idea that carriers behave atomistically due to the competitive pressure exerted by fringe carriers. Finally, Brueckner and Van Dender (2008) try to reach a consensus in the internalization debate considering different competitive scenarios, and using a simple and tractable way of capturing congestion which consists in collapsing the peak and offpeak periods from Brueckner's (2002) analysis into a single period where congestion is always present. They conclude that, under Cournot behavior, carriers internalize their own congestion; under a Stackelberg model with a Cournot follower the leader internalizes a lower share of its congestion; and under a Stackelberg model with fringe carriers the leader behaves atomistically and does not internalize any congestion. We extend the single-route congestion analysis in Brueckner and Van Dender (2008) to a simple hub-and-spoke network structure.

In fact, the originality of the present paper lies in putting together the simple way of capturing congestion proposed in Brueckner and Van Dender (2008) and the modeling elements in Brueckner and Flores-Fillol (2007) that allow to study the effects of frequency and aircraft size choices on airport congestion in the framework of a duopoly model. Other important modeling elements inherited from Brueckner and Flores-Fillol (2007) are the fact that travelers exhibit *brand loyalty* (i.e., they have a utility gain from using a particular airline)<sup>5</sup> and the presence of *economies of traffic density* (i.e., economies from operating a larger aircraft), which are unequivocal elements in the airline industry.

Concerning carriers' internalization of congestion, we conclude that airlines only internalize their own congestion, neglecting the congestion they impose on the other airline and on all passengers. The fact that airlines fail to internalize the congestion inflicted on other carriers is well documented in the literature, but the failure to internalize passenger

congestion is an important contribution of the present analysis. As a consequence, the main novelty with respect to other toll structures analyzed in the previous literature, is that the toll we suggest takes into account two elements: the congestion imposed on all passengers and the congestion imposed on the other airlines. We refer to these kind of tolls as *two-part congestion tolls*.

An important element in our setting is that markets are fully served and therefore airlines exert no monopoly power over any passenger and the exercise of market power only affects the division of a fixed traffic pool between the carriers through their choices of fares and frequencies. In a more general version of the model with partially-served markets, airlines would internalize part of the congestion inflicted on passengers. However, for the sake of simplicity and to have clear results, we rule out this possibility.<sup>6</sup> In presence of market power, it is difficult to have unambiguous effects because a reduction in a carrier's flight volume mitigates airport congestion but raises 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. While uninternalized congestion again tends to make flight volumes excessive, carriers have a new incentive to limit their flight volumes in order to raise fares, yielding to an unclear net effect.

Finally, from a broader perspective, the paper also elucidates the particular effect of different factors affecting hub congestion through a number of extensions. We conclude that: (i) larger networks lead to a more inefficient choice of frequency and aircraft size; (ii) a more inefficient choice of layover time is associated to a less efficient choice of flight frequency; (iii) larger airport capacity leads to smaller congestion tolls; and (iv) the formation of airline alliances yields a more efficient situation because allied carriers internalize the congestion they impose on partner airlines.

The plan of the paper is as follows. Section 2 presents the base case obtaining the equilibrium and the social optimum and computing the corresponding congestion tolls. Section 3 suggests a range of possible elements at stake that may affect the choice of frequency and congestion tolls. Finally, a brief concluding section closes the paper.

## 2 The base case

In a rather simple setting, this section presents the equilibrium, the social optimum and the congestion tolls required to achieve efficiency.

### 2.1 Derivation of traffic levels

We assume the simplest possible network with three cities (a hub  $H$  and two spoke airports  $A$ ,  $B$ ), two airlines (1 and 2), and three city-pair markets: two *local* markets ( $AH$  and  $BH$ ) and a *connecting* market ( $AB$ ), as shown in Figure 1. Markets  $AH$  and  $BH$  are served nonstop and  $AB$  is served indirectly via hub  $H$ . Passenger population size in each of the city-pair markets is normalized to unity and it is assumed that all the passengers undertake travel.

In some previous models, a reduction in a carrier's flight volume reduces airport congestion while 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. While uninternalized congestion again tends to make flight volumes excessive, carriers have a new incentive to limit their flight volumes in order to raise fares yielding to an ambiguous net effect. To focus solely on the congestion issue, we limit the effect of market power by assuming fully-served markets so that airlines exert no monopoly power over any passenger.<sup>7</sup> Therefore, the exercise of market power only affects the division of a fixed traffic pool between the carriers through their choices of fares and frequencies.

Utility for a consumer is given by  $c + \text{travel benefit} - \text{expected schedule delay} - \text{congestion damage}$ . Firstly,  $c$  is consumption expenditure and equals  $y - p_i$  for consumers using airline  $i$  with  $i = 1, 2$ , where  $y$  denotes income which is assumed to be uniform across consumers without loss of generality, and  $p_i$  is airline  $i$ 's fare.

Secondly, *travel benefit* has two components as in Brueckner and Flores-Fillol (2007):  $b$ , equal to the gain from travel; and  $a$ , the airline brand-loyalty variable. Without brand loyalty, the airline with the most attractive frequency/fare combination would attract all the passengers in the market. However, in presence of brand loyalty, consumers are presumed to have a preference for a particular carrier, which means that an airline with an inferior frequency/airfare combination can still attract some passengers. This approach is formalized by specifying a utility gain from using airline 1 rather than airline 2, denoted

$a$ , and assuming that this gain is uniformly distributed over the range  $[-\alpha/2, \alpha/2]$ , so that half the consumers prefer airline 1 (and have  $a > 0$ ) and half prefer airline 2 (and have  $a < 0$ ). Therefore,  $a$  varies across consumers. Interestingly,  $\alpha$  is a measure of (exogenous) product differentiation in the sense that a small  $\alpha$  indicates similar products and thus small gain from using one airline or the other; whereas a big  $\alpha$  allows for significant utility gains depending on passenger's preferred carrier.

Thirdly, the *expected schedule delay* is modeled as in Brueckner (2004) and Brueckner and Flores-Fillol (2007). Let  $H$  denote the time circumference and assume a uniform distribution of consumers in terms of preferred departure time along the circle. In this framework, the schedule delay is defined as the difference between the preferred and actual departure times, and the expected schedule delay of airline  $i$  equals  $H/4f_i$ , where  $f_i$  is number of (evenly spaced) flights operated by carrier  $i$ , with  $i = 1, 2$ . Introducing a parameter  $\delta > 0$  capturing the disutility of the delay we get  $\delta H/4f_i$  and, defining  $\gamma \equiv \delta H/4$ , we obtain the final expression  $\gamma/f_i$ .<sup>8</sup>

Finally, *congestion damage* captures the extra time cost per passenger due to congestion and the resulting delays, and differs between local and connecting city-pair markets. As in Brueckner and Van Dender (2008), we collapse the peak and offpeak periods from Brueckner's (2002) analysis into a single period where congestion is always present. In this setup, congestion depends on the overall aircraft movements registered at each airport. Note that the spoke airports only serve one route that connects with the hub, and thus the number of aircraft movements at these two airports is  $f_1 + f_2$ , i.e., the sum of both carriers' frequency on the mentioned route. On the other hand, the number of aircraft movements at the hub is  $2f_1 + 2f_2$  because the hub connects with the two spoke airports, reflecting that congestion is typically a hub-related phenomenon. Consequently, congestion damage for local passengers equals  $\lambda(3f_1 + 3f_2)$  since they experience congestion at one spoke and at the hub, where  $\lambda \geq 0$  is the disutility of congestion; whereas congestion damage for connecting passengers is  $\lambda(4f_1 + 4f_2)$  because they experience congestion at the three airports.

Hence, the utility of a *local* passenger traveling with carrier 1 is

$$u_1 = y - p_1 - \frac{\gamma}{f_1} - \lambda(3f_1 + 3f_2) + a. \quad (1)$$

In a similar way, the utility of a *connecting* passenger traveling with carrier 1 is

$$U_1 = y - P_1 - \frac{\gamma}{f_1} - \lambda(4f_1 + 4f_2) + a, \quad (2)$$

where  $P_i$  denotes the  $AB$  fare with  $i = 1, 2$ .<sup>9</sup>

The analysis that follows derives the demand functions. It is just presented for carrier 1 for simplicity reasons. The corresponding expressions for carrier 2 are derived analogously. A local passenger loyal to 1 (thus with  $a > 0$ ) will fly with her preferred carrier when  $y - p_1 - \gamma/f_1 + a > y - p_2 - \gamma/f_2$ , or equivalently when  $a > p_1 - p_2 + \gamma/f_1 - \gamma/f_2 \equiv \hat{a}$ . Therefore, there is a minimum required brand-loyalty  $\hat{a}$  and only those passengers with  $a > \hat{a}$  will undertake air travel with airline 1. We observe that the brand-loyalty threshold  $\hat{a}$  increases with carrier 1's airfare and schedule delay, relative to the ones determined by carrier 2. Otherwise, passengers will choose airline 2. Then, carrier 1's local traffic is given by  $q_1 = \int_{\hat{a}}^{\alpha/2} \frac{1}{\alpha} da$  where  $1/\alpha$  gives the density of  $a$ .

Carrying out the integration, we obtain the following expression:

$$q_1 = \frac{1}{2} - \frac{1}{\alpha}(p_1 - p_2 + \gamma/f_1 - \gamma/f_2). \quad (3)$$

Similarly, connecting traffic is given by

$$Q_1 = \frac{1}{2} - \frac{1}{\alpha}(P_1 - P_2 + \gamma/f_1 - \gamma/f_2), \quad (4)$$

where capital letters denote traffic and fares in market  $AB$ . Carrier 2's demand functions are identical after interchanging subscripts.

Quite interestingly, both demand functions are independent of *passengers' congestion damage* because this term cancels out when comparing utilities. As a consequence, airlines will not take into account the congestion they impose on passengers.

## 2.2 Airline costs and profits

To characterize the equilibrium in airfares and frequencies, we need to specify the airline's cost structure. Airport congestion also affects airlines through an increase of operating costs. It is important to point out that costs borne by airlines are route-dependent (and not market-dependent),<sup>10</sup> so that they depend on the number of links operated by the airline (i.e.,  $AH$  and  $BH$ ).

Thus, a *flight's operating cost* on a certain route is given by  $\theta + \tau s_1 + \eta(3f_1 + 3f_2)$  where  $s_1$  stands for carrier 1's aircraft size (i.e., the number of seats). The parameter  $\theta$  is the marginal cost per departure (or aircraft fixed cost) that captures the cost of fuel,

airport maintenance, renting the gate to board and disembark the passengers and landing and air-traffic control fees; and the parameter  $\tau$  is the marginal cost per seat which involves serving the passenger on the ground and on the air. Finally,  $\eta(3f_1 + 3f_2)$  is the airline's congestion cost on the considered route, with  $\eta \geq 0$ . Note that the level of congestion on a route is caused by aircraft movements both at the hub ( $2f_1 + 2f_2$ ) and at the spoke airport ( $f_1 + f_2$ ).

As in Brueckner (2004), it is assumed that all seats are filled,<sup>11</sup> so that load factor equals 100% and therefore

$$s_1 = (q_1 + Q_1) / f_1, \quad (5)$$

i.e., aircraft size can be determined residually dividing airline's total traffic on a route (i.e., local+connecting) by the the number of planes. Note that cost per seat, that can be written  $[\theta + \eta(3f_1 + 3f_2)] / s_1 + \tau$ , visibly decreases with  $s_1$ , capturing the presence of economies of traffic density (i.e., economies from operating a larger aircraft) that are unequivocal in the airline industry. In other words, having a larger traffic density on a certain route reduces the impact of the cost associated with frequency.

Therefore, carrier 1's total cost from operating on a route is  $f_1[\theta + \tau s_1 + \eta(3f_1 + 3f_2)]$  or equivalently

$$c_1 = \theta f_1 + \tau(q_1 + Q_1) + f_1 \eta(3f_1 + 3f_2). \quad (6)$$

Thus, airline 1's profit is  $\pi_1 = 2p_1 q_1 + P_1 Q_1 - 2c_1$ , that can be rewritten using (6) as

$$\pi_1 = \underbrace{2(p_1 - \tau) q_1}_{\text{Local margin}} + \underbrace{(P_1 - 2\tau) Q_1}_{\text{Connecting margin}} - \underbrace{2f_1[\theta + 3\eta(f_1 + f_2)]}_{\text{Congestion and fixed cost}}, \quad (7)$$

indicating that variable costs are independent of the number of flights, and that there are two local markets using one route and one connecting market making use of two routes. The corresponding expression for carrier 2 is identical to (7) after interchanging subscripts.<sup>12</sup>

### 2.3 Equilibrium and comparative statics

After plugging (3) and (4) into (7) and maximizing, the first-order conditions for fares are

$$\frac{\partial \pi_1}{\partial p_1} = 1 - \frac{2}{\alpha} (2p_1 - p_2 + \gamma/f_1 - \gamma/f_2 - \tau) = 0, \quad (8)$$

$$\frac{\partial \pi_1}{\partial P_1} = \frac{1}{2} - \frac{1}{\alpha} (2P_1 - P_2 + \gamma/f_1 - \gamma/f_2 - 2\tau) = 0. \quad (9)$$

Since carriers are symmetric, the symmetric equilibrium is the natural focus, and this equilibrium is found by setting  $p_1 = p_2 = p$  and  $P_1 = P_2 = P$ . In this case, after substituting into (8) and (9), we obtain

$$p^* = \tau + \alpha/2 \text{ and } P^* = 2\tau + \alpha/2, \quad (10)$$

so that the local airfare equals the marginal cost of a seat ( $\tau$ ) plus a markup that depends on the degree of product differentiation ( $\alpha/2$ ); and the connecting fare is similar but takes into account the fact that two routes are needed to serve this market. As differentiation disappears, the fare converges to the marginal cost, recovering the Bertrand-equilibrium outcome. The same results are obtained in Flores-Fillol (2009b) in a similar setup.

The first-order condition for frequency is  $\frac{\partial \pi_1}{\partial f_1} = \frac{\gamma}{\alpha f_1^2} (2p_1 + P_1 - 4\tau) - 6\eta (2f_1 + f_2) - 2\theta = 0$  and,<sup>13</sup> taking into account the equilibrium fare values in (10), we obtain

$$\frac{3\gamma}{\alpha f_1^2} - 2\theta - 6\eta (2f_1 + f_2) = 0. \quad (11)$$

Setting  $f_1 = f_2 = f$  and rearranging, we get the following equilibrium condition for flight frequency

$$\underbrace{6\eta f^3}_{A(f)} = \underbrace{\frac{\gamma}{2} - \frac{2\theta}{3} f^2}_{B(f)}. \quad (12)$$

The equilibrium frequency is found graphically, as shown in Figure 2, where we observe that the  $f^*$  solution occurs at the intersection between a cubic expression ( $A(f)$ ) and a quadratic expression ( $B(f)$ ). Looking at (12) along with Figure 2, it is easy to carry out a comparative-static analysis for all the parameters in the model.

An increase in carriers' congestion cost ( $\eta$ ) raises the height of the cubic curve, leading to a decrease in  $f^*$ . The reduction of the equilibrium flight frequency is a natural reaction to a more damaging congestion.

When the disutility of schedule delay ( $\gamma$ ) rises, the intercept of the quadratic expression increases, leading to a higher  $f^*$ . Quite intuitively, carriers respond to a rise in the disutility of schedule delay by increasing flight frequency.

An increase in the aircraft fixed cost ( $\theta$ ) leads to a higher  $f^2$ -coefficient and, as a consequence,  $B(f)$  becomes more concave and  $f^*$  decreases. As expected, equilibrium frequency falls when the cost of frequency rises.

Finally, looking at travel volumes, we observe that  $q_1^* = Q_1^* = 1/2$  so that each airline carries half of the demand in every city-pair market.

## 2.4 Social optimum and congestion tolls

With the comparative-static properties of the equilibrium understood, attention now shifts to welfare analysis, where a social planner dictates flight frequency. The planner maximizes a welfare function composed by consumer surplus and carriers' profits, i.e.,  $W = \underbrace{2(u_1 + u_2)}_{AH+BH} + \underbrace{U_1 + U_2}_{AB} + \pi_1 + \pi_2$ . After integrating across agents,<sup>14</sup> this expression can be rewritten as

$$W = 3y + \frac{3\alpha}{4} - \underbrace{\frac{3}{2} \left( \frac{\gamma}{f_1} + \frac{\gamma}{f_2} \right)}_{\text{Schedule delay cost}} - \underbrace{2(\theta f_1 + \theta f_2 + 2\tau)}_{\text{Fixed and seat cost}} - \underbrace{2(f_1 + f_2)[5\lambda + 3f_1\eta + 3f_2\eta]}_{\text{Congestion costs for airlines and passengers}}, \quad (13)$$

where fares cancel out since they are pure transfers between passengers and airlines. The planner tries to minimize the costs related to scheduled delay, aircraft operations and congestion. The condition for optimal choice of  $f_1$  is

$$\frac{3\gamma}{2f_1^2} - 2\theta - 10\lambda - 12\eta(f_1 + f_2) = 0, \quad (14)$$

and, after applying symmetry, we obtain the following social-optimum condition:

$$\underbrace{8\eta f^3}_{C(f)} = \underbrace{\frac{\gamma}{2} - \frac{2(\theta + 5\lambda)}{3} f^2}_{D(f)}. \quad (15)$$

The expressions (12) and (15) are easily compared. Note that the social-optimum cubic function is higher than the equilibrium one (i.e.,  $A(f) < C(f)$  for  $f > 0$ ); and that the  $f^2$ -coefficient in the quadratic function is larger in the social-optimum condition, as depicted in Figure 3. Therefore, equilibrium frequencies are excessive compared with the optimum (i.e.,  $f^* > f^{SO}$ ). On the other hand, we also observe that the equilibrium aircraft size is inefficiently small (i.e.,  $s^* < s^{SO}$ ) since markets are fully served and all seats are filled (see (5)). Interestingly, in absence of congestion (i.e.,  $\eta = \lambda = 0$ ) the inefficiency disappears and  $f^* = f^{SO} = \left(\frac{3\gamma}{4\theta}\right)^{1/2}$  and  $s^* = s^{SO} = \left(\frac{3\gamma}{4\theta}\right)^{-1/2}$ .

**Proposition 1** *In presence of congestion, there is an overprovision of flight frequency and aircraft size is suboptimal. In absence of congestion, both frequency and aircraft size are efficient.*

As argued by many observers, when there is congestion airlines operate too many flights using overly small aircraft (regional jets or even turboprops), and a socially preferred outcome would require less frequent flights and larger aircraft. Some authors have argued

that there is an apparent overprovision of flight frequency in the current airline unregulated environment related to the adoption of hubbing strategies that led to a concentration of traffic on the spoke routes and to an increase of flight frequency,<sup>15</sup> which were widely viewed as excessive.<sup>16</sup> Our result suggests that this may be related to the presence of congestion.

In fact, the inefficient choice of flight frequency can be seen by comparing the first-order conditions corresponding to the equilibrium analysis and the social-optimum analysis. The marginal *social* congestion cost from operating an extra flight on both of the segments equals  $10\lambda + 24\eta f$  (after imposing symmetry in (14)); and the marginal congestion costs that are taken into account by airlines are  $18\eta f$  (after imposing symmetry in (11)). Therefore, the difference between these two expressions is  $10\lambda + 6\eta f$ , which captures the part of social congestion costs that are not internalized by each airline. More precisely,  $6\eta f$  represents the congestion inflicted on the other carrier on both routes, and  $10\lambda$  is the congestion experienced by *all* passengers (including its own passengers).

In this situation, a congestion toll is needed to reach the social optimum. Since there are two routes in the network, the toll per flight will be exactly half of the previous expression evaluated at the social optimum, i.e.,

$$T = 5\lambda + 3\eta f^{SO}. \quad (16)$$

Note that the marginal congestion damage (*MCD*) from an extra flight *on a route* is given by  $5\lambda + 3\eta f_1 + 3\eta f_2$  (see (13)); and thus each carrier is charged a toll equal to the marginal congestion damage evaluated at the social optimum ( $MCD^{SO}$ ) after subtracting carrier's own internalized congestion. The inefficiency in flight frequency (and thus in aircraft size) arises because airlines only internalize their own congestion, neglecting the congestion they impose on the other airline and on all passengers.

The fact that airlines fail to internalize the congestion inflicted on other carriers is well documented in the literature.<sup>17</sup> However, the failure to internalize passenger congestion is an important contribution of the present analysis. Consequently, the main novelty with respect to other toll structures analyzed in the previous literature, is that the toll we suggest takes into account the congestion imposed on all passengers. We refer to these kind of tolls as *two-part congestion tolls*, as it is spelt out in the proposition below.

**Proposition 2** *Two-part congestion tolls are required to recover efficiency. The toll an*

*airline is charged accounts for two elements: the congestion imposed on other airlines and the congestion imposed on all passengers.*

Therefore, the rule pointed out in Brueckner and Van Dender (2008) suggesting that each airline is charged  $MCD^{SO}$  times its airport flight share (which equals 1/2 in the symmetric equilibrium) does not apply to our setting because airlines are also charged by the congestion imposed on all passengers.<sup>18</sup> In fact, the asymmetry in the treatment between passenger and airline congestion brings more realism in the analysis of airport congestion and is explained by the richer structure of our model in comparison to previous models.

Airlines' failure to internalize is balanced by levying congestion tolls that are computed taking into account carriers' neglected congestion. With the congestion-toll liabilities of  $2f_1T$  and  $2f_2T$  subtracted from the respective profit functions, the social-optimum frequency is recovered.

### **3 Other factors affecting congestion**

Having the base case as a reference, several extensions are explored in this section with the purpose of understanding other factors affecting congestion as well as possible ways to mitigate its effects.<sup>19</sup>

#### **3.1 Network size**

An element that may affect congestion is network size. Suppose that an airline that connects a hub with  $n$  non-hub cities. For any new spoke city that is incorporated into the existing network, the carrier would gain access to one new local market and  $n$  new connecting markets. Thus, airlines' connecting possibilities grow exponentially by adding new routes but, at the same time, hub airports become overloaded. In this section, we explore the effects of network size when airlines connect their hub airport with  $n$  non-hub cities.

In this case, the amount of aircraft movements at the hub airport becomes  $nf_1 + nf_2$  whereas it remains  $f_1 + f_2$  at spoke airports, as in the base case. Therefore, the utility

functions for consumers traveling with airline 1 become

$$\begin{aligned} u_1 &= y - p_1 - \frac{\gamma}{f_1} - \lambda [(n+1)f_1 + (n+1)f_2] + a \text{ and} \\ U_1 &= y - P_1 - \frac{\gamma}{f_1} - \lambda [(n+2)f_1 + (n+2)f_2] + a, \end{aligned} \tag{17}$$

where we recover the base case with  $n = 2$ . The demand functions are exactly the same as in the base case (see (3) and (4)) since the congestion term cancels out when comparing utilities. Taking into account that there are  $n$  local markets and  $n(n-1)/2$  connecting markets,<sup>20</sup> the profit function for carrier 1 becomes

$$\pi_1 = \underbrace{n(p_1 - \tau)q_1}_{\text{Local margin}} + \underbrace{\frac{n(n-1)}{2}(P_1 - 2\tau)Q_1}_{\text{Connecting margin}} - \underbrace{n\{\theta f_1 + f_1\eta[(n+1)f_1 + (n+1)f_2]\}}_{\text{Congestion and fixed cost}}. \tag{18}$$

Solving the maximization problem and applying symmetry, we obtain the same equilibrium fares as in the base case (see (10)) but the condition determining the equilibrium frequency is now

$$\underbrace{6\eta f^3}_{A(f)} = \underbrace{\frac{\gamma}{2} - \frac{2\theta}{(n+1)}f^2}_{B(f)}, \tag{19}$$

which shows that the equilibrium frequency is logically increasing with network size since the  $f^2$ -coefficient in the quadratic function falls with  $n$  whereas the cubic expression is as in the base case. The social optimum condition is now

$$\underbrace{8\eta f^3}_{C(f)} = \underbrace{\frac{\gamma}{2} - \frac{2\theta + \lambda n(n+3)}{(n+1)}f^2}_{D(f)}, \tag{20}$$

which illustrates that  $f^{SO}$  is also increasing with network size. Comparing (19) and (20), it is easy to check that equilibrium flight frequency remains excessive and thus aircraft size remains suboptimal, as in the base case. Quite interestingly, we can show that the overprovision of flight frequency is aggravated as  $n$  increases for for  $n > n^* = \sqrt{\frac{2\theta}{\lambda} - 2} - 1$ , i.e.,  $|f^* - f^{SO}|$  increases with  $n$ , as it is summarized in the proposition below. The proof is provided in the Appendix.

**Proposition 3** *If the number of non-hub cities is sufficiently large, then larger networks make airline frequency and aircraft size choices more inefficient in presence of congestion.*

Hence, it seems that airlines decide indirectly the amount of hub congestion they produce by choosing network size, as put forward by Mayer and Sinai (2003). In fact, carriers face the following trade-off: by limiting the number of routes, they reduce their own congestion and choose flight frequency and aircraft size more efficiently; but, at the same time, they incur in a loss in terms of lower possibilities of feeding connecting markets with passengers coming from different endpoints.<sup>21</sup>

### 3.2 Layover time

As airlines keep on scheduling flights on top of successful banks, aircraft may become too full to accommodate the extra traffic. In fact, as Mayer and Sinai (2003) argue, better hub connections comes at the cost of higher congestion. In this framework, depeaking traffic banks at hubs by increasing passengers' layover time seems an easy way to reduce hub congestion. This strategy consisting on depeaking hubs has been firstly adopted by American Airlines at Chicago O'Hare airport (which has been also called the "rolling hub" system). We will explore the effects of such an strategy by analyzing the effects of layover time on congestion.

Brueckner (2004), in a monopoly model without congestion, introduces layover time as a shift factor reducing the utility of connecting passengers since they dislike waiting. However, as we suggested before, layover time may also have a positive effect when we introduce congestion in the analysis. In this extension, we let carriers choose layover time in a competitive framework in presence of congestion.

It is important to notice that, spreading flights at hubs yields an analogous effect in the spoke airports that serve the hub. Therefore, the choice of layover time has an overall impact in the way carriers organize their frequencies. Thus, the utility functions for consumers traveling with airline 1 become

$$\begin{aligned}
 u_1 &= y - p_1 - \frac{\gamma}{f_1} - \lambda \left( \frac{3f_1}{\mu_1} + \frac{3f_2}{\mu_2} \right) + a \text{ and} \\
 U_1 &= y - P_1 - \mu_1 - \frac{\gamma}{f_1} - \lambda \left( \frac{4f_1}{\mu_1} + \frac{4f_2}{\mu_2} \right) + a,
 \end{aligned}
 \tag{21}$$

where  $\mu_i$  stands for carrier  $i$  layover time with  $i = 1, 2$ . Thus, layover time reduces connecting passengers' utility since they dislike waiting and relaxes carriers' own congestion effect (note that the base case is recovered with  $\mu_1 = \mu_2 = 1$ ). The demand function for

local passengers is the same as in the base case (see (3)) and the demand function for connecting passengers becomes

$$Q_1 = \frac{1}{2} - \frac{1}{\alpha}(P_1 - P_2 + \mu_1 - \mu_2 + \gamma/f_1 - \gamma/f_2), \quad (22)$$

where we observe that layover time introduces another source of competition with respect to the base case (now airlines compete in fares, frequencies and layover time).

Finally, the congestion supported by airlines is also reduced and the cost function turns into

$$c_1 = \theta f_1 + \tau (q_1 + Q_1) + f_1 \eta \left( \frac{3f_1}{\mu_1} + \frac{3f_2}{\mu_2} \right). \quad (23)$$

Maximizing profits and applying symmetry, we obtain the same equilibrium fares as in the base case (see (10)) but the condition determining the equilibrium frequency is now

$$6 \frac{\eta}{\mu^*} f^3 = \frac{\gamma}{2} - \frac{2\theta}{3} f^2. \quad (24)$$

Thus, as suggested before, layover time has a double effect because it has a negative impact on passengers (that dislike waiting) but, at the same time, it mitigates the negative effects of congestion on carriers, allowing them to increase flight frequency. From the first-order condition for  $\mu_1$ , after applying symmetry, we obtain

$$\mu^* = [2f^* (6\eta f^*)]^{1/2}, \quad (25)$$

where  $6\eta f^*$  is carrier's own congestion on the two routes it serves. Quite naturally, carriers increase layover time as their own congestion problem becomes more serious. Additionally, the positive relationship between  $f$  and  $\mu$  in equilibrium is corroborated.

Shifting attention to the social optimum, the optimal frequency is determined by

$$8 \frac{\eta}{\mu^{SO}} f^3 = \frac{\gamma}{2} - \frac{2}{3} \left( \theta + 5 \frac{\lambda}{\mu^{SO}} \right) f^2, \quad (26)$$

where layover time reduces the congestion damage both for airlines and passengers. The socially-optimal layover time is given by

$$\mu^{SO} = [2f^{SO} (10\lambda + 12\eta f^{SO})]^{1/2}. \quad (27)$$

In this case, we observe that the socially optimal layover time increases with overall congestion (that includes congestion supported both by passengers and carriers). Therefore,

carriers' choice of  $\mu$  is inefficient because they do not take into account the congestion imposed on passengers and on the other carrier, as it is spelt out in the lemma that follows.

**Lemma 1** *When carriers compete in layover time, the equilibrium layover time is inefficient because carriers do not take into account the congestion imposed on other airlines and the congestion imposed on all passengers.*

Furthermore, by comparing the equilibrium and the social-optimum layover time, it can be checked that  $\mu^{SO} > \mu^*$  requires  $\frac{5\lambda}{3\eta} > \frac{(f^*)^2 - 2(f^{SO})^2}{f^{SO}}$ , a condition that will hold in a framework with moderate overprovision of frequency (a sufficient condition for  $\mu^{SO} > \mu^*$  would be  $f^* < \sqrt{2}f^{SO}$ ). Therefore, in a situation with overprovision of flight frequency, as the gap  $f^* - f^{SO}$  closes (i.e., less overprovision), the gap  $\mu^{SO} - \mu^*$  increases. Thus, a sub-optimal choice of layover time can partially correct a situation with excessive frequencies. The proposition below summarizes these findings.

**Proposition 4** *In a situation with overprovision of flight frequency, a more inefficient choice of layover time is associated to a less inefficient choice of flight frequency.*

The conclusion from Proposition 4 is that layover time seems to be a valid tool for carriers to reduce their own congestion and mitigate the overprovision of flight frequency.

### 3.3 Airport capacity

A potential solution to airport congestion is to invest in new runways. We introduce airport capacity in the model just by replacing  $\lambda$  and  $\eta$  in the base case by  $\lambda/K$  and  $\eta/K$  respectively where  $K$  stands for airport capacity. We assume that the decisions on airport capacity are made by the airport management authority so that airlines cannot influence the actual level of  $K$  (the base case is recovered with  $K = 1$ ).

The new equilibrium and social-optimum conditions are as in the base case, after replacing  $\eta$  and  $\lambda$  by  $\eta/K$  and  $\lambda/K$  respectively, i.e.,

$$6\frac{\eta}{K}f^3 = \frac{\gamma}{2} - \frac{2\theta}{3}f^2 \quad \text{and} \quad 8\frac{\eta}{K}f^3 = \frac{\gamma}{2} - \frac{2}{3}\left(\theta + 5\frac{\lambda}{K}\right)f^2, \quad (28)$$

where we observe that new runway capacity mitigates the negative effects of congestion both on passengers and carriers, allowing them to increase flight frequency. Thus, an increase in  $K$  leads to a higher  $f^*$ , as a natural airline reaction to a lower congestion

cost. Logically, the new social-optimum condition changes in a similar fashion and the overprovision of frequency persists. Two-part congestion tolls are now given by  $T = 5\frac{\lambda}{K} + 3\frac{\eta}{K}f^{SO}$  and, therefore, they decrease with  $K$ .

**Proposition 5** *As airport capacity increases, smaller two-part congestion tolls are needed.*

As  $K$  increases, the inefficiency associated to the presence of congestion is reduced and the needed tolls are smaller. In the limit, when  $K \rightarrow \infty$  the inefficiency disappears ( $f^* \rightarrow f^{SO}$ ) and tolls vanish ( $T \rightarrow 0$ ). Therefore, in the hypothetical case of costless and unrestricted airport expansions, congestion could be eradicated.

Unfortunately, in reality airport capacity is costly. In the reasoning that follows, we consider the decision of the airport management authority that dictates the optimal level of  $K$ . Assuming a public-owned airport, the airport management authority behaves as a social planner but the welfare function (equivalent to (13)) is decreased by the cost of capacity  $P_K K$ , where  $P_K$  is the cost of one unit of  $K$ . In this case, the optimal choice for  $K$  is

$$K^{SO} = [4f^{SO} (5\lambda + 6\eta f^{SO}) / P_K]^{1/2}, \quad (29)$$

which is intuitively decreasing with  $P_K$ . From (29) and the congestion toll expression, it is easy to observe that  $P_K = 2 \underbrace{\left[ 2f^{SO} \left( 5\frac{\lambda}{K^{SO}} + 3\frac{\eta}{K^{SO}} f^{SO} \right) \right]}_{\text{Toll revenue}} + 12\frac{\eta}{K^{SO}} (f^{SO})^2$ . Therefore,

isolating the toll revenue, we get

$$\text{Toll revenue} = \frac{P_K K^{SO}}{2} - 6\frac{\eta}{K^{SO}} (f^{SO})^2, \quad (30)$$

which implies that the toll revenue pays less than half of the cost of optimal airport capacity.

**Corollary 1** *If the planner has the power to determine airport capacity, the revenue generated from levying congestion tolls fails to cover the cost of the optimal-size airport.*

This result overturns the well-known self-financing rule for congested facilities which says that toll revenue exactly covers the construction cost of a congested facility built with constant returns to scale, confirming the result in Brueckner (2002) in a different setup.

Airport expansions already constitute a restricted tool to mitigate congestion because of their long gestation period (between 10 and 15 years), and the existing physical constraints (e.g., New York-LaGuardia) and/or environmental constraints (e.g., Long Beach-Daugherty Field). In addition, as a consequence of the above corollary, they also seem to involve an important financial burden.

### 3.4 Alliances

Let us consider an international context in which we have four carriers, each of them being a duopolist in a local market as shown in Figure 4. We assume that airlines 1 and 3 (respectively 2 and 4) are partners in the connecting market, which is also a duopoly market, as in Flores-Fillol (2009a). Therefore, the base case can be reconsidered as the result of a double complementary alliance (or merger)<sup>22</sup> between airlines 1 and 3 and airlines 2 and 4 (compare Figures 1 and 4).

Now we proceed to analyze the *pre-alliance* case.<sup>23</sup> In a scenario without alliances, the amount of aircraft movements is given by  $f_1 + f_2$  at airport  $A$ , and by  $f_3 + f_4$  at airport  $B$ . All the four carriers make use of the hub, and thus aircraft movements at  $H$  are  $f_1 + f_2 + f_3 + f_4$ . The utility for a local passenger flying with carrier 1 is now  $u_1 = y - p_1 - \frac{\gamma}{f_1} - \lambda(2f_1 + 2f_2 + f_3 + f_4) + a$  and the corresponding demand function is as in the base case (see (3)) because the congestion term cancels out when comparing  $u_1$  and  $u_2$ .

Two assumptions need to be made to present the utility for connecting passengers. Firstly, we assume that each partner sets noncooperatively a *subfare* for its portion of the connecting flight, with the sum of the subfares giving the connecting fare, as in Brueckner (2001) and Flores-Fillol (2009a).<sup>24</sup> Secondly, the schedule delay in the connecting markets is computed as the *average* between the schedule delay on the two routes of the network.<sup>25</sup> With this information, the utility for a connecting passenger traveling with carriers 1 and 3 is given by

$$U_{13} = y - (s_1 + s_3) - \frac{2\gamma}{f_1 + f_3} - \lambda(2f_1 + 2f_2 + 2f_3 + 2f_4) + a, \quad (31)$$

where  $s_1$  and  $s_3$  are the subfares set by both partners; and the corresponding demand function is

$$Q_{13} = \frac{1}{2} - \frac{1}{\alpha} \left( s_1 + s_3 - s_2 - s_4 + \frac{2\gamma}{f_1 + f_3} - \frac{2\gamma}{f_2 + f_4} \right). \quad (32)$$

Therefore, the profit function for carrier 1 becomes

$$\pi_1 = \underbrace{(p_1 - \tau) q_1}_{\text{Local margin}} + \underbrace{(s_1 - \tau) Q_{13}}_{\text{Connecting margin}} - \underbrace{f_1 [\theta + \eta (2f_1 + 2f_2 + f_3 + f_4)]}_{\text{Congestion and fixed cost}}. \quad (33)$$

Solving the maximization problem and applying symmetry, we obtain that  $p^* = s^* = \tau + \alpha/2$  which means that local fares are not affected by complementary alliances. Yet, the condition determining the equilibrium frequency is now

$$\underbrace{\frac{16}{3}\eta f^3}_{E(f)} = \underbrace{\frac{\gamma}{2} - \frac{2}{3}\theta f^2}_{B(f)}, \quad (34)$$

where the quadratic expression is as in the base case but the cubic expression is now smaller (i.e.,  $E(f) < A(f)$  for  $f > 0$ ). Consequently, the equilibrium frequency is larger without alliances. Since the social optimum is not affected by alliances, we observe that  $f^{SO} < f^* < f_{NA}^*$  as depicted in Figure 5 where  $f_{NA}^*$  is the equilibrium frequency in the pre-alliance case; and consequently  $s^{SO} > s^* > s_{NA}^*$ . Therefore, alliances reduce the equilibrium flight frequency and mitigate the problem of the overprovision of flight frequency.

Quite intuitively, the congestion toll without alliances is larger than in the base case and is given by

$$T_{NA} = 5\lambda + 4\eta f^{SO}. \quad (35)$$

This *two-part congestion toll* accounts for the congestion inflicted on carrier 2 on route  $AH$  ( $2\eta f_2$ ), plus the congestion imposed on carriers 3 and 4 ( $\eta f_3 + \eta f_4$ ) that operate on the other route and are affected through the congestion at the hub airport. Finally, the toll also includes the congestion imposed on all passengers, which is the same than under the base case ( $5\lambda$ ) since aircraft movements are the same with and without alliances.

Recalling that airlines 1 and 3 are partners in the connecting market, the main difference with respect to the base-case toll is that now carriers do not internalize the congestion imposed on the partner airline. The proposition below summarizes the results obtained, analyzing the move from a *pre-alliance* to an *alliance* situation.

**Proposition 6** *Allied carriers internalize the congestion imposed on partner airlines. As a consequence, the equilibrium frequency is lower (and the equilibrium aircraft size is larger) under alliances; congestion tolls are also lower under alliances.*

Brueckner (2002) and in Mayer and Sinai (2003) provide empirical evidence that there is more congestion internalization in highly concentrated airports. Our result confirms this evidence since alliances make the hub become more concentrated and, consequently, more congestion is internalized. Given the increasing importance of airport congestion, the result in the above proposition provides a powerful argument in favor of airline consolidation.

## 4 Concluding remarks

Airport congestion depends on the number of aircraft operations. By using overly small aircraft, airlines schedule too many flights, aggravating the situation at hubs. At a first glance, investing in airport capacity could be seen as a policy measure to overcome the problem of congestion. However, runway expansions are constrained by physical and environmental concerns, have long gestation periods and imply a very high financial cost that make this remedy typically inefficient and impracticable. Thus, the problems associated to congestion are difficult to solve.

Policy makers should try to create the "right" incentives for carriers to use larger aircraft and thus reduce flight frequency and alleviate airport congestion. The current landing-fees system based on aircraft weight should be revised because it makes cheaper the use of small aircraft, aggravating the problem of congestion. In fact, the FAA has recently suggested the implementation of a new "two-part landing fee structure" consisting of both an operation charge and a weight-based charge (in lieu of the standard weight-based charge). Such a two-part fee would serve as an incentive for carriers to use larger aircraft and increase the number of passengers served with the same or fewer operations.<sup>26</sup>

Additionally, in an oligopoly context in which carriers internalize only a share of the congestion they generate, we suggest *two-part congestion tolls* that provide a way to implement a more efficient outcome. These tolls account for the congestion imposed on passengers and on the other airlines.

We also conclude that larger networks and lower layover times may exaggerate the inefficiency associated to congestion. However, policy implications oriented to modify airlines' network size and layover time seem difficult to justify from first principles.

More interestingly, from the interaction between congestion and airline cooperation, we

conclude that (complementary) alliances make carriers internalize their partner's congestion. This is a new argument to take into account when evaluating the pros and the cons of carrier-consolidation agreements. Thus, if the integration trend observed in the last years keeps stepping forward, in a new environment with few major players (around *SkyTeam*, *Star Alliance* and *oneworld*) a larger share of congestion should be internalized by airlines.

## References

- [1] BRUECKNER, J.K., 2001. The economics of international codesharing: an analysis of airline alliances, *International Journal of Industrial Organization* 19, 1475–1498.
- [2] BRUECKNER, J.K., 2002. Airport congestion when carriers have market power, *American Economic Review* 92, 1357-1375.
- [3] BRUECKNER, J.K., 2004. Network structure and airline scheduling, *Journal of Industrial Economics* 52, 291-312.
- [4] BRUECKNER, J.K., 2005. Internalization of airport congestion: a network analysis, *International Journal of Industrial Organization* 23, 599-614.
- [5] BRUECKNER, J.K., FLORES-FILLOL, R., 2007. Airline schedule competition, *Review of Industrial Organization* 30, 161-177.
- [6] BRUECKNER, J.K., SPILLER, P.T., 1991. Competition and mergers in airline networks, *International Journal of Industrial Organization* 9, 323-342.
- [7] BRUECKNER, J.K., ZHANG, Y., 2001. A model of scheduling in airline networks, *Journal of Transport Economics and Policy* 35, 195-222.
- [8] BRUECKNER, J.K., VAN DENDER, K., 2008. Atomistic congestion tolls at concentrated airports? Seeking a unified view in the internalization debate, *Journal of Urban Economics* 64, 288-295.
- [9] DANIEL, J.I., 1995. Congestion pricing and capacity of large hub airports: a bottleneck model with stochastic queues, *Econometrica* 63, 327-370.
- [10] DANIEL, J.I., HARBACK, K.T., 2008. (When) Do hub airlines internalize their self-imposed congestion delays?, *Journal of Urban Economics* 63, 583-612.
- [11] DOGANIS, R., 1985. *Flying off Course: The Economics of International Airlines*, George Allen & Unwin, London.
- [12] DOUGLAS, G.W., MILLER III, J.C., 1974. *Economic Regulation of Domestic Air Transport: Theory and Policy*, Brookings Institution, Washington, D.C.

- [13] FAA, 2008. Policy regarding airport rates and charges, Federal Register/Vol. 73, No. 12/Thursday, January 17, 2008/Notices.
- [14] FLORES-FILLOL, R., 2009a. Airline alliances: parallel or complementary?, *Applied Economics Letters* 16, 585-590.
- [15] FLORES-FILLOL, R., 2009b. Airline competition and network structure, *Transportation Research Part B*, forthcoming.
- [16] FLORES-FILLOL, R., MONER-COLONQUES, R., 2007. Strategic formation of airline alliances, *Journal of Transport Economics and Policy* 41, 427-449.
- [17] GIVONI, M., RIETVELD, P., 2006. Choice of aircraft size - explanations and implications, *Tinbergen Institute discussion paper* TI 2006-113/3.
- [18] 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.
- [19] MORRISON, S.A., WINSTON, C., 1995. *The Evolution of the Airline Industry*, Brookings Institution, Washington, D.C.

## Notes

<sup>1</sup>With the deregulation of the airline sector, airlines became free to set fares and make strategic network choices. In the years following the deregulation, we observe the success of hub-and-spoke structures leading the airline sector to an intense wide-ranging network reorganization.

<sup>2</sup>See FAA (2008) for more detailed information.

<sup>3</sup>Data from IATA ([www.iata.org](http://www.iata.org)).

<sup>4</sup>Another reason that creates incentives for airlines to use narrow-body aircraft is the current landing-fees system based on aircraft weight, which makes cheaper the use of small aircraft.

<sup>5</sup>In fact, brand loyalty is an important element of the airline industry, especially since the proliferation of frequent-flyer-programs and worldwide alliances. Brand loyalty may also reflect idiosyncratic consumer preferences for particular aspects of airline service that may differ across carriers.

<sup>6</sup>Brueckner and Van Dender (2008) also rule out market-power but they follow a different approach, by assuming a perfectly elastic demand for air travel so that fares are fixed and are not chosen by airlines. Then they also assume that aircraft have a fixed seat capacity and, therefore, flight frequency and air traffic are parallel measures.

<sup>7</sup>We follow the approach in Brueckner and Flores-Fillol (2007) when all passengers are high types, i.e., they always travel and markets are fully served. Notice, however, that Brueckner and Flores-Fillol (2007) also consider the possibility of having low-type passengers that are characterized by a lower valuation of travel and may not undertake air travel, which means that markets can be partially served.

<sup>8</sup>Therefore consumers compare fares ( $p_1$  and  $p_2$ ) and expected schedule delay ( $\gamma/f_1$  and  $\gamma/f_2$ ) of both airlines. While this approach may not be fully accurate for individual consumers, it appears to capture the choice setting of a corporate travel department, which must sign an exclusive contract with a particular airline for transporting its employees. The travel department cares about the average schedule delay for the company employees, while also seeking low fares. It signs an exclusive contract with the airline providing the best combination of these features. Alternatively, the model could apply to individual business travelers, who cannot predict their travel times and thus purchase refundable full-fare tickets, which allow them to board the next flight upon arriving at the airport. In either case, the precise departure times of individual flights are not relevant, accounting for the simplicity of the overall approach.

<sup>9</sup>Analogously, the utility of a *domestic* passenger traveling with carrier 2 is  $u_2 = y - p_2 - \frac{\gamma}{f_2} - \lambda(3f_1 + 3f_2) - a$ ; and the utility of a *connecting* passenger traveling with carrier 2 is  $U_2 = y - P_2 - \frac{\gamma}{f_2} - \lambda(4f_1 + 4f_2) - a$ , with  $a < 0$  for passengers loyal to carrier 2 and  $a > 0$  for passengers loyal to carrier 1.

<sup>10</sup>As suggested in Brueckner and Spiller (1991) under a very different specification; and pointed out in Flores-Fillol (2009b) in a more similar specification.

<sup>11</sup>As we said before, high load factors are a reality in the airline industry since they constitute a prerequisite for profitable operations.

<sup>12</sup>As suggested before, when maximizing profits, carriers do not take into account the congestion inflicted on passengers since demand functions are independent of passengers' congestion damage.

<sup>13</sup>The second-order conditions  $\partial^2\pi_1/\partial p_1^2, \partial^2\pi_1/\partial P_1^2, \partial^2\pi_1/\partial f_1^2 < 0$  are satisfied by inspection. The

remaining positivity condition on the Hessian determinant, which is assumed to hold, requires  $2(p_1 - \tau) + (P_1 - 2\tau) > 3\left(\frac{\gamma}{4f_1} - \frac{2\alpha\eta f_1^3}{\gamma}\right)$ , i.e., the sum of margins in the three markets operated by an airline has to be sufficiently large.

<sup>14</sup>Note that we have to carry out the integration across all the passengers, i.e.,

$$\begin{aligned}
W = & 2 \underbrace{\int_0^{\alpha/2} \left[ y - p_1 - \frac{\gamma}{f_1} - \lambda(3f_1 + 3f_2) + a \right] \frac{1}{\alpha} da}_{u_1} + 2 \underbrace{\int_{-\alpha/2}^0 \left[ y - p_2 - \frac{\gamma}{f_2} - \lambda(3f_1 + 3f_2) - a \right] \frac{1}{\alpha} da}_{u_2} \\
& + \underbrace{\int_0^{\alpha/2} \left[ y - P_1 - \frac{\gamma}{f_1} - \lambda(4f_1 + 4f_2) + a \right] \frac{1}{\alpha} da}_{U_1} + \underbrace{\int_{-\alpha/2}^0 \left[ y - P_2 - \frac{\gamma}{f_2} - \lambda(4f_1 + 4f_2) - a \right] \frac{1}{\alpha} da}_{U_2} \\
& + 2(p_1 - \tau) \underbrace{\int_0^{\alpha/2} \frac{1}{\alpha} da + (P_1 - 2\tau) \int_0^{\alpha/2} \frac{1}{\alpha} da - 2f_1 [\theta + 3\eta(f_1 + f_2)]}_{\pi_1} \\
& + 2(p_2 - \tau) \underbrace{\int_{-\alpha/2}^0 \frac{1}{\alpha} da + (P_2 - 2\tau) \int_{-\alpha/2}^0 \frac{1}{\alpha} da - 2f_2 [\theta + 3\eta(f_1 + f_2)]}_{\pi_2}.
\end{aligned}$$

<sup>15</sup>See Morrison and Winston (1995) and Brueckner (2004).

<sup>16</sup>See Douglas and Miller (1974), Brueckner and Zhang (2001) and Flores-Fillol (2009b).

<sup>17</sup>See Brueckner (2002 and 2005), Mayer and Sinai (2003) and Brueckner and Van Dender (2008).

<sup>18</sup>Levyng atomistic tolls would imply charging  $MCD^{SO}$  to each airline since these kind of tolls ignore carriers' own-congestion internalization.

<sup>19</sup>If needed, the details of the computations for each extension are available from the author upon request.

<sup>20</sup> $(n - 1)Q_1$  is the connecting traffic on a certain route.

<sup>21</sup>Treating  $n$  as a choice variable would complicate the analysis without any further insight. Thus, determining the optimal network size is out of the scope of this paper.

<sup>22</sup>As pointed out in Brueckner (2001), it makes sense to differentiate between mergers and alliances when there is network overlapping, which is not the case in our setup. When some routes are operated by both partners, mergers are superior to parallel alliances since they make larger efficiency gains by pooling passengers in larger aircrafts and thus profiting of economies of traffic density.

<sup>23</sup>The strategic formation of complementary airline alliances is studied in Flores-Fillol and Moner-Colonques (2007).

<sup>24</sup>Doganis (1985) argues that, due to the decreasing influence of IATA, there is a certain degree of coordination between carriers when determining some fares even in the absence of alliances or codesharing agreements.

<sup>25</sup>Connecting passengers care about the schedule delay on both routes. It could be argued that the relevant frequency for connecting passengers is  $\min\{f_1, f_3\}$ . For simplicity reasons, we consider the average  $(f_1 + f_3)/2$  although both approaches converge in the case with symmetric carriers.

<sup>26</sup>See FAA (2008).

# Figures

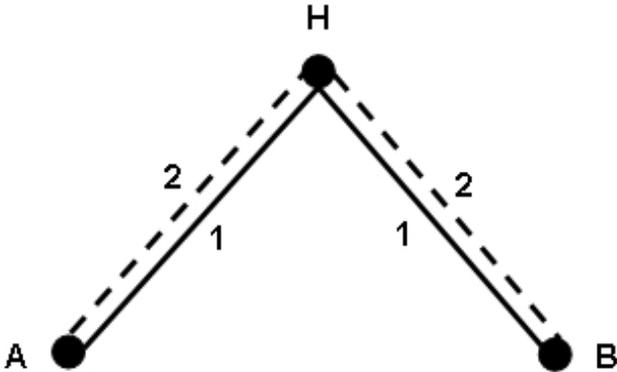


Figure 1: The network

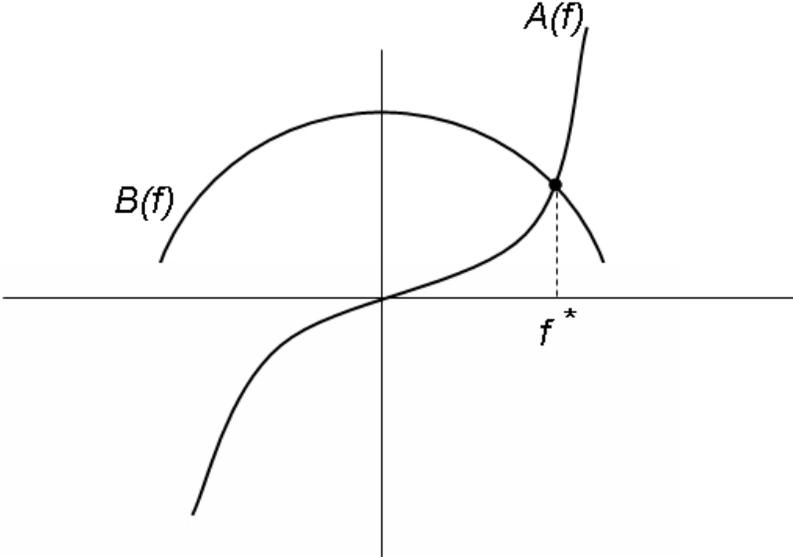


Figure 2: The  $f$  solution

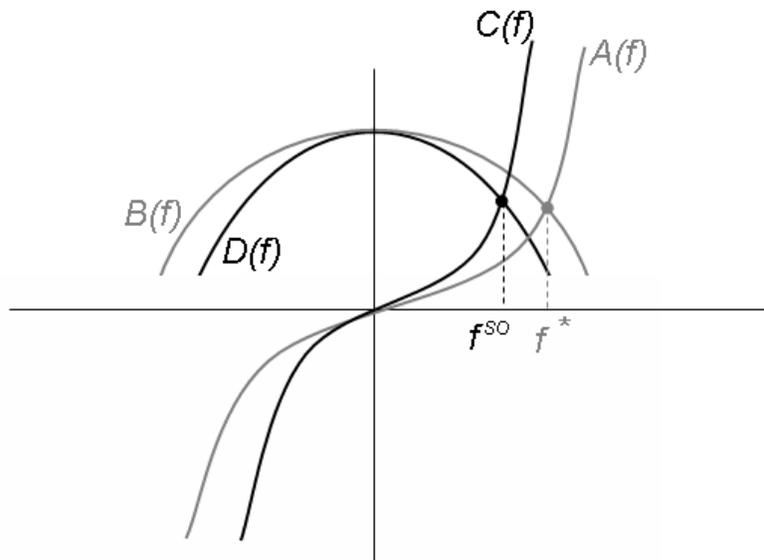


Figure 3: Overprovision of frequency

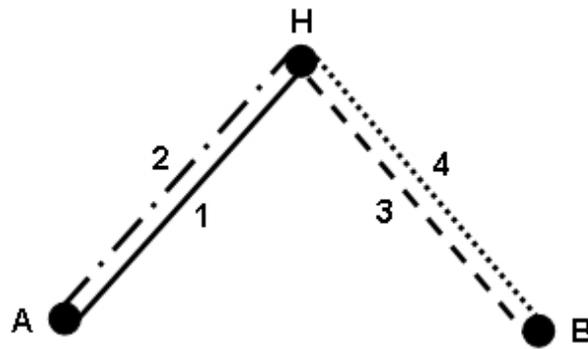


Figure 4: The no alliance case

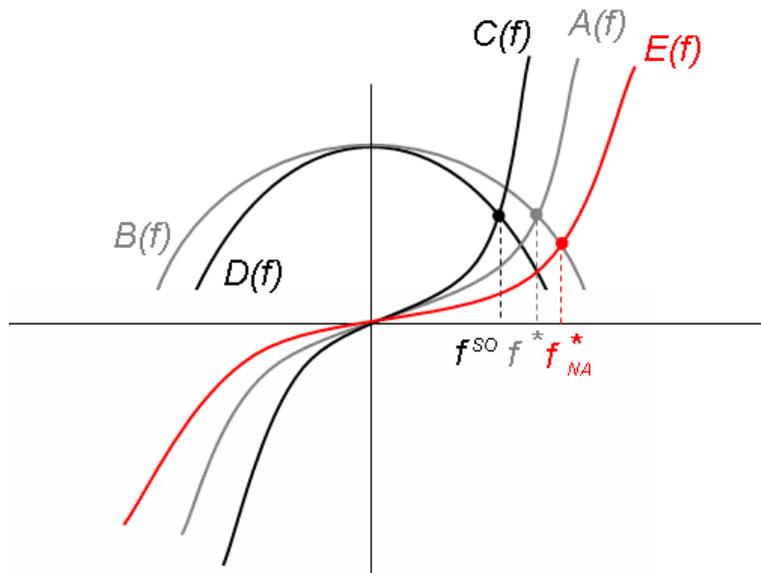


Figure 5: Alliances reduce frequency

## Appendix: Proof of Proposition 3

From (19), let us define  $\Omega^* \equiv A(f) - B(f) = 0$ , that is  $\Omega^* = 6\eta f^3 - \frac{\gamma}{2} + \frac{2\theta}{(n+1)}f^2 = 0$ . The total differential of the equilibrium frequency with respect  $n$  is

$$\frac{df^*}{dn} = \frac{-\partial\Omega^*/\partial n}{\partial\Omega^*/\partial f} = \frac{\frac{2\theta}{(n+1)^2}f^2}{18\eta f^2 + \frac{4\theta}{(n+1)}f}, \quad (\text{A1})$$

and it is easy to check that  $\frac{df^*}{dn} > 0$ .

Equivalently, from (20) we define  $\Omega^{SO} \equiv C(f) - D(f) = 0$ , that is  $\Omega^{SO} \equiv 8\eta f^3 - \frac{\gamma}{2} + \frac{2\theta + \lambda n(n+3)}{(n+1)}f^2$ . The total differential of the socially-optimal frequency with respect  $n$  equals

$$\frac{df^{SO}}{dn} = \frac{-\partial\Omega^{SO}/\partial n}{\partial\Omega^{SO}/\partial f} = \frac{-\left[-\frac{2\theta}{(n+1)^2}f^2 + \frac{\lambda(n^2+2n+3)}{(n+1)^2}f^2\right]}{24\eta f^2 + \frac{4\theta}{(n+1)}f + \frac{2\lambda n(n+3)}{(n+1)}f}. \quad (\text{A2})$$

From (A2) we observe that  $\frac{df^{SO}}{dn} < 0$  if  $\lambda(n^2 + 2n + 3) - 2\theta > 0$ . Solving this expression for  $n$ , we get  $-1 \pm \sqrt{\frac{2\theta}{\lambda} - 2}$ . Thus, for  $n > n^* = \sqrt{\frac{2\theta}{\lambda} - 2} - 1$ , then  $\frac{df^{SO}}{dn} < 0$  (note: if  $\lambda \geq \theta$ , then  $\frac{df^{SO}}{dn} < 0$  is always observed). In this case, it is easy to observe that  $|f^* - f^{SO}|$  increases with  $n$  since  $f^*(n)$  and  $f^{SO}(n)$ . ■