Published version:

doi.org/10.1016/j.euroecorev.2012.05.001

On the optimal distribution of traffic of network airlines*

Xavier Fageda[†] and Ricardo Flores-Fillol[‡]

May 2012

Abstract

Network airlines have increasingly focused their operations on hub airports through the exploitation of connecting traffi c. However, in this paper we show that they may also have incentives to divert traffi c away from their hubs. More precisely, we examine how the optimal distribution of traffi c of network carriers can be affected by the two major recent innovations in the airline industry: the regional jet technology and the low-cost business model. On the one hand, we show that a network airline may nd it pro table to serve thin point-to-point routes with regional jets when the distance between endpoints is suffi ciently short and there is a high proportion of business travelers. On the other hand, we observe that a network airline may be interested in serving thin point-to-point routes by means of a low-cost subsidiary when the distance between endpoints is longer and there is a high proportion of leisure travelers. We conclude that network airlines are using those innovations to provide services on thin routes out of the hubs.

Keywords: regional jet technology; low-cost business model; point-to-point network; hub-and-spoke network

JEL Classification Numbers: L13; L2; L93

*We are grateful to Leonardo J. Basso, Jan K. Brueckner, and Jordi Perdiguero for helpful comments. The authors acknowledge financial support from the Spanish Ministry of Science and Innovation (ECO2010-19733, ECO2010-17113, and ECO2009-06946/ECON), Generalitat de Catalunya (2009SGR900 and 2009SGR1066), Barcelona Graduate School of Economics (Research Recognition Program), and Ramón Areces Foundation. An earlier version of the paper appeared as a working paper of Fundación de las Cajas de Ahorros (FUNCAS), working paper number 628/2011.

[†]Department of Economic Policy, Universitat de Barcelona, Avinguda Diagonal 690, 08034 Barcelona, Spain. Tel.: +34934039721; fax: +34934024573; email: xfageda@ub.edu.

[‡]Departament d'Economia and CREIP, Universitat Rovira i Virgili, Avinguda de la Universitat 1, 43204 Reus, Spain. Tel.: +34977759851; fax: +34977759810; email: ricardo.flores@urv.cat.

1 Introduction

Network airlines have increasingly focused their operations on hub airports through the exploitation of connecting traffic, which has allowed them to take advantage of the economies of traffic density that characterize the airline industry. Several papers have examined airlines' network structure.[1] Less attention has been devoted to decisions of network airlines on thin point-to-point (PP) routes, which are those connecting two non-hub airports. PP routes can be served using different aircraft technologies (i.e., turboprops, regional jets and mainline jets) and different business models (i.e., using either the main brand or a low-cost subsidiary).

This paper examines the influence of two major innovations in the distribution of traffic of network airlines. First, the emergence of regional jets constitutes an important technological innovation since these aircraft can provide high-frequency services on longer routes than turboprops. Second, the emergence of a low-cost business model represents an important managerial innovation, making it possible to offer seats at lower fares (with lower flight frequency). With the adoption of these innovations, we investigate whether network airlines may have more incentives to provide services out of the hub on thin PP routes.

Using a monopoly model, Brueckner and Pai (2009) analyze the effect of the introduction of regional jets on the emergence of new PP routes, referred to as the "new routes hypothesis". While theoretically sound, this hypothesis fails to find empirical support from US data. Similarly, studying the case of Continental Airlines (focusing on its hubs in Cleveland and Houston), Dresner *et al.* (2002) find that regional jets are mainly used on new HS routes (longer than routes served with turboprops), and appear to increase demand on denser routes where they replace turboprops.

Although we borrow most of the theoretical analysis from Brueckner and Pai (2009), we also introduce some important modeling novelties. First, we explicitly consider PP routes as thin routes, which brings more realism to the model. Second, we introduce route distance as an important element conditioning airlines' choices, which allows us to study airlines' optimal distribution of traffic for each distance range. Finally, we extend the model of Brueckner and Pai (2009) (which only applies to regional jet PP connections) to examine the effect of new low-cost PP routes.[2]

The econometric analysis in Brueckner and Pai (2009) focuses on routes that gained regional jet service in the period 1996-2005, using data of four US major carriers. Their results suggest that those routes would have likely gained jet service in the absence of the regional jet technology. In our paper, the econometric analysis focuses on the actual use of the three differ-

[1] Brueckner (2004) analyses the monopoly case and Flores-Fillol (2009) extends it to a duopoly. Barla and Constantatos (2005) examine the effect of capacity decisions under demand uncertainty on network structure.

[2] Regarding the provision of air services by low-cost carriers, the existing literature finds that entry is more likely to occur on dense routes (Bogulaski *et al.*, 2004; Gil-Moltó and Piga, 2008).

ent types of aircraft (turboprops, regional jets, and jets) by all major American and European network airlines in 2009. Along with the factors explored in Brueckner and Pai (2009), our multivariate econometric analysis also considers route competition and an additional proxy for the proportion of business passengers. Furthermore, we examine the use of low-cost subsidiaries by European network airlines.

In contrast to the previous literature, we show both theoretically and empirically that the "new routes hypothesis" is corroborated for some distance range when either regional jet or low-cost connections become available to network airlines. More precisely, our theoretical model predicts that a network airline may find it profitable to offer services on thin PP routes with regional jets for sufficiently short distances. This service would be aimed at business travelers, since the smaller size of regional jet aircraft may allow airlines to increase service quality (i.e., flight frequency) at higher fares. Additionally, a network airline may find it profitable to provide flights on thin PP routes with a low-cost subsidiary for longer distances to serve leisure travelers who are more fare-sensitive. Our empirical application uses data from the US and the EU to show that regional jets are widely used on thin and short PP routes with a high proportion of business travelers. Finally, we also find that European network airlines tend to use low-cost subsidiaries on thin and relatively long PP routes with a high proportion of leisure travelers.

Therefore, our analysis suggests that network airlines may have incentives to divert traffic away from their hubs by making use of either regional aircraft or low-cost subsidiaries. This phenomenon can act as a brake on the *hubbing* network strategy followed by major airlines since the deregulation of the sector, and it has important implications at the regional level.

The plan of the paper is as follows. A theoretical model analyzing the optimal traffic division in a simple network is presented in Section 2. Section 3 uses data of selected carriers to illustrate some of the theoretical results. Finally, a brief conclusion closes the paper. Some details on the theoretical analysis are presented in Appendices A and C. All the proofs are provided in Appendix B.

2 The model

We borrow most of the theoretical analysis from Brueckner and Pai (2009), but we also introduce some important modeling novelties (as explained before). Thus, the analysis that follows stresses the differences with respect to Brueckner and Pai (2009) and omits redundant material that can be found in that paper.

We assume the simplest possible network with three cities (A, B and H) and three city-pair markets (AH, BH and AB) as shown in Fig. 1.

–Insert Fig. 1 here–

AH and BH are "local" markets, which are always served nonstop, and market AB can be served either directly or indirectly with a one-stop trip via hub H, depending on the airline's network choice. The distance of routes AH and BH is assumed to be constant and equal to 1, whereas the distance of route AB is given by d, with $d \in (0, 2]$. The magnitude of d is an important factor influencing the airline's network choice.

Let the utility consumer i gets from air travel be $u_i^{air} = z_i - p_i - \gamma_i/f - \xi\mu_i$, where z_i denotes income (which includes travel benefit), p_i is the airline's fare, γ_i/f captures schedule delay with γ_i being the schedule delay disutility and f flight frequency, and μ_i stands for layover time (connecting passengers dislike waiting), where $\xi = \{0, 1\}$ with $\xi = 0$ for direct flights and $\xi = 1$ for connecting flights. As in Brueckner and Pai (2009), we assume two types of consumers, $i = H, L$, where H-types and L-types have the characteristics of business and leisure travelers, respectively. With respect to the L-types, the H-types have higher income, higher layover-time disutility, and a stronger aversion to schedule delay, i.e., $z_H > z_L$, $\mu_H > \mu_L$, and $\gamma_H > \gamma_L$. We assume a perfectly discriminating monopolist that charges fares to AB passengers depending on their type and routing. Normalizing the utility of the outside option to 0, surplus extraction implies $p_i = z_i - \gamma_i/f - \xi\mu_i$.

Passenger population size in market AB is normalized to unity, whereas $N > 1$ denotes population in markets AH and BH. In such a way, we explicitly consider PP routes as thin routes. In market AB, we assume that there is a share δ of type-H passengers and a share $1-\delta$ of type-L passengers. Further, the shares of H-types and L-types flying direct are θ_H and θ_L, respectively. Therefore, direct traffic on route AB and the connecting traffic on routes AH and BH are given by $q^d = \delta\theta_H + (1-\delta)\theta_L$ and $q^c = N + 1 - q^d$.

Turning our attention to local passengers in markets AH and BH, we assume that there is a share λ of type-H passengers and a share $1 - \lambda$ of type-L passengers, and thus $\widetilde{z} = \lambda z_H + (1-\lambda) z_L$ and $\widetilde{\gamma} = \lambda\gamma_H + (1-\lambda)\gamma_L$. Finally, the number of flight departures on route AB is given by $f^d = q^d/n^d$, where n^d is the *number of passengers per flight* on route AB. Both aircraft size and load factor determine the number of passengers per flight, which is given by $n^d = l^d s^d$, where s^d stands for aircraft size and $l^d \in [0, 1]$ for load factor.[3] Equivalently, flight frequency on routes AH and BH is $f^c = q^c/n^c$, with $n^c = l^c s^c$.

As a consequence of all these modeling assumptions, airline's revenue is given by

[3] We extend the approach in the existing literature, which typically assumes 100% load factor (see Brueckner, 2004; Flores-Fillol, 2009; Brueckner and Pai, 2009; Flores-Fillol, 2010; and Bilotkach *et al.*, 2010).

$$R = \underbrace{2N\left(\widetilde{z}-\frac{\widetilde{\gamma}n^c}{q^c}\right)}_{local} + \underbrace{\theta_H\delta\left(z_H-\frac{\gamma_H n^d}{q^d}\right)}_{direct\ H\text{-}types} + \underbrace{\theta_L(1-\delta)\left(z_L-\frac{\gamma_L n^d}{q^d}\right)}_{direct\ L\text{-}types} + \tag{1}$$

$$+\underbrace{(1-\theta_H)\delta\left(z_H-\mu_H-\frac{\gamma_H n^c}{q^c}\right)}_{connecting\ H\text{-}types} + \underbrace{(1-\theta_L)(1-\delta)\left(z_L-\mu_L-\frac{\gamma_L n^c}{q^c}\right)}_{connecting\ L\text{-}types}.$$

Similarly to Bilotkach *et al.* (2010), a *flight's operating cost* on route AB is given by $\omega(d)+\tau^d n^d$, where the parameter τ^d is the marginal cost per seat of serving the passenger on the ground and in the air, and the function $\omega(d)$ stands for the cost of frequency (or cost per departure), which captures the aircraft fixed cost (including landing and navigation fees, renting gates, airport maintenance and the cost of fuel). The function $\omega(d)$ is assumed to be continuously differentiable with respect to $d>0$ with $\omega'(d)>0$ because fuel consumption increases with distance. Note that cost per passenger, which can be written $\omega(d)/n^d+\tau^d$, visibly decreases with n^d capturing the presence of economies of traffic density (i.e., economies from serving a larger number of passengers on a certain route), the existence of which is beyond dispute in the airline industry.[4] Further, to generate determinate results, $\omega(d)$ is assumed to be linear, i.e., $\omega(d)=\omega d$ with a positive marginal cost per departure $\omega>0$.[5] Therefore, the airline's total cost from operating on route AB is $C^d=f^d\left[\omega d+\tau n^d\right]$ and, using $f^d=q^d/n^d$, we obtain $C^d=q^d\left(\frac{\omega d}{n^d}+\tau^d\right)$. Proceeding analogously for routes AH and BH, we obtain $C^c=q^c\left(\frac{\omega}{n^c}+\tau^c\right)$ since distance of routes AH and BH is assumed to be constant and equal to 1. Therefore, the *airline's total cost* from operating all routes is

$$C = 2\underbrace{q^c\left(\frac{\omega}{n^c}+\tau^c\right)}_{C^c} + \underbrace{q^d\left(\frac{\omega d}{n^d}+\tau^d\right)}_{C^d}. \tag{2}$$

Quite naturally, as d increases and the triangle in Fig. 1 flattens, direct connections become less profitable. The airline's objective is to maximize profits, which are given by $\pi=R-C$.

As in Brueckner and Pai (2009), we assume that airline's only choice variables are θ_H and θ_L, i.e., the division of H-type and L-type traffic between direct and connecting service (note that q^c and q^d depend on θ_H and θ_L). On the one hand, we observe that $\pi(\theta_H,\theta_L)$ is a strictly convex function of θ_H for γ_H sufficiently large with respect to γ_L,[6] so that the optimal θ_H is a

[4] See Caves *et al.* (1984), Brueckner and Spiller (1994), and Berry *et al.* (2006).

[5] Since fuel consumption is higher during landing and take off operations, $\omega''(d)<0$ might be a natural assumption. Assuming a concave function of the type $\omega(d)=\omega d^r$ with $r\in(0,1)$ would have no qualitative effect on our results; the critical distances that will be computed would simply need to be raised to the power $1/r$. Swan and Adler (2006) study the linearity of airlines' costs with respect to distance.

[6] As in Brueckner and Pai (2009), strict convexity requires $\gamma_H>2\widetilde{\gamma}$ or, equivalently, $\gamma_H(1-2\lambda)>2\gamma_L(1-\lambda)$. This condition requires γ_H sufficiently large with respect to γ_L and $\lambda<1/2$, i.e., there are more L-types than H-types among local passengers. Computations are available upon request.

corner solution, equal to either 0 or 1. On the other hand, it can be checked that $\pi(\theta_H, \theta_L)$ is a strictly concave function of θ_L, meaning that the optimal θ_L lies in the interval $[0, 1]$.

Starting from a situation in which the airline operates a hub-and-spoke network (i.e., AB passengers make a one-stop trip via hub H and $q^d = 0$), in the two following subsections we will consider other simple divisions of traffic between direct and connecting traffic when either a regional jet (RJ) or a low-cost (LC) direct connection between A and B is established by the network airline. Even though the AB market is relatively thin (as compared to local markets, which are denser), the network airline may be interested in sending either H-types or L-types direct (or both). The result $(\theta_H, \theta_L) = (0, 0)$ represents a hub-and-spoke (HS) network, and $(1, 1)$ denotes a fully-connected (FC) network. Finally, passenger segmentation occurs when only one type of passengers flies direct: $(1, 0)$ occurs when only H-types fly direct, and $(0, 1)$ occurs when only L-types fly direct.

2.1 The emergence of a RJ technology

The RJ technology is characterized by a lower aircraft size and a higher marginal cost per seat. Let us consider a network airline that operates in a HS manner (i.e., there is no direct service between A and B). In this situation, we study the emergence of a new direct service on route AB to carry type-H passengers when a RJ technology becomes available. This seems a natural airline reaction, since the lower aircraft size implies a higher flight frequency (because $f^d = q^d/n^d$, with $n^d = l^d s^d$) and H-types are more sensitive to schedule delay. Therefore, we assume $\tau^d > \tau^c$ (since cost per passenger in regional connections is higher) and $s^d < s^c$ (since regional aircraft are smaller). The latter assumption implies $n^d < n^c$, supposing that load factor remains the same on the three routes of the network (i.e., $l^d = l^c$). Hence, as pointed out in Brueckner and Pai (2009), for the outcome $(\theta_H, \theta_L) = (1, 0)$ to be optimal, the following conditions need to be met

$$\Omega \equiv \frac{\partial \pi(1,0)}{\partial \theta_L} < 0, \tag{3}$$

$$\Phi \equiv \pi(1,0) - \pi(0,0) > 0, \tag{4}$$

$$\Lambda \equiv \frac{\partial \pi(0,0)}{\partial \theta_L} < 0, \tag{5}$$

where Eqs. (3) and (4) ensure that there is no incentive to either increase θ_L or reduce θ_H (remember that $\theta_H = \{0, 1\}$), and Eq. (5) is needed to rule out $\pi(1, 0) < \pi(0, \theta_L)$ for $\theta_L \in [0, 1]$.

Carrying out the needed computations, Eqs. (3), (4), and (5) yield three expressions that depend on the parameters of the model (expressions provided in Appendix A). In a situation in which the airline operates a HS network, then $\theta_H^* = \theta_L^* = 0$ and $\Omega, \Phi, \Lambda < 0$ are satisfied. As in Brueckner and Pai (2009), the airline will send H-types direct (i.e., $(\theta_H^*, \theta_L^*) = (1, 0)$ becomes the equilibrium) when Φ reverses its sign from negative to positive. This will occur when $-\delta\Delta\tau^d - \frac{\delta\omega d}{\Delta n^d} - \gamma_H \Delta n^d > 0$, where the first and the second terms have a negative impact, whereas the third term has a positive effect since $\Delta n^d = n^d - n^c < 0$ and $\Delta\tau^d = \tau^d - \tau^c > 0$. On the one hand, a higher cost associated to route AB and a longer distance between cities A and B make the emergence of a direct connection more difficult. On the other hand, type-H passengers' aversion to schedule delay makes a new direct connection easier.

2.2 The emergence of a LC business model

Compared to the standard HS business model (using mainline jets), the LC business model is characterized by a higher load factor and a lower marginal cost per seat. As before, let us consider a network airline that initially operates a HS network (i.e., there is no direct service between A and B). In this situation, we consider that the network airline can set up a subsidiary LC carrier to provide a direct service on route AB to carry type-L passengers. The airline may be interested in establishing this connection because the higher load factor implies a lower flight frequency and thus a lower fare (because $p_L^d = z_L - \gamma_L/f^d$) and L-types are less sensitive to schedule delay and more fare-sensitive. Therefore, we assume $\tau^d < \tau^c$ (since cost per passenger in LC connections is lower) and $l^d > l^c$ (since load factor in LC air services is higher). The latter assumption implies $n^d > n^c$, supposing that the airline uses similar mainline jets on all routes (i.e., $s^d = s^c$). Although these two considerations are favorable to the adoption of a LC business model, there is still a trade-off since setting up a new direct connection implies a new cost element, as shown in Eq. (2). For the outcome $(\theta_H, \theta_L) = (0, 1)$ to be optimal, the following conditions need to be observed

$$\Psi \equiv -\frac{\partial \pi(0,1)}{\partial \theta_L} < 0, \tag{6}$$

$$\Gamma \equiv \pi(1,1) - \pi(0,1) < 0, \tag{7}$$

$$\Upsilon \equiv -\frac{\partial \pi(1,1)}{\partial \theta_L} < 0, \tag{8}$$

where Eqs. (6) and (7) ensure that there is no incentive either to decrease θ_L or to raise θ_H (remember that $\theta_H = \{0, 1\}$), and Eq. (8) is needed to rule out $\pi(0,1) < \pi(1, \theta_L)$ for $\theta_L \in [0, 1]$.

Carrying out the needed computations, Eqs. (6), (7), and (8) yield three expressions that depend on the parameters of the model (expressions provided in Appendix A). At this point, we analyze the emergence of a direct LC connection to serve L-type passengers starting from a situation in which the airline operates a HS network where $(\theta_H^*, \theta_L^*) = (0, \theta_L^*)$ with $\theta_L^* \in [0, 1)$, so that all H-types and at least some L-types fly connecting, where θ_L^* approaches 0 as the distance between A and B increases.[7] To sustain this distribution of passengers, we need

[7] Since $\pi(\theta_H, \theta_L)$ is a strictly concave function of θ_L, although the result $\theta_L^* = 0$ is a possibility, the only statement that can be made is that $\theta_L^* \in [0, 1)$.

to observe $\Psi, \Upsilon > 0$, so that $\theta_L = 1$ is not optimal, meaning that (at least) some L-types travel connecting through the hub. Concerning H-types, the airline will send them connecting when $\Sigma \equiv \pi(1,\theta_L) - \pi(0,\theta_L) < 0$ with $\theta_L \in [0,1]$. Note that Γ is a particular case of Σ with $\theta_L = 1$ (the expression for Σ is given in Appendix B) and thus $\Sigma < 0$ implies $\Gamma < 0$.[8] Therefore, $\Psi, \Upsilon > 0$ and $\Sigma < 0$ are assumed to hold. We can define $\Delta n^d = n^d - n^c > 0$ and $\Delta\tau^d = \tau^d - \tau^c < 0$ since the LC connection is characterized by a higher load-factor aircraft with a lower cost per passenger. In this situation, Ψ and Υ become negative for sufficiently important Δn^d and $\Delta\tau^d$ since they increase in τ^d and decrease in n^d; and Σ (and thus Γ) remains negative (i.e., H-types still fly connecting) as long as $-\Delta\tau^d - \frac{\omega d}{\Delta n^d} - \frac{\gamma_H - \gamma_L}{\delta + \theta_L(1-\delta)}\Delta n^d < 0$, where the first and the second terms have a positive impact, whereas the third term has a negative effect. The interpretation of this expression is similar as in the RJ case.

2.3 The effect of distance

After studying the setting in which either a RJ or a LC direct connection may arise, our attention now shifts to the effect of distance between endpoints on PP routes because network airlines may use different aircraft and business models depending on the characteristics of each city-pair market (and route distance is an important element). We discern distance intervals in which a new PP connection can optimally arise, analyzing the differences between the two types of connection (either RJ or LC). This also provides us with some predictions to test in the econometric application in Section 3.

2.3.1 RJ technology

Focusing on the effect of distance, from $\Omega < 0$ and $\Lambda < 0$ we can derive two lower bounds, i.e., $d > d_\Omega$ and $d > d_\Lambda$. In the same way, from $\Phi > 0$, we obtain the upper bound $d < d_\Phi$ (these bounds are provided in Appendix B). Therefore, the following lemma can be stated.

Lemma 1 *Focusing on the effect of distance between endpoints A and B, for a sufficiently low n^d relative to n^c, the optimal division of passengers is*
i) $(\theta_H^*, \theta_L^*) = (1,0)$, *for* $d \in (\max\{d_\Omega, d_\Lambda, 0\}, d_\Phi)$, *and*
ii) $(\theta_H^*, \theta_L^*) = (0,0)$, *for* $d > d_\Phi$.

The condition requiring a sufficiently low n^d relative to n^c (i.e., RJs are sufficiently small as compared to mainline jets) ensures that $d_\Phi > \max\{d_\Omega, d_\Lambda\}$. Lemma 1($i$) suggests that the

[8] Although Σ is a more general expression than Φ (in the RJ case) and than Γ (in the LC case), we decided to present first Φ and Γ because these expressions are simpler (since they do not depend on θ_L), and Σ is not needed to derive the main results of the paper: the existence of equilibria of the type (1,0) and (0,1), i.e., equilibria with passenger segmentation.

network airline would segregate passengers for moderately short distances, by sending H-types direct and L-types connecting. Thus, a network airline may find it profitable to offer services on PP routes with RJs (for business travelers) for sufficiently short distances, since the smaller size of RJ aircraft may allow airlines to increase service quality (i.e., flight frequency) at higher fares. Naturally, as captured in Lemma 1(ii), sending passengers direct becomes less profitable as distance increases, and the airline operates in a HS manner for sufficiently long distances.

In addition, whenever $\max\{d_\Omega, d_\Lambda\} > 0$, it could happen that $d \in (0, \max\{d_\Omega, d_\Lambda\})$. In this case, both high and low types may fly direct, as captured in the following corollary.

Corollary 1 *When $d_\Omega > 0$ and $d \in (0, \min\{d_\Omega, d_\Sigma\})$, then the optimal division of passengers is $(\theta_H^*, \theta_L^*) = (1, \theta_L^*)$ with $\theta_L^* \in (0, 1]$.*

The condition $d < d_\Sigma$, which implies $\pi(1, \theta_L) > \pi(0, \theta_L)$ for $\theta_L \in [0, 1]$, ensures that all H-types still fly direct (the bound d_Σ is explained in Appendix B); and $d < d_\Omega$, which implies $\frac{\partial \pi(1,0)}{\partial \theta_L} > 0$, guarantees that the network airline sends (at least) some L-type passengers direct.[9]

Thus, the corollary above states that the airline would send all H-types and a certain number of L-types direct for short distances, because connecting becomes increasingly inefficient.

2.3.2 LC business model

Focusing on the effect of distance, from $\Gamma < 0$, $\Psi < 0$ and $\Upsilon < 0$, we can derive the lower bound $d > d_\Gamma$ and the upper bounds $d < d_\Psi$ and $d < d_\Upsilon$ (note that d_Γ, d_Ψ and d_Υ can be trivially computed and are provided in Appendix B). Therefore, the following lemma follows.

Lemma 2 *Focusing on the effect of distance between endpoints A and B, for a sufficiently high n^d relative to n^c, the optimal division of passengers is*
i) $(\theta_H^*, \theta_L^*) = (1, 1)$, for $d < d_\Gamma$, and
ii) $(\theta_H^*, \theta_L^*) = (0, 1)$, for $d \in (d_\Gamma, \min\{d_\Psi, d_\Upsilon\})$.

The condition requiring a sufficiently high n^d relative to n^c (i.e., the load factor in the LC flights on route AB is sufficiently high as compared to the load factor on routes AH and BH) ensures that $\min\{d_\Psi, d_\Upsilon\} > d_\Gamma$. When a LC business model is set up on route AB, Lemma 2(i) suggests that the airline would send all passengers direct for short distances. For longer distances, the network airline would segregate passengers sending only L-types direct, as captured in Lemma 2(ii). Naturally, as distance increases, sending passengers direct becomes less profitable and airlines end up adopting HS networks for sufficiently long distances, as captured in the following corollary.

[9] Note that the condition $d < d_\Lambda$ (which implies $\frac{\partial \pi(0,0)}{\partial \theta_L} > 0$) is no longer needed with $d < d_\Sigma$ (which implies $\pi(1, \theta_L) > \pi(0, \theta_L)$ for $\theta_L \in [0, 1]$).

Corollary 2 *When $d > \max\{d_\Psi, d_\Sigma\}$, then the optimal division of passengers is $(\theta_H^*, \theta_L^*) = (0, \theta_L^*)$ with $\theta_L^* \in [0,1)$.*

The condition $d > d_\Sigma$, which implies $\pi(1,\theta_L) < \pi(0,\theta_L)$ for $\theta_L \in [0,1]$, ensures that all H-types still fly connecting (the bound d_Σ is explained in Appendix B); and $d > d_\Psi$ implies $-\frac{\partial \pi(0,1)}{\partial \theta_L} > 0$, so that the airline sends (at least) some L-type passengers connecting.[10]

Therefore, the result in the corollary above states that, for sufficiently long distances, the network airline would send all H-types and a certain number of L-types connecting, adopting a HS network. Quite naturally, as distance increases, direct flights become less profitable.

2.4 Discussion

Considering an environment in which both a RJ technology may be available and a LC business model can be adopted by network airlines on thin routes, we can contemplate a numerical example where the previous results arise (since the solutions are complex). Given the stylized nature of the model, parameter choices are necessarily arbitrary and the analysis is not exhaustive. However, it reveals some interesting insights which are in line with the empirical evidence. Let $z_L = 10$, $\gamma_L = 0.1$, $\mu_L = 0.7$, $z_H = 30$, $\gamma_H = 1.5$ and $\mu_H = 9$, so that income, schedule-delay and connection disutilities are much higher for the H-types. Let $\delta = 0.5$, so that AB passengers are composed by both H and L-types in equal parts. However $\lambda = 0.45$ indicates that H-types are relatively scarce among local passengers (remember that a sufficient condition for strict convexity of $\pi(\theta_H, \theta_L)$ with respect to θ_H is $\lambda < 1/2$). Let $N = 1.3$ (remember that $N > 1$ is assumed), indicating that local spoke-to-hub markets (i.e., markets AH and BH) are normally denser than spoke-to-spoke markets (i.e., market AB). The marginal cost per departure is $\omega = 32$, which is substantially larger than the marginal cost per passenger on hub-to-spoke routes, which is given by $\tau^c = 4.2$. Logically, the condition $\tau^d_{LC} < \tau^c < \tau^d_{RJ}$ is observed, with $\tau^d_{LC} = 4$ and $\tau^d_{RJ} = 7$ (where subscripts denote the type of PP connection between endpoints A and B). Finally, the number of passengers per flight on routes AH and BH is given by $n^c = 9.3$, and the condition $n^d_{RJ} < n^c < n^d_{LC}$ is respected, with $n^d_{RJ} = 2.5$ and $n^d_{LC} = 12.8$, since RJ aircraft are smaller and the load factor is higher when a low cost business model is implemented. Given this parameter constellation, the optimal choice of θ_H and θ_L depends on the value of d, in a way made clear in Fig. 2 below

–Insert Fig. 2 here–

The critical values of d that determine the different relevant regions are $d_\Omega = 0.852$, $d_\Phi = 0.896$, $d_\Gamma = 1.027$ and $d_\Psi = 1.959$ (Appendix C explains why these are the critical values of

[10] Note that the condition $d > d_\Upsilon$ (which implies $-\frac{\partial \pi(1,1)}{\partial \theta_L} > 0$) is no longer needed with $d > d_\Sigma$ (which implies $\pi(1,\theta_L) < \pi(0,\theta_L)$ for $\theta_L \in [0,1]$).

d), and the equilibrium in network structure depends crucially on the type of PP connection adopted on route AB (either RJ or LC). With the parameter values chosen above, we can compute the profit obtained by the airline for different values of θ_H and θ_L. More precisely, we will consider the cases $\theta_H, \theta_L = \{0, 1\}$, i.e., assuming that the airline has to send all passengers of the same type through the same routing. This is not a strong assumption since, looking at Fig. 2 above, one can observe that the optimal values of θ_H and θ_L are either 0 or 1 in all cases except in the following two regions. First, the region $d < d_\Omega$ when a RJ model is adopted and $(\theta_H^*, \theta_L^*) = (1, \theta_L^*)$ with $\theta_L^* \in (0, 1]$, with $\theta_L^* \to 1$ as d decreases, so that a FC network arises for a sufficiently small distance between A and B. Second, the region $d > d_\Psi$ when a LC model is adopted and $(\theta_H^*, \theta_L^*) = (0, \theta_L^*)$ with $\theta_L^* \in [0, 1)$, with $\theta_L^* \to 0$ as d increases, so that a HS network arises for a sufficiently long distance between A and B. Table 1 below presents the value of $\pi(0, 0)$, $\pi(1, 0)$, $\pi(0, 1)$ and $\pi(1, 1)$ for some particular values of d in the different regions shown in Fig. 2. The values in Table 1 confirm the results shown in Fig. 2 above.[11]

−Insert Table 1 here−

As we can see, the choice of θ_H and θ_L gives rise to a certain network structure, where shorter distances between endpoints A and B support FC structures and higher levels of d favor HS network configurations. Interestingly, for $d \in (d_\Phi, d_\Gamma)$, the HS network is the outcome when a RJ technology is available and the FC network is the outcome when airlines implement a LC business model. As a consequence, we can conclude that adopting either a RJ model or a LC model on certain PP routes can significantly affect airlines' network structure.

Additionally, focusing on the cases in which there is passenger segmentation (i.e., $(\theta_H^*, \theta_L^*) = (1, 0)$ when a RJ model is adopted, and $(\theta_H^*, \theta_L^*) = (0, 1)$ when a LC model is adopted), we observe that $(1, 0)$ arises for shorter distances than $(0, 1)$. This result is also confirmed by the empirical evidence, as will be shown in the next section.

3 An empirical application

In this section, we conduct an empirical application of the issues developed in the theoretical model, using data corresponding to 2009. First, we explain the criterion for the selection of the sample of routes and describe the variables used in the empirical analysis. Then, we examine data and estimate equations to identify how route features (distance, competition, demand, proportion of business and leisure travelers) influence aircraft technology and business models.

[11] Note that when a RJ model is adopted in the region $d < d_\Omega$, then $\pi(1, 0) > \pi(1, 1)$ is possible for values of d close to d_Ω (the optimal result is $(\theta_H^*, \theta_L^*) = (1, \theta_L^*)$ with $\theta_L^* \in (0, 1]$). In addition, when a LC model is adopted in the region $d > d_\Psi$, then $\pi(0, 0) < \pi(0, 1)$ is possible for values of d close to d_Ψ (the optimal result is $(\theta_H^*, \theta_L^*) = (0, \theta_L^*)$ with $\theta_L^* \in [0, 1)$).

3.1 Data

Data on airline supply on each route both for the US and the EU (frequencies, type of aircraft and total number of seats) have been obtained from RDC aviation and data on distance of the route are from the Official Airlines Guide (OAG) and the webflyer web site.[12]

Our sample includes all routes with direct flights served within continental US by the six major American network carriers (American Airlines, Continental, Delta, Northwest, United Airlines, and US Airways) and their subsidiaries, and all routes with direct flights served within the EU (EU of 27 countries + Switzerland and Norway) by the four major network airlines (Air France/KLM, British Airways, Iberia, and Lufthansa) and their subsidiaries. Altogether, at the airline-route level, we have 5031 observations for US carriers and 1033 for EU airlines.[13]

We account for routes with different market structures, including monopoly and oligopoly routes. Monopoly routes are defined as those routes where the dominant airline has a market share larger than 90% in terms of total annual seats.

Regarding the type of aircraft, the most used turboprops in our sample are: ATR 42/72, British Aerospace ATP, De Havilland DHC-8, Embraer 120, Fairchild Dornier 328, Fokker 50, Saab 340/2000. The most used regional jets (RJs) are: Avro RJ 70/85/100, Bae 146, Canadian Regional Jet, Embraer RJ 135/140/145/270/175/190/195, Fokker 70/100. Finally, the most used mainline jets are: Airbus 318/319/320/321, Boeing 717/737/757, and MD 80/90.

Network airlines can provide regional services either directly or by means of a subsidiary or partner airline (see Forbes and Lederman, 2009). On routes where regional aircraft are dominant, we cannot determine whether the provision of services is undertaken by a regional carrier that is a subsidiary of the network airline, or by an independent regional carrier that has signed a contract with the network airline because our dataset always allocates these regional flights to the network carrier.

In addition to the type of aircraft being used, we are also interested in the business model implemented by the airline: either full-service or low-cost (LC) service. This analysis focuses on European airlines because the American network carriers did not have any LC subsidiaries in 2009.[14] Among the European airlines, we have Transavia (LC subsidiary of Air France/KLM), Vueling (LC subsidiary of Iberia), and Germanwings and Bmi Baby (LC subsidiaries of Lufthansa).[15] There are at least three reasons for this difference between the US and

[12] See *http://webflyer.com.*

[13] Since data for some explanatory variables are not available for the American carriers, the sample used in the regressions is reduced to 4895 observations. The Delta-Northwest merger was not completed until early 2010. Hence, we treat Delta and Northwest as separate airlines regarding their choice of aircraft. The Lufthansa-Austrian merger was not completed until 2010, while the Iberia-Vueling merger was completed in 2009.

[14] Ted was a LC subsidiary of United but it was diluted into the mainline brand in the beginning of 2009. Another LC subsidiary, Song, was folded into the Delta mainline brand in 2006.

[15] Some of the largest European LC carriers like Ryanair or Easyjet are not subsidiaries of network airlines

the EU. First, the national interests of the former flag carriers in Europe make them operate in non-hub national airports to prevent competition in their home markets. Second, Europe has a higher number of airports specialized in leisure traffic. Finally, it could be argued that LC carriers in the US have experienced a certain upmarket movement that bring them closer to network carriers. In this context, setting up a LC subsidiary can be inadvisable for American network carriers.[16]

We consider the following US hub airports: Dallas (DFW), New York (JFK), Miami (MIA) and Chicago (ORD) for American Airlines; Cleveland (CLE), Houston (IAH) and New York (EWR) for Continental; Atlanta (ATL), Cincinnatti (CVG), New York (JFK) and Salt Lake City (SLC) for Delta; Detroit (DTW), Memphis (MEM) and Minneapolis (MSP) for Northwest; Chicago (ORD), Denver (DEN), Los Angeles (LAX), San Francisco (SFO) and Washington Dulles (IAD) for United Airlines; and Charlotte (CLT), Philadephia (PHX) and Phoenix (PHX) for US Airways. We consider the following European hubs: Amsterdam (AMS) and Paris (CDG and ORY) for Air France; London (LHR) for British Airways; Madrid (MAD) for Iberia; and Frankfurt (FRA), Munich (MUC) and Zurich (ZRH) for Lufthansa.

Data on population and Gross Domestic Product per Capita (GDPC) of American endpoints refer to the Metropolitan Statistical Area (MSA) and the information has been obtained from the US census. Some routes located in Micropolitan Statistical Areas are excluded because of the difficulties in obtaining sound comparable data. In the case of the EU, these data refer to the NUTS 3 level (statistical unit used by Eurostat), provided by Cambridge Econometrics (European Regional Database publication). We are aware that MSAs and NUTS 3, as defined by Eurostat, are not strictly comparable. Hence, it is difficult to make joint estimations using the whole sample of routes that include airlines from both the US and the EU.

In the EU, airports located in the following islands are considered tourist destinations: the Balearic and Canary Islands (Spain), Sardinia and Sicily (Italy), Corsica (France), and many Greek islands,[17] and also the airports of Alicante (ALC), Faro (FAO), Malaga (AGP) and Nice (NCE). In the US, we consider as tourist destinations the airports of Las Vegas (LAS), Orlando (MCO), Grand Canyon (FLG), Spokane (GEG), Vail (EGE), and some coastal cities of Florida and California, which are the two most popular states for tourism. Some ski resorts airports (like Aspen) are not in our sample because they are located in Micropolitan Statistical Areas.

We have built an airport access variable that measures the distance between the airport

and, as a consequence, they are not included in our empirical analysis.

[16] Graham and Vowles (2006) and Morrell (2005) undertake a broad examination of the establishment of LC subsidiaries by network carriers, but fail to find indisputable evidence of the success of this strategy. In the US, it seems that the diffculties in effectively separating network operations from those of the LC subsidiary may lead to a cannibalization and dilution of the main brand. Furthermore, network carriers may find it difficult to differentiate the pay scales of employees due to union activism.

[17] Details available from the authors on request.

and the city center using Google Maps. In most cases, the identity of the relevant cities was self-evident. For airports located between cities, we calculated the distance from the airport to the closest city with more than $100,000$ inhabitants.

Fig. 3 below shows that regional aircraft are the type most used by the main American network carriers up to a route distance of 900 miles. In fact, US major airlines mainly serve PP routes in the distance range 300-900 miles with RJs, and RJs are still widely used on routes in the distance range 900-1200 miles. Turboprops are widely used on routes shorter than 300 miles. Mainline jets are obviously the dominant type of aircraft on routes longer than 1200 miles. The upshot of this exploratory examination of data is that the high number of PP routes in the distance range of 300-1200 (and particularly in the distance range 300-900 miles), may be related to the advantages that US network airlines have gained from using RJs.

–Insert Fig. 3 here–

Fig. 4 shows that RJs are the most used aircraft by the main European network carriers up to a route distance of 600 miles, especially the distance range 300-600 miles. Turboprops are also widely used on routes shorter than 300 miles. Interestingly, the use of mainline jets with a LC subsidiary is the dominant model on routes longer than 600 miles. Thus, these data provide some evidence that the relatively high number of PP routes in the distance range 300-600 miles has to do with the use of RJs. Furthermore, the viability of PP routes on routes longer than 600 miles seems to be associated (in many cases) with the use of LC subsidiaries.

–Insert Fig. 4 here–

3.2 The emergence of a RJ technology

To examine airlines' aircraft choices, we estimate the following equation for the network airline i offering services on route k

$$Type\_of\_aircraft_{ik} = \alpha + \beta_1 Distance_k + \beta_2 Population_k + \beta_3 Population_k^2 + \beta_4 GDPC_k + \\ +\beta_5 D_k^{tourism} + \beta_6 Dist\_to\_city\_center_k + \beta_7 D_k^{monopoly} + \beta_8 D_{ik}^{hub} + \varepsilon_k. \tag{9}$$

Note that different types of aircraft may be used on the same route. Hence, we need to compute the market share of all aircraft used by airlines from the same category (turboprops, RJs or mainline jets) in terms of the total number of seats offered on the route. The dependent variable for the type of aircraft used is then constructed. This variable takes the value zero for routes where RJs have the largest market share (which will be the reference case); it takes the value one for routes where the turboprops have the largest market share, and it takes the value two for routes where mainline jets have the largest market share. Note that typically the market share of the category of aircraft that is dominant is well above 50%. We consider the following variables as exogenous explanatory variables of the type of aircraft used by airlines.

1. $Distance_k$: Number of kilometers in the case of European routes and number of miles in the case of American routes flown to link the endpoints of the route.

2. $Population_k$: Weighted average of population at the origin and destination regions of the route. We also include the square of the population as an explanatory variable because the effect of this variable is concentrated around the median values of its statistical distribution.[18]

3. $GDPC_k$: Weighted average of Gross Domestic Product per capita at the origin and destination regions of the route. Weights are based on population.

4. $D_k^{tourism}$: Dummy variable that takes the value one for routes in which at least one of the endpoints is a major tourist destination.

5. $Dist\_to\_city\_center_k$: The sum of the distances between the origin and the destination city center and the respective airports.

6. $D_k^{monopoly}$: Dummy variable that takes the value one on routes where one airline has a market share larger than 90% in terms of total annual seats.

7. D_{ik}^{hub}: Dummy variable that takes the value one on routes in which at least one of the endpoints is a hub airport.

We include airline fixed effects in the regression. We consider the airline with the highest number of observations as the reference: Delta for the US and Air France/KLM for the EU.

The cost superiority of mainline jets in relation to RJs increases with distance, while on very short-haul routes turboprops are less costly than RJs. Thus, as route distance increases, we can expect RJs to be used less than mainline jets and more than turboprops. The longer range of RJs with respect to turboprops yields a clear prediction on the expected effect of the distance variable. However, the expected results for the rest of explanatory variables in the choice of RJs in relation to turboprops are not clear a priori.

Demand should be higher in more populated and richer endpoints. Additionally, monopoly routes should generally be thinner than routes where several airlines offer air services. As compared to mainline jets, we expect RJs to be used more on both monopoly routes and thinner routes, i.e., routes with less populated endpoints.

Note that the $GDPC_k$ variable may capture two different effects. On the one hand, the proportion of business travelers should be higher in richer endpoints but, on the other hand, demand may also be higher.

Our analysis also tries to identify routes with a higher proportion of leisure travelers. These routes are the ones with a tourist destination as endpoint and the ones with airports further

[18] The same could be argued for the distance variable, but the square of distance is highly insignificant when we include it in the regressions. As a consequence, this variable is not considered.

away from the city center. The relatively higher frequency of RJs makes them particularly convenient for business travelers, so that we expect RJs (in relation to mainline jets) to be used less on tourist routes with a higher proportion of leisure travelers.

Finally the dummy variable for hub airports allows us to determine whether RJs are more likely to be used either to feed hubs or to provide services on PP routes. Recall that hub-to-spoke routes may be generally denser than spoke-to-spoke routes.

We estimate Eq. (9) using a multinomial logit, which is appropriate when the dependent variable is based on more than two discrete alternatives that do not have a natural ordering. In our context, we have three different alternatives: turboprops, RJs, or jets. The multinomial logit estimates the probability for an airline to choose one of these alternatives. To compute the probabilities for the different alternatives, it is needed to set one of them as a reference case. Then the reported results show separately the probability for each alternative to be chosen as compared to the one that is considered as the reference. Thus, the sign and the statistical significance of the explanatory variables may differ depending on the considered alternative.

In our empirical model, the use of RJs is assumed to be the reference case. Then we assign value zero to those observations where the use of RJs prevails. We assign value one to those observations where turboprops are mostly used, and value two to those observations where jets are dominant. Thus, we consider separately the choice of turboprops or mainline jets in relation to RJs. Given that the observations with RJs are assigned the value zero, we can infer that a higher value of the corresponding explanatory variable would mean that the use of RJs will be more (less) likely if the sign of the coefficient associated to this variable is negative (positive).

Tables 2 and 3 report separately the results of airlines' choice of turboprops in relation to RJs, and the choice of jets in relation to RJs. Thus, we may expect that the sign and the statistical significance of the coefficients associated to the explanatory variables differ in each different choice. Recall that we are not able to consider jointly the choices of American and European airlines, which in fact operate in very different markets because the statistical definition of urban areas is not comparable. Table 2 shows the coefficients estimated and their respective standard errors. Table 3 shows the predicted change in the probability for an outcome to take place (i.e., the use of RJs in relation either to turboprops or to mainline jets) as each independent variable changes from its minimum to its maximum value (i.e., from 0 to 1 for discrete variables) while all other independent variables are held constant at their mean values. The results in Table 2 report the statistical significance of the considered relationships, while the results in Table 3 report the quantitative impact of each explanatory variable.

−Insert Tables 2 and 3 here−

First, we compare the use of RJs as compared to mainline jets. Looking at the effect of route distance, RJs are used more on shorter routes, as expected. The impact of the distance variable is really important: the predicted increase in the probability of using mainline jets in relation to RJs as distance shifts from its minimum to its maximum value is about 95% in the case of American network airlines and 85% in the case of European network airlines.

Additionally, we find that RJs are more likely to be used on thinner routes than mainline jets. Our results show that mainline jets are used more than RJs on routes with more populated and richer endpoints (although the variable of GDP per capita is not statistically significant in the case of European airlines). In contrast, mainline jets are less used on monopoly routes. Recall that the $GDPC_k$ variable may capture two different effects. On the one hand, the proportion of business travelers should be higher in richer endpoints but, on the other hand, demand may also be higher. The predicted change in probabilities is quite high for all these variables and similar for US and EU network airlines. Only the effect of population on the predicted change in probabilities seems to be clearly higher in the case of European airlines.

Interestingly, RJs seem to be more used on routes with a higher proportion of business travelers. We make this conclusion in view of the fact that RJs are less used than mainline jets on tourist routes and on routes where airports are further from the city center. The predicted change in probabilities is also high for both variables. Note that the dummy variable for tourist destinations and the variable of airport distance from the city center seem to capture better the proportion of business travelers than the variable of GDP per capita.

Finally, European network airlines use RJs more on spoke-to-spoke routes (i.e., PP routes) than on hub-to-spoke routes. Although we do not find statistical differences between hub-to-spoke routes and spoke-to-spoke routes considering US network airlines as a whole, this result can be qualified by analyzing each carrier independently and focusing on airline-specific effects. Results from regressions for each airline show that these differences are generally related with the magnitude of the effect but not with its direction or its statistical significance. An important exception is the result of the dummy for hub-to-spoke routes (i.e., D_{ik}^{hub}) for US network airlines. Table 4 explores this effect, showing the results of this variable for each American network airline.[19] Table 4 suggests that several US network airlines use RJs more on spoke-to-spoke routes than on hub-to-spoke routes as is the case for European network airlines.

–Insert Table 4 here–

Shifting our attention to the analysis of the use of RJs with respect to turboprops, as expected, we can derive only one strong inference: turboprops are used more than RJs on shorter routes. The predicted decrease in the use turboprops with respect to RJs when distance shifts from its minimum to its maximum value is about 41% in the case of US network airlines and 60% in the case of European ones. From a statistical point of view, there are other significant variables such as the dummies for monopoly routes and tourist endpoints. However, their impact in terms of the change in the predicted probabilities is very small (almost zero).

[19] The full report of the estimates of airline specific regressions is available upon request from the authors.

Looking at our previous theoretical results, we observe that the result $(\theta_H^*, \theta_L^*) = (1, 0)$, i.e., only business passengers travel direct, is confirmed empirically. Our empirical results show that RJs are mostly used by business travelers for intermediate-distance routes, and are mostly used on PP routes (for EU carriers and several US carriers). Consequently, new direct connections may be related to the advent of a RJ technology. In terms of Brueckner and Pai (2009), the "new routes hypothesis" based on RJ direct connections seems plausible.

3.3 The emergence of a LC business model

Here we focus our attention on routes where mainline jets are used. Our interest is to examine when a network airline is more likely to choose to operate the route with a LC subsidiary instead of the main brand. Recall that this analysis focuses only on European network airlines. We estimate the following equation for an airline i offering services on route k

$$\begin{aligned} D^{LC\_subsidiary} &= \alpha + \beta_1 Distance_k + \beta_2 Population_k + \beta_3 GDPC_k + \beta_4 D_k^{tourism} + \\ &+ \beta_5 Dist\_to\_city\_center_k + \beta_6 D_k^{monopoly} + \beta_7 D_{ik}^{hub} + \varepsilon_k, \end{aligned} \tag{10}$$

where the dependent variable is dichotomous and takes the value one on routes where network airlines make use of a LC subsidiary. We use the same explanatory variables as in Eq. (9).[20]

A priori, it is not clear whether the LC subsidiary is used more than the main brand either on longer or on shorter routes. However, following the theoretical analysis, we would expect the LC subsidiary to be widely used on thin PP routes with a high proportion of leisure travelers and relatively long distances. Thus, we expect LC subsidiaries to be used more on spoke-to-spoke routes (than on hub-to-spoke routes), on monopoly routes, on routes with poorer and less populated endpoints, and on routes with a high proportion of leisure travelers, i.e., routes from/to tourist destinations and routes with airports further away from the city center.

The estimation of Eq. (10) is made using the logit technique. A higher value of the coefficient associated to an explanatory variable means that the LC subsidiary is more (less) likely to be used if the sign of this coefficient is positive (negative). Table 5 below shows the results.

–Insert Table 5 here–

The results above confirm our hypotheses. Indeed, all the coefficients are statistically significant and have the expected sign, except the one corresponding to the variable of the distance from the airport to the city center, which is not statistically significant. The impact in terms of change in the predicted probabilities is also high for all the significant variables.

[20] We exclude the observations of British Airways in the regression because this airline did not have a LC subsidiary in the period considered. Given the reduced number of observations in this regression, we consider that airline fixed effects are inappropriate. The low number of observations also advises against including the square of population as explanatory variable. In any case, this latter variable is highly insignificant when included in the regression.

Since the coefficient associated to the distance variable is positive and significant, we find evidence that the LC subsidiary is used more than the main brand on longer routes. For an airline, the predicted increase in the probability of using a LC subsidiary instead of the main brand as route distance shifts from its minimum to its maximum value is about 73%.

Furthermore, the LC subsidiary is used more on PP routes because the coefficient associated to the dummy variable for hub routes is negative and statistically significant. This result may be expected because network airlines concentrate connecting traffic in their hubs. The predicted decrease in the probability of using LC subsidiaries when routes involve a hub is about 76%.

The LC subsidiary is more likely to be used on monopoly routes and on routes with poorer and less populated endpoints. Thus, we conclude that LC subsidiaries are used more on thinner routes. The predicted change in the probability of using LC subsidiaries is notable for these variables.

Finally, it seems that the LC subsidiary is more likely to be used on routes with a high proportion of leisure travelers because the coefficient associated to the dummy for tourist routes is positive and statistically significant. The predicted increase in the probability of using LC subsidiaries when routes have a tourist major destination as an endpoint is about 24%.

These results corroborate our theoretical results, and the optimal passenger division $(\theta_H^*, \theta_L^*) = (0, 1)$, i.e., only leisure passengers travel direct, is confirmed. Therefore, LC subsidiaries are mostly used to carry leisure travelers on relatively long and thin PP routes. Consequently, new direct connections may be related to the emergence of this new business model.

4 Concluding remarks

Network airlines may benefit from concentrating operations in their hub airports through the exploitation of density economies and a higher connectivity. However, HS networks may have negative consequences, such as congestion, lower competition due to airport dominance (by the *hubbing* airline), and lower service quality for citizens living in cities far from hub airports.

This paper shows that network airlines may also have incentives to divert passengers away from the hub. Our main contribution is the analysis of the influence of two innovations, the RJ technology and the LC business model, in the provision of services on PP routes.

We find that the RJ technology and the LC business model are intensively used by network airlines on thin PP routes. On the one hand, a network airline finds it profitable to offer services on thin PP routes with RJs for sufficiently short distances. This direct connection is mostly addressed to business travelers, since the smaller size of RJ aircraft may allow network airlines to increase flight frequency. On the other hand, a network carrier could be interested in serving a thin PP route by means of a subsidiary LC carrier for sufficiently long distances. This direct connection will be used mainly by leisure travelers who are more fare-sensitive.

The research question raised in this paper is especially relevant, because setting up new RJ or LC direct connections may have very different implications in terms of network structure, fares and flight frequency. In addition, the regional impact of the different airline network configurations may also differ widely. Policy makers and airport operators should assess which type of airline networks they want to foster in their sphere of influence. If they wish to promote direct connections away from the hub, they should use tools such as airport charges (both the level and the relation with the weight of the aircraft), investment in capacities, and marketing of the cities where the airports are located.

References

[1] Barla, P., Constantatos, C., 2005, 'Strategic interactions and airline network morphology under demand uncertainty,' *European Economic Review*, 49, pp. 703-716.

[2] Berry, S., Carnall, M., Spiller, P., 2006, 'Airline hubs: costs, markups and the implications of customer heterogeneity,' In: Lee, D. (Ed.), *Advances in Airline Economics*, vol. 1, Elsevier, Amsterdam, pp. 183-214.

[3] Bilotkach, V., Fageda, X., Flores-Fillol, R., 2010, 'Scheduled service versus private transportation: the role of distance,' *Regional Science and Urban Economics*, 40, pp. 60-72.

[4] Bogulaski, C., Ito, H., Lee, D., 2004, 'Entry patterns in the Southwest Airlines route system,' *Review of Industrial Organization*, 25, pp. 317-350.

[5] Brueckner, J.K., 2004, 'Network structure and airline scheduling,' *Journal of Industrial Economics*, 52, pp. 291-312.

[6] Brueckner, J.K., Pai, V., 2009, 'Technological innovation in the airline industry: the impact of regional jets,' *International Journal of Industrial Organization*, 27, pp. 110-120.

[7] Brueckner, J.K., Spiller, P.T., 1994, 'Economies of traffic density in the deregulated airline industry,' *Journal of Law and Economics*, 37, pp. 379-415.

[8] Caves, D.W., Christensen, L.R., Tretheway, M.W., 1984, 'Economies of density versus economies of scale: why trunk and local service airline costs differ,' *RAND Journal of Economics*, 15, pp. 471-489.

[9] Dresner, M., Windle, R., Zhou, M., 2002, 'Regional jet services: supply and demand,' *Journal of Air Transport Management*, 8, pp. 267-273.

[10] Flores-Fillol, R., 2009, 'Airline competition and network structure,' *Transportation Research Part B*, 43, pp. 966-983.

[11] Flores-Fillol, R., 2010, 'Congested hubs,' *Transportation Research Part B*, 44, pp. 358-370.

[12] Forbes, S.J., Lederman, M., 2009, 'Adaptation and vertical integration in the airline industry,' *American Economic Review*, 99, pp. 1831-1849.

[13] Gil-Moltó, M.J., Piga, C., 2008, 'Entry and exit by European low cost and traditional carriers,' *Tourism Economics*, 14, pp. 577-598.

[14] Graham, B., Vowles, T.M., 2006, 'Carriers within carriers: a strategic response to low-cost airline competition,' *Transport Reviews*, 26, pp. 105-126.

[15] Morrell, P., 2005, 'Airline within airlines: an analysis of US network airline responses to low cost carriers,' *Journal of Air Transport Management*, 11, pp. 303-312.

[16] Swan, W, Adler, N, 2006, 'Aircraft trip cost parameters: a function of stage length and seat capacity,' *Transportation Research Part E*, 42, pp. 105-115.

Figures and Tables

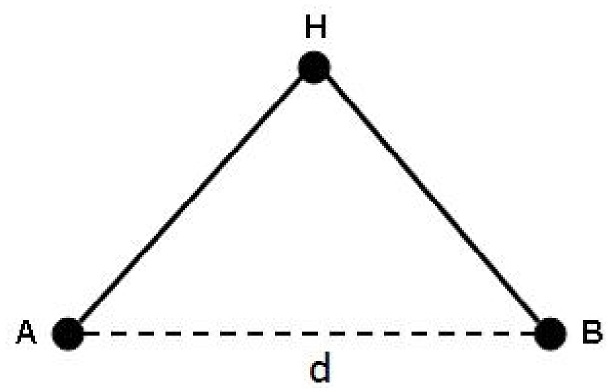

Fig. 1: Network

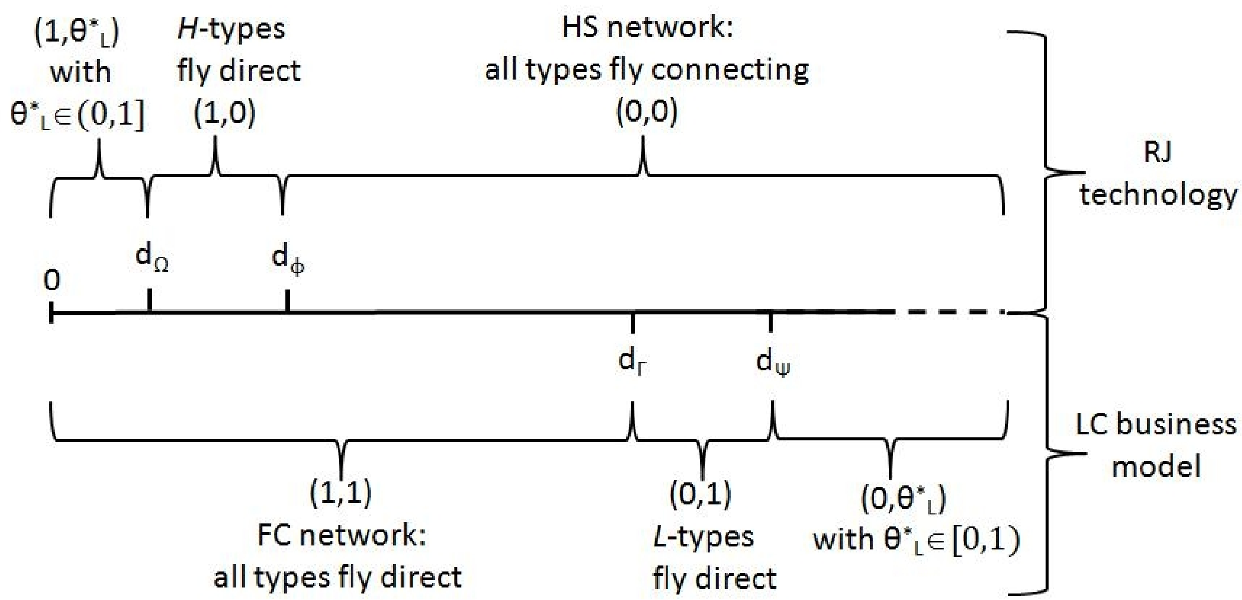

Fig. 2: Optimal network choice

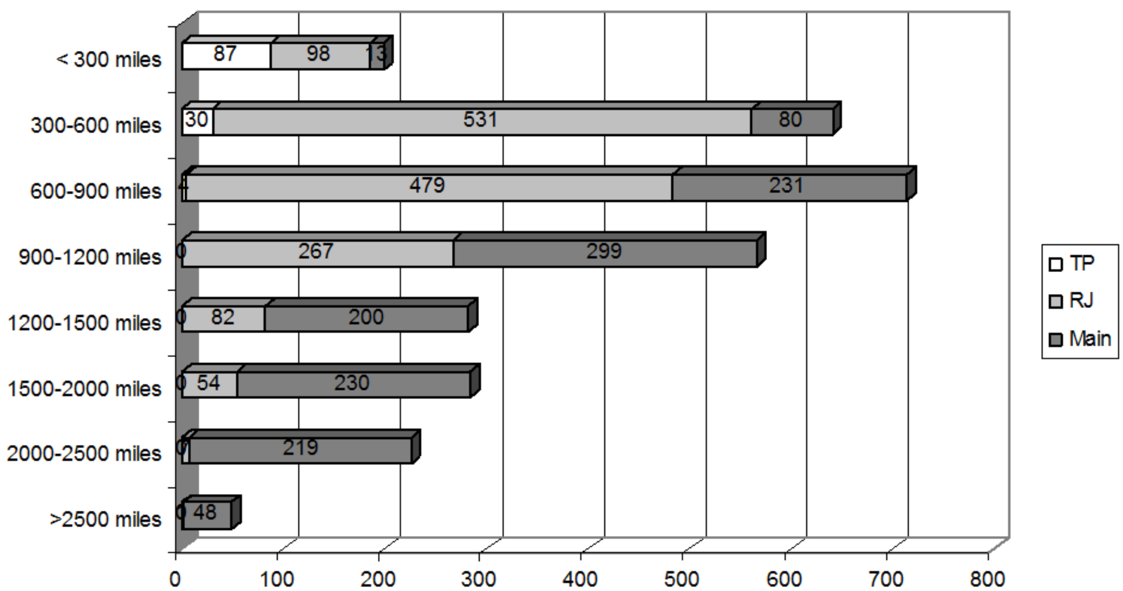

Fig. 3: Aircraft technology by distance (PP routes - US)

Note 1: Data refer to the number of routes where each considered type of aircraft is dominant.

Note 2: TP are turboprops, RJ are regional jets, and Main are mainline jets.

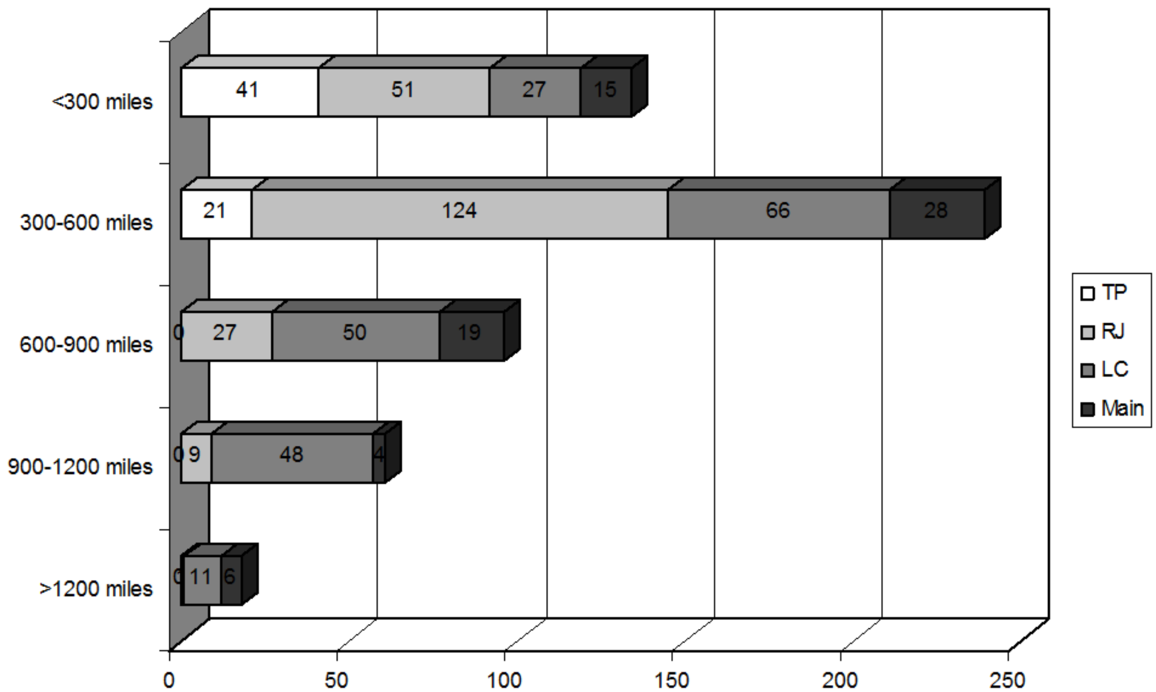

Fig. 4: Aircraft technology and business model by distance (PP routes - EU)

Note 1: Data refer to the number of routes where each considered type of aircraft and business model is dominant.

Note 2: TP are turboprops, RJ are regional jets, LC are mainline jets with a low-cost subsidiary, and Main are mainline jets with the main brand.

Table 1: Example of network choice when RJ and LC models are available on route AB

	$d = 0.20$ ($d < d_\Omega$)		$d = 0.87$ ($d \in (d_\Omega, d_\Phi)$)		$d = 0.95$ ($d \in (d_\Phi, d_\Gamma)$)		$d = 1.50$ ($d \in (d_\Gamma, d_\Psi)$)		$d = 1.98$ ($d > d_\Psi$)	
	RJ	LC	RJ	LC	RJ	LC	RJ	LC	RJ	LC
$\pi(0,0)$	18.49	18.49	18.49	18.49	<u>18.49</u>	18.49	<u>18.49</u>	18.49	<u>18.49</u>	<u>18.49</u>
$\pi(1,0)$	22.93	10.03	<u>18.66</u>	9.19	18.15	9.09	14.63	8.40	11.56	7.80
$\pi(0,1)$	18.68	20.18	14.39	19.34	13.88	19.24	10.36	<u>18.56</u>	7.29	17.96
$\pi(1,1)$	<u>24.40</u>	<u>21.22</u>	15.82	<u>19.54</u>	14.80	<u>19.34</u>	7.76	17.97	1.61	16.77

Note: d_Ω= 0.852, d_Φ= 0.896, d_Γ= 1.027 and d_Ψ= 1.959.

Table 2: Results of estimates of the aircraft choice (mlogit) - US sample

	US sample ($N = 4895$)		EU sample ($N = 1033$)	
	Dependent variable: RJ=0, turboprop=1	Dependent variable: RJ=0, mainline jet=2	Dependent variable: RJ=0, turboprop=1	Dependent variable: RJ=0, mainline jet=2
$Distance_k$	-0.0099 (0.0006)***	0.0025 (0.00009)***	-0.006 (0.0006)***	0.0015 (0.00017)***
$Population_k$	-3.35e-07 (9.30e-08)***	9.39e-08 (3.67e-08)***	0.00023 (0.00020)	0.00037 (0.00013)***
$Population_k^2$	2.05e14 (4.63e-15)***	-6.62e-15 (1.88e-15)***	-1.44e-08 (1.84e-08)	-2.63e-08 (1.09e-08)**
$GDPC_k$	0.000014 (0.00002)	0.000026 (0.00001)**	0.003 (0.005)	0.002 (0.002)
$D_k^{tourism}$	1.44 (0.30)***	1.35 (0.12)***	0.86 (0.37)**	0.92 (0.26)***
$Dist\_to\_city\_center_k$	-0.04 (0.017)**	0.009 (0.003)***	-0.011 (0011)	0.015 (0.005)***
$D_k^{monopoly}$	1.98 (0.30)***	-1.15 (0.08)***	0.91 (0.32)***	-0.81 (0.16)***
D_{ik}^{hub}	-0.31 (0.21)	0.14 (0.10)	-0.064 (0.37)	0.38 (0.18)**
$D_{American}$	1.37 (0.52)***	1.79 (0.12)***	–	–
$D_{Continental}$	3.02 (0.38)***	0.25 (0.19)	–	–
$D_{Northwest}$	1.65 (0.37)***	−0.17 (0.12)***	–	–
D_{United}	3.77 (0.42)***	0.23 (0.14)	–	–
$D_{US\ Airways}$	1.69 (0.37)***	0.03 (0.11)	–	–
$D_{British\ Airways}$	–	–	0.21 (0.96)	0.96 (0.40)**
$D_{Lufthansa}$	–	–	-0.35 (0.35)	0.78 (0.20)**
D_{Iberia}	–	–	-0.45 (0.38)	-0.12 (0.24)
$Constant$	-1.04 (0.80)	-3.63 (0.36)***	0.69 (0.97)	-2.84 (0.61)***
R^2	0.42		0.25	
$F(joint\ sig.)$	1725.65***		322.99***	

Note 1: Standard errors in parenthesis (robust to heteroscedasticity).
Note 2: Statistical significance at 1% (***), 5% (**), 10% (*).

Table 3: Change in the predicted probabilities

	US sample ($N = 4895$)		EU sample ($N = 1033$)	
	Dependent variable: RJ=0, turboprop=1	Dependent variable: RJ=0, mainline jet=2	Dependent variable: RJ=0, turboprop=1	Dependent variable: RJ=0, mainline jet=2
$Distance_k$	-44.76%	95.94%	-60.52%	84.61%
$Population_k$	-0.045%	41.43%	0.26%	66.71%
$GDPC_k$	0.0010%	16.48%	0.32%	13.17%
$D_k^{tourism}$	0.007%	32.32%	0.23%	19.17%
$Dist\_to\_city\_center_k$	-0.019%	21.43%	1.73%	37.34%
$D_k^{monopoly}$	0.022%	-27.61%	1.15%	-19.38%
D_{ik}^{hub}	-0.0029%	3.47%	0.31%	9.04%

Table 4: Results from regressions for the variable D_{ik}^{hub} - US sample

	Dependent variable: RJ=0, mainline jet=2	
	Coefficient	Change in the predicted probabilities
Delta ($N = 1214$)	0.30 (0.17)*	6.67%
American ($N = 808$)	-0.15 (0.36)	-0.08%
Continental ($N = 268$)	1.41 (0.68)**	13.78%
Northwest ($N = 1085$)	0.89 (0.20)***	16.35%
United ($N = 528$)	-4.51 (0.85)***	-58.88%
US Airways ($N = 992$)	1.07 (0.24)***	25.57%

Note 1: Standard errors in parenthesis (robust to heteroscedasticity).
Note 2: Statistical significance at 1% (***), 5% (**), 10% (*).

Table 5: Results of estimates of the business model (logit) – EU routes with mainline jets ($N = 493$)

	Dependent variable: main brand=0, LC subsidiary=1	
	Coefficient	Change in the predicted probabilities
$Distance_k$	0.0013 (0.00029)***	72.67%
$Population_k$	-0.00017 (0.00005)***	-43.71%
$GDPC_k$	-0.013 (0.0046)***	-59.77%
$D_k^{tourism}$	1.001 (0.45)**	24.08%
$Dist\_to\_city\_center_k$	-0.005 (0.009)	-12.09%
$D_k^{monopoly}$	2.27 (0.37)***	51.08%
D_{ik}^{hub}	-4.02 (0.42)***	-76.36%
$Constant$	2.40 (0.87)***	–
R^2	0.58	
$F(joint\ sig.)$	114.40***	

Note 1: Standard errors in parenthesis (robust to heteroscedasticity).

Note 2: Statistical significance at 1% (***), 5% (**), 10% (*).

A Appendix: Details on Subsections 2.1 and 2.2

Carrying out the needed computations, Eq. (3) becomes

$$\Omega \equiv (1-\delta)\left[\mu_L + 2\tau^c - \tau^d + \omega\left(\frac{2}{n^c} - \frac{d}{n^d}\right) + n^d\frac{\gamma_H - \gamma_L}{\delta} - Nn^c\frac{2\widetilde{\gamma} - \gamma_L}{(N+1-\delta)^2}\right], \tag{A1}$$

which shows the gains and losses for the network airline from increasing θ_L (i.e., sending more L-types direct). On the one hand, the airline saves the connecting discount to compensate for layover time disutility (μ_L) and the costs corresponding to routes AH and BH: the passenger cost ($2\tau^c$) and the cost of frequency ($\frac{2\omega}{n^c}$). Note that the cost of frequency decreases in s^c (since $n^c = l^c s^c$) because there is a negative relationship between flight frequency and aircraft size. On the other hand, it incurs the costs associated to the new direct service on route AB: the passenger cost (τ^d) and the cost of frequency ($\frac{\omega d}{n^d}$), which increases with distance since longer routes are more costly to serve. The two last terms capture the gain of sending more L-types direct as aircraft size is larger on route AB and smaller on routes AH and BH. Thus, there is an advantage associated to larger aircraft, which implies lower flight frequency and lower fares, since L-types are fare-sensitive.

Equivalently, Eq. (4) reduces to

$$\Phi \equiv \delta\left[\mu_H + 2\tau^c - \tau^d + \omega\left(\frac{2}{n^c} - \frac{d}{n^d}\right) - n^d\frac{\gamma_H}{\delta} + n^c\frac{(1-\delta)(\gamma_H - \gamma_L) + N(\gamma_H - 2\widetilde{\gamma})}{(1+N)(1+N-\delta)}\right], \tag{A2}$$

which indicates that the gain from sending all the H-types direct increases with their layover time disutility (μ_H) and with the costs corresponding to routes AH and BH ($2\tau^c + \frac{2\omega}{n^c}$). In contrast, the network airline incurs the costs associated to the new direct service on route AB ($\tau^d + \frac{\omega d}{n^d}$). The negative effect $n^d\frac{\gamma_H}{\delta}$ shows that the benefit from shifting all the H-types to direct service decreases with aircraft size and thus increases with frequency, capturing the advantage in terms of schedule delay stemming from a higher flight frequency and a smaller aircraft size. The last positive term, which increases with n^c and thus decreasing with f^c, captures the fact that sending all the H-types direct is more beneficial if the service quality (i.e. flight frequency) of the connecting service is poor.

Eq. (5) yields this condition

$$\Lambda \equiv (1-\delta)\left[\mu_L + 2\tau^c - \tau^d + \omega\left(\frac{2}{n^c} - \frac{d}{n^d}\right) - n^c\frac{\delta(\gamma_H - \gamma_L) - N(2\widetilde{\gamma} - \gamma_L)}{(1+N)^2}\right], \tag{A3}$$

which has a similar interpretation as Eq. (A1), except for the last term that has a more complex intuitive explanation.

Carrying out the necessary computations, Eq. (6) becomes

$$\Psi \equiv (1-\delta)\left[-\mu_L - 2\tau^c + \tau^d - \omega\left(\frac{2}{n^c} - \frac{d}{n^d}\right) + n^c\frac{\delta(\gamma_H - \gamma_L) + N(2\widetilde{\gamma} - \gamma_L)}{(N+\delta)^2}\right], \tag{A4}$$

which shows the gains and losses for the network airline from decreasing θ_L (i.e., sending fewer L-types direct). First, the airline incurs the connecting discount to compensate for layover time disutility (μ_L) for those passengers who switch from the direct to the connecting service. Second, the airline incurs the passenger cost ($2\tau^c$) and the frequency cost ($\frac{2\omega}{n^c}$) associated to routes AH and BH, whereas it saves the passenger cost (τ^d) and the frequency cost ($\frac{\omega d}{n^d}$) associated to the direct service on route AB. Finally, the last term captures the fact that savings from sending fewer L-types direct increase with load factor of connecting aircraft, capturing the cost advantage in terms of economies of traffic density stemming from larger aircraft size (and lower frequency), which leads to lower fares.

Equivalently, Eq. (7) reduces to

$$\Gamma \equiv \delta \left[\mu_H + 2\tau^c - \tau^d + \omega \left(\frac{2}{n^c} - \frac{d}{n^d} \right) - n^d \left(\gamma_H - \gamma_L \right) + n^c \frac{\gamma_H - 2\widetilde{\gamma}}{N + \delta} \right], \tag{A5}$$

which indicates that the gain from sending all the H-types direct logically increases with their layover time disutility (μ_H) and with the costs corresponding to routes AH and BH ($2\tau^c + \frac{2\omega}{n^c}$). In contrast, the network airline incurs the costs associated to the direct service on route AB ($\tau^d + \frac{\omega d}{n^d}$). The last two terms show the preference of H-types for service quality (i.e., flight frequency). Thus, the higher the load factor on route AB (which increases n^d), the lower the frequency and the higher the cost for H-types to fly direct. Equivalently, the higher the load factor on routes AH and BH (which increases n^c), the lower the frequency and the higher the savings from switching to a direct connection.

Finally, Eq. (8) yields this condition

$$\Upsilon \equiv (1 - \delta) \left[-\mu_L - 2\tau^c + \tau^d - \omega \left(\frac{2}{n^c} - \frac{d}{n^d} \right) - \delta n^d \left(\gamma_H - \gamma_L \right) + n^c \frac{2\widetilde{\gamma} - \gamma_L}{N} \right], \tag{A6}$$

which has a similar interpretation to Eq. (A4). ■

B Appendix: Proofs

Proof of Lemma 1.

From Eqs. (A1), (A2) and (A3), we obtain the following threshold values for distance

$$d_\Omega = \tfrac{n^d}{\omega} \left[\mu_L + 2\tau^c - \tau^d + \tfrac{2\omega}{n^c} + n^d \tfrac{\gamma_H - \gamma_L}{\delta} - N n^c \tfrac{2\widetilde{\gamma} - \gamma_L}{(N + 1 - \delta)^2} \right], \tag{A7}$$

$$d_\Phi = \tfrac{n^d}{\omega} \left[\mu_H + 2\tau^c - \tau^d + \tfrac{2\omega}{n^c} - n^d \tfrac{\gamma_H}{\delta} + n^c \tfrac{(1-\delta)(\gamma_H - \gamma_L) + N(\gamma_H - 2\widetilde{\gamma})}{(1+N)(1+N-\delta)} \right], \tag{A8}$$

$$d_\Lambda = \tfrac{n^d}{\omega} \left[\mu_L + 2\tau^c - \tau^d + \tfrac{2\omega}{n^c} - n^c \tfrac{\delta(\gamma_H - \gamma_L) - N(2\widetilde{\gamma} - \gamma_L)}{(1+N)^2} \right], \tag{A9}$$

where $\Omega, \Lambda < 0$ imply $d > d_\Omega, d_\Lambda$, and $\Phi > 0$ implies $d < d_\Phi$. Therefore, $(\theta_H^*, \theta_L^*) = (1, 0)$ arises for $d \in (\max\{d_\Omega, d_\Lambda, 0\}, d_\Phi)$. We assume that this interval is non-empty, a condition that is guaranteed for a sufficiently small n^d relative to n^c (i.e., RJs need to be sufficiently small as compared to mainline jets).[21] Finally, since $\Phi < 0$ implies $d > d_\Phi$, then $(\theta_H^*, \theta_L^*) = (0,0)$ arises for $d > d_\Phi$. ■

Proof of Corollary 1.

This corollary explains the requirements that must hold to sustain the optimal distribution of passengers $(\theta_H^*, \theta_L^*) = (1, \theta_L^*)$ with $\theta_L^* \in (0,1]$. To have (at least) some L-types traveling direct, i.e., $\theta_L^* \in (0,1]$, we need $\min\{d_\Omega, d_\Lambda\} > 0$ and $d \in (0, \min\{d_\Omega, d_\Lambda\})$. In addition, $d < d_\Phi$ ensures $\pi(1,0) > \pi(0,0)$, but it does not guarantee to observe $\theta_H^* = 1$ for any θ_L^*. At this point, let us define $\Sigma \equiv \pi(1,\theta_L) - \pi(0,\theta_L) > 0$, where

$$\Sigma \equiv \delta \left[\mu_H + 2\tau^c - \tau^d + \omega\left(\tfrac{2}{n^c} - \tfrac{d}{n^d}\right) - n^d \tfrac{\gamma_H - \gamma_L}{\delta + \theta_L(1-\delta)} + n^c \tfrac{(1-\delta)(1-\theta_L)(\gamma_H - \gamma_L) + N(\gamma_H - 2\widetilde{\gamma})}{[N + (1-\delta)(1-\theta_L)][1 + N - (1-\delta)\theta_L]}\right]. \tag{A10}$$

Therefore $d < d_\Sigma$ implies $\Sigma \equiv \pi(1,\theta_L) - \pi(0,\theta_L) > 0$ for any $\theta_L \in [0,1]$, ensuring that all H-types still fly direct, where

$$d_\Sigma = \tfrac{n^d}{\omega}\left[\mu_H + 2\tau^c - \tau^d + \tfrac{2\omega}{n^c} - n^d \tfrac{\gamma_H - \gamma_L}{\delta + \theta_L(1-\delta)} + n^c \tfrac{(1-\delta)(1-\theta_L)(\gamma_H - \gamma_L) + N(\gamma_H - 2\widetilde{\gamma})}{[N + (1-\delta)(1-\theta_L)][1 + N - (1-\delta)\theta_L]}\right]. \tag{A11}$$

Finally, imposing $d < d_\Omega$ (which implies $\frac{\partial \pi(1,0)}{\partial \theta_L} > 0$) is sufficient to guarantee that the airline sends (at least) some L-type passengers direct (and the condition $d < d_\Lambda$ is not needed anymore). In conclusion, $d < \min\{d_\Omega, d_\Sigma\}$ sustains the optimal division of passengers $(\theta_H^*, \theta_L^*) = (1, \theta_L^*)$ with $\theta_L^* \in (0,1]$. Note that $d_\Omega < d_\Sigma$ is satisfied for a sufficiently small n^d relative to n^c.

Note that, from the expression for $\Sigma \equiv \pi(1,\theta_L) - \pi(0,\theta_L)$ above, we cannot recover $\Phi \equiv \pi(1,0) - \pi(0,0)$ by setting $\theta_L = 0$ (observe the element that multiplies n^d in the expressions for Φ and Σ). The reason is that there is a discontinuity in $\pi(0,\theta_L)$ between $\theta_L = 0$ and $\theta_L > 0$ because $\theta_L = 0$ implies dismantling the direct route between cities A and B and sending all passengers through the hub (i.e., adopting a HS network). ■

Proof of Lemma 2.

From Eqs. (A4), (A5) and (A6), we obtain the following threshold values for distance

$$d_\Psi = \tfrac{n^d}{\omega}\left[\mu_L + 2\tau^c - \tau^d + \tfrac{2\omega}{n^c} - n^c \tfrac{\delta(\gamma_H - \gamma_L) + N(2\widetilde{\gamma} - \gamma_L)}{(N+\delta)^2}\right], \tag{A12}$$

$$d_\Gamma = \tfrac{n^d}{\omega}\left[\mu_H + 2\tau^c - \tau^d + \tfrac{2\omega}{n^c} - n^d(\gamma_H - \gamma_L) + n^c \tfrac{\gamma_H - 2\widetilde{\gamma}}{N+\delta}\right], \tag{A13}$$

$$d_\Upsilon = \tfrac{n^d}{\omega}\left[\mu_L + 2\tau^c - \tau^d + \tfrac{2\omega}{n^c} + \delta n^d(\gamma_H - \gamma_L) - n^c \tfrac{2\widetilde{\gamma} - \gamma_L}{N}\right], \tag{A14}$$

where $\Psi, \Upsilon < 0$ imply $d < d_\Psi, d_\Upsilon$, and $\Gamma < 0$ implies $d > d_\Gamma$. Therefore, $(\theta_H^*, \theta_L^*) = (0,1)$ for $d \in (d_\Gamma, \min\{d_\Psi, d_\Upsilon\})$. We assume that this interval is non-empty, a condition that is guaranteed for a sufficiently large n^d relative to n^c (i.e., the load factor in the low-cost flights on

[21] Computations available from the autors on request.

route AB is sufficiently high as compared to the load factor in regular flights on routes AH and BH).[22] Finally, when $\Gamma > 0$ then $d < d_\Gamma$ and $(\theta_H^*, \theta_L^*) = (1, 1)$. ■

Proof of Corollary 2.
This corollary explains the requirements that must hold to sustain the optimal distribution of passengers $(\theta_H^*, \theta_L^*) = (0, \theta_L^*)$ with $\theta_L^* \in [0, 1)$. To have (at least) some L-types traveling connecting, i.e., $\theta_L^* \in [0, 1)$, we need $d > \max\{d_\Psi, d_\Upsilon\}$. However, this condition does not guarantee that all H-types still fly connecting (i.e., $\theta_H^* = 0$), which requires $\Sigma < 0$ or, equivalently, $d > d_\Sigma$ (the expressions for Σ and d_Σ are given in the proof of Corollary 1). Therefore, $d > \max\{d_\Psi, d_\Sigma\}$ sustains the optimal division of passengers $(\theta_H^*, \theta_L^*) = (0, \theta_L^*)$ with $\theta_L^* \in [0, 1)$. Note that $d_\Psi > d_\Sigma$ for a sufficiently large n^d relative to n^c. ■

C Appendix: Details on the numerical analysis

These are the values for all the critical values of distance: $d_\Lambda = 0.848$, $d_\Omega = 0.852$, $d_\Phi = 0.896$, $d_\Gamma = 1.027$, $d_\Psi = 1.959$ and $d_\Upsilon = 4.485$. Finally let us denote d_Σ^{RJ} and d_Σ^{LC} the values of d_Σ, depending on the type of PP connection between endpoints A and B. Note that d_Σ^{RJ} and d_Σ^{LC} are functions of θ_L. On the one hand, d_Σ^{RJ} is a concave function that takes values between 0.935 (when $\theta_L = 0$) and 1.099 (when $\theta_L = 0.83$). On the other hand, d_Σ^{LC} is an increasing and concave function that takes values between -5.547 (when $\theta_L = 0$) and 1.027 (when $\theta_L = 1$).
There are a number of restrictions that must hold to carry out this numerical analysis. Lemma 1 states that $(\theta_H^*, \theta_L^*) = (1, 0)$ arises for $d \in (\max\{d_\Omega, d_\Lambda, 0\}, d_\Phi)$ and, since $d_\Omega > d_\Lambda > 0$, the relevant value is d_Ω. Looking at Lemma 2, $(\theta_H^*, \theta_L^*) = (0, 1)$ arises for $d \in (d_\Gamma, \min\{d_\Psi, d_\Upsilon\})$ and, since $d_\Psi < d_\Upsilon$, the relevant value is d_Ψ. Following Corollary 1, $(\theta_H^*, \theta_L^*) = (1, \theta_L^*)$ with $\theta_L^* \in (0, 1]$ is observed when $d_\Omega > 0$ and $d \in \left(0, \min\left\{d_\Omega, d_\Sigma^{RJ}\right\}\right)$ and, since $d_\Omega < d_\Sigma^{RJ}$ holds for any $\theta_L \in [0, 1]$, the relevant value is d_Ω. Finally, looking at Corollary 2, $(\theta_H^*, \theta_L^*) = (0, \theta_L^*)$ with $\theta_L^* \in [0, 1)$ occurs when $d > \max\left\{d_\Psi, d_\Sigma^{LC}\right\}$ and, since $d_\Psi > d_\Sigma^{LC}$ holds for any $\theta_L \in [0, 1]$, the relevant value is d_Ψ. ■

[22] Computations available from the autors on request.