The Role of Demand Uncertainty in Airline Network Structure

by

Philippe BARLA
and
Christos CONSTANTATOS

Département d'économique and GREEN
Université Laval
Pavillon J.A. De Sève
Québec, Canada, G1K 7P4
e-mail: philippe.barla@ecn.ulaval.ca

Abstract

In this paper, we provide another reason that may explain the adoption of the hub-and-spoke network structure in the airline industry. We show that when an airline has to decide on its capacity before the demand conditions are perfectly known, a hub-and-spoke network structure by pooling passengers from several markets into the same plane helps the firm to lower its cost of excess capacities in the case of low demand and to reduce its opportunity cost of rationing in the case of high demand.

Key words: Airlines, Network, Demand uncertainty, Hub

JEL: L11, L13, L93

We would like to thank M. Roland for his helpful remarks.

The Role of Demand Uncertainty in Airline Network Structure

1. Introduction

In the airline industry, one of the most striking changes precipitated by deregulation has been the restructuring of carrier networks from a mostly linear to a hub-and-spoke (h&s) structure (Levine 1987).[1] Economies of traffic have been identified as the major factor explaining this change. The usual argument is that an h&s network, through increased traffic density on the links to the hub (the spokes), allows airlines to use larger, more efficient aircraft and to spread fixed costs over more passengers. It also allows better quality service through increasing flight frequencies (see Bailey, Graham and Kaplan 1985, Caves, Christensen and Tretheway 1984, Hendricks, Piccione and Tan 1995). In this paper, we provide another reason for hubbing, that has been surprisingly neglected in the literature, namely better allocation of capacity under demand uncertainty.[2]

The main feature of our analysis is that an airline has to decide on its network structure, including the level of capacities it offers, before the demand conditions are completely known. In this context we show that hubbing by pooling passengers from several markets into the same plane allows the firm to adjust the allocation of capacity after the demand conditions are revealed. This flexibility means that if the demand on one market turns out to be low, thereby creating excess capacity, the firm can increase sales in other markets. Moreover, if the demand in one market ends up being high with consequent binding capacity constraints, the opportunity cost of rationing passengers is reduced by hubbing, since the firm can first ration the low valuation travelers on several markets before rationing travelers with higher willingness to pay.

[1] In a linear structure city-pairs are linked through direct service, while in an h&s network, cities are linked by direct service only to a few central airports, the hubs. Other city-pairs are linked indirectly, through the hubs.

[2] Indeed, to our knowledge, the interaction between network structure and capacity allocation across markets has not been formally examined. This is all the more surprising since the main idea constitutes an application of the notion of yield management.

2. The Model

In this paper, we develop a stylized model of network structure choice that sets asides many of the determinants of that choice that have been already studied in the literature in order to focus the analysis on the effect of demand uncertainty. Consider an airline, as in fig. 1, that serves three markets AH, BH and AB. We assume that only city H can be developed as a hub. The airline decision is thus whether to adopt a linear network, in which case market AB is served directly, or an h&s structure, in which case market AB is served indirectly through the hub H. In a linear network structure, the firm must decide on the level of capacity to offer on three links AH, BH and AB while in an h&s structure, it must choose capacities on two links, AH and BH. The airline network decisions (structure and capacity levels) are made in the context of uncertain demand conditions. The latter is resolved immediately after the network is determined and before the firm decides its price. This timing reflects that, contrary to pricing, altering a decision related to network structure in the light of new demand information can be costly. To simplify the analysis, we consider any such *ex post* adjustments as impossible.[3]

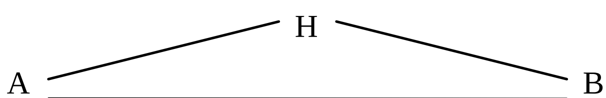


Figure 1. Network structure

We assume that the demand for air travel on the three markets is represented by $P_r = a_r - Q_r$, $r = AH, BH, AB$ (with P_r, the price and Q_r, the total number of passengers on route r). We model the uncertainty about the demand conditions by assuming that a_{AB} is a random variable with distribution function given by F_{AB}. To keep the analysis tractable, we assume that a_{AB} follows a uniform distribution on the support [0,1]. Furthermore, to reduce the

[3] For instance, canceling a flight in case of low demand may have consequences on an airline's reputation. Adjustments may be particularly difficult at slot-congested airports.

number of sub-cases to be studied, we set $a_r = 1$, for $r \neq AB$ which rules out demand uncertainty on markets AH and BH.[4]

On the cost side, we assume a very simple structure that allows us to focus on demand considerations in the choice of network structure. First, the per unit capacity cost (i.e. the cost of offering one seat) is independent of the number of passengers carried on a route. Second, this cost is c on the links that include the hub (AH, BH) while it is $2c$ on the link AB (if it is served directly). The first assumption rules out traditional economies of traffic density, already studied in the literature. By imposing the same cost for carrying a passenger from A to B, whether directly or through the hub, the second assumption eliminates some obvious cost considerations on the choice of network structure.

Finally, we assume that if the airline chooses to serve AB directly, an overflow on that market cannot be accommodated though the hub. While it may appear restrictive, this assumption is close in spirit to Oum, Zhang and Zhang (1995) and therefore could be replaced by the introduction of a fixed cost necessary to insure convenient traveling through the hub.[5] Note that the main point of the paper, namely that demand uncertainty favors hubbing, is robust to the elimination of this assumption, whose purpose is simply to highlight some of the mechanisms that bring about this result.[6]

---

[4] We further assume that travelers are not allowed to do any arbitrage. If a firm serves AB through its hub, two types of arbitrage are possible: first, AB travelers could, if profitable, buy separate tickets for each sub-route. That is buy a ticket AH and BH to travel between A and B. In our model however, given the demand structure, such an option is never profitable at the profit maximizing prices. The second possibility of arbitrage is for consumers on the hub markets to buy a ticket on AB and only use the portion corresponding to their actual journey. Following Hendricks, Piccione and Tan (1997), we exclude this option since the "carrier can stop this practice by requiring travelers to board their outgoing and return flights at the city designated on the tickets. This is indeed current practice among airlines."

[5] This cost may involve, for example, the development of an efficient system for transferring passengers and their luggage or acquiring time slots that are compatible. Introducing such a fixed cost C complicates the exposition of the analysis, for one can always find sufficiently high values of C that destroy the attractiveness of the h&s structure.

[6] For a more detailed discussion of its role, see below.

3. The Network Choice

Before examining the airline network choice, let us introduce the notation used below: hereafter, superscript I refers to the h&s structure (AB is served indirectly) while D refers to the linear network structure (AB is served directly). The subscripts AB refers to market between A and B while we use H to refer to the two connections (and/or markets) to the hub (AH and BH).

Let us first examine the airline capacity decision if the airline decides to offer a direct connection between A and B. Since there is no demand uncertainty in the hub markets ($a_H = 1$), the monopolist chooses capacities just equal to the profit maximizing quantities (the usual monopoly solution). That is, we have: $K_H^{D*} = Q_H^{D*} = \frac{1}{2}(1-c)$. For market AB, the demand is uncertain. The optimal capacity on that link is the solution to the following expected profit maximization program:

$$\underset{wrt\ K_{AB}^D}{Max} \quad E(\Pi(K_{AB}^D)) = \int_0^{2K_{AB}^D} \tfrac{1}{4} a_{AB}^2 \, da_{AB} + \int_{2K_{AB}^D}^{1} (a_{AB} - K_{AB}^D) K_{AB}^D \, da_{AB} - 2cK_{AB}^D \quad [1]$$

This function reflects the fact that, since capacities cannot be adjusted after the demand is revealed, the firm's objective is to maximize its revenue subject to a capacity constraint. If the demand state is such that $a_{AB} \le 2K_{AB}^D$ then the capacity constraint is not binding and the firm sells its revenues maximizing quantity (i.e. $Q_{AB} = \frac{1}{2} a_{AB}$). If on the other hand, the demand state is such that the capacity constraint is binding ($a_{AB} > 2K_{AB}^D$), the firm will sell all of its capacity ($Q_{AB} = K_{AB}^D$). It is easy to verify that the optimal K_{AB}^D solving [1] is given by $K_{AB}^{D*} = \frac{1}{2} - \sqrt{c}$. Note that market AB is only served if $c \le \frac{1}{4}$.

If the firm adopts an h&s network serving AB through H, it has to choose its capacities on links AH and BH, which will be used to serve both the hub and non-hub passengers. Given the symmetry of the two hub links (same demand and cost conditions), the capacities chosen on these two links must be similar. For a specific choice K_H^I and after the demand condition on AB is revealed, the firm maximizes its revenues subject to potential capacity limits. There are two cases to be considered depending on whether K_H^I is higher than $\frac{1}{2}$ or not.

Case 1- $K_H^I > \frac{1}{2}$: if the demand state is such that $a_{AB} \leq 2K_H^I - 1$, the monopolist's capacity constraints are not binding and therefore it sells its revenue maximization quantities in all three markets after the demand is revealed (i.e. $Q_{AB} = \frac{1}{2} a_{AB}$ and $Q_H = \frac{1}{2}$). If the demand state is such that $a_{AB} \geq 2K_H^I - 1$, then the firm is constrained by its capacities. The capacity on each link must be allocated between the hub and the non-hub markets. The optimal allocation is such that the marginal profit of selling one more seat in market AB is equal to its marginal cost. This cost is nothing more than the opportunity cost associated with having to ration one consumer in market AH and BH. Taking into account that the capacity constraints are binding (i.e. $Q_H = (K_H^I - Q_{AB})$), the optimal allocation condition is $(a_{AB} - 2Q_{AB}) = 2(1 - 2(K_H^I - Q_{AB}))$, leading to:

$$\overline{Q}_{AB} = \tfrac{4}{6} K_H^I + \tfrac{1}{6}(a_{AB} - 2) \text{ and } \overline{Q}_H = \tfrac{2}{6} K_H^I - \tfrac{1}{6}(a_{AB} - 2) \qquad [2]$$

where the bar indicates that capacity constraints are binding. The firm therefore maximizes the following expected profit function:

$$E(\Pi(K_H^I)) = \int_0^{2K_H^I - 1} (\tfrac{1}{4} a_{AB}^2 + 2 \times \tfrac{1}{4}) da_{AB} + \int_{2K_H^I - 1}^{1} (a_{AB} - \overline{Q}_{AB}) \overline{Q}_{AB} + 2(1 - \overline{Q}_H) \overline{Q}_H \, da_{AB} - c(2K_H^I) \qquad [3]$$

which yields $K_H^{I*} = \frac{1}{2}(2 - \sqrt{6c})$. This solution is only valid for $c \in [0, \frac{1}{6}]$ since the objective function is valid for $K_H^{I*} \geq \frac{1}{2}$.

Case 2- $K_H^I \leq \frac{1}{2}$: in this case, the monopolist always sells its entire capacity on both links. However, for low values of a_{AB} (i.e., $a_{AB} \leq 2 - 4K_H^I$), it only serves the hub markets since the marginal profit of selling to a first passenger in market AB is lower than the marginal opportunity cost of this sale. For $a_{AB} \geq 2 - 4K_H^I$, the optimal allocation between the hub and non-hub markets is given by $\overline{Q}_{AB}$ and $\overline{Q}_{H}$ in equation [2]. To determine what capacities to offer, the monopolist maximizes the following objective function:

$$E(\Pi(K_H^I)) = \int_0^{2-4K_H^I} 2(1-K_H^I)K_H^I \, da_{AB} + \int_{2-4K_H^I}^{1} (a_{AB} - \overline{Q}_{AB})\overline{Q}_{AB} + 2(1-\overline{Q}_H)\overline{Q}_H da_{AB} - c(2K_H^I) \quad [4]$$

which yields $K_H^{I*} = \frac{1}{32}(20 - \sqrt{384c - 48})$, for $c \in [\frac{1}{6}, \frac{1}{2}]$.

We can now compare the difference in maximun expected profits between the indirect and direct network structures in order to determine the firm's optimal network choice. Figure 2 shows this difference to be positive for all relevant values of *c*, implying that the monopolist will always adopt an h&s network structure.[7]

The advantage of hubbing in the presence of demand uncertainty results from the fact that by pooling consumers, the h&s structure offers the flexibility to adjust the allocation of capacity across markets after the demand has been revealed. Such flexibility is not available to a firm that chooses the direct network. This advantage has various implications depending on the demand state that can be best illustrated by examining the *ex post* difference in profit as a function of the demand state a for a specific value of *c*. Since our objective is to understand why any given investment in capacity yields higher profits under the hub network, the profit differences between the two structures are computed at capacities leading to the same total capacity costs. Figure 3 reproduces the profit difference for *c*=0.1. The *ex post* profits of the direct network have been computed at optimal capacity levels (K_{AB}^{D*}, K_H^{D*}) and those of the h&s network structure, by setting $K_H^I = K_{AB}^{D*} + K_H^{D*}$, thereby ensuring similar total investment.[8] Four zones (Z) can be distinguished in figure 3: in Z1 ($a_{AB} \leq 0.2674$), the demand conditions in market AB are such

[7] Obviously, if there is a fixed cost associated with the development of the hub, the network structure choice will depend upon the comparison between the amount of this fixed cost and the expected profit difference.

[8] I.e., $c(2K_H^I) = 2cK_{AB}^{D*} + c(2K_H^{D*})$. This assumption helps to highlight the effects in favors of the h&s network by eliminating *ex post* profit differences due to different capacity choices *ex ante*. The actual difference in expected profit is actually higher than what can be derived from Fig 3, since in the latter, the h&s monopolist has been constrained to a non-optimal capacity level. Note, however, that if computed using *ex ante* optimal capacities for both structures, profit differences could be negative for some demand states. This happens because, for some realizations of a, the difference between the *ex ante* capacity choice and its *ex post* optimal level turns out to be lower for the direct network.

that capacity constraints do not matter in either network structures. In this zone, the difference in *ex post* profits is given by:

$$(\tfrac{1}{4}a_{AB}^2 - \tfrac{1}{4}a_{AB}^2) + 2(\tfrac{1}{4} - \tfrac{1}{4}(1-c)(1+c)) \qquad [5]$$

While revenues from market AB are the same in both network structures (the first term is zero), revenues from the hub markets are higher in the h&s structure (the second term is positive). In this case, the flexibility provided by hubbing allows the firm to use part of the excess capacities created by a low demand in AB to increase sales in the hub markets. In zone Z2 ($0.2674 < a_{AB} < 0.3674$), capacity constraints are binding in the h&s network structure while they are not in the linear network structure. The *ex post* profit difference in this zone is given by:

$$[(a_{AB} - \overline{Q}_{AB})\overline{Q}_{AB} - \tfrac{1}{4}a_{AB}^2] + 2[(1-\overline{Q}_H)\overline{Q}_H - \tfrac{1}{4}(1-c)(1+c)] \qquad [6]$$

Revenues from market AB are higher under the D structure (the first term is negative). This difference is, however, more than offset by the h&s's revenue advantage in the hub markets (the second term), due to its flexibility in allocating capacities *ex post*. The D network structure induces an inefficient capacity allocation: there is excess capacity on AB whereas the marginal revenue in the hub markets is positive. This distortion, and thus the profit difference, declines with a_{AB}. In Z3 and Z4, capacity constraints are binding in both network structures. The *ex post* profit difference is then:

$$[(a_{AB} - \overline{Q}_{AB})\overline{Q}_{AB} - (a_{AB} - K_{AB}^D)K_{AB}^D] + 2[(1-\overline{Q}_H)\overline{Q}_H - \tfrac{1}{4}(1-c)(1+c)] \qquad [7]$$

In Z3, $a_{AB} < a_{AB}^* = 0.5675$, the value of a for which the *ex ante* capacity allocation of the D structure turns out to be efficient *ex post*. The distortion in capacity allocation is similar to that in Z2: given its total capacity investment, the direct monopolist is selling too much in AB and not enough in the hub markets. As a_{AB} approaches a_{AB}^*, the profit difference decreases since *ex post* flexibility becomes less of an issue. In Z4, $a_{AB} > a_{AB}^*$ and flexibility in capacity allocation once again becomes important. Hubbing now allows the firm to reduce its opportunity cost of rationing passengers: low valuation AH and BH consumers can be dropped in order to serve high valuation AB customers. This option is not available to the firm serving AB directly since we have ruled out for that firm the possibility of rerouting an overflow on AB through the hub. Allowing for such rerouting would allow the two structures to perform equally in high demand states (Z4). However, it would not challenge the overall superiority of the h&s, due to its better performance in low demand states. Furthermore, note that if one relaxes the assumption of no demand uncertainty in markets AH and BH, the advantage of hubbing is likely to be enhanced, since if demand in one of the hub market turns out to be high, the h&s monopolist is able to ration AB passengers to increase sales in the high demand hub market.[9] In the case, even if we allow rerouting in the direct network structure, an overflow on one of the hub markets, say AH could only be accomadated through B. While flying distances have not been explicitly modeled in this paper, an A-B-H itinerary usually involves significant circuity and can be easily ruled out under even mild assumptions about consumer preferences.

It is finally worth noting that, in our setting, the h&s network is always superior to the linear structure in terms of expected consumer surplus. This is so even if the total investment in capacity is larger in the linear structure than in the h&s configuration since the flexibility provided by hubbing allows to serve more passengers with less capacity. Moreover, this flexibility helps to avoid the rationing of high valuation travelers, thereby increasing consumer surplus even further. For instance, when capacity cost is high (i.e., $c>0.25$) no AB passenger will be served under the linear structure, while an h&s network will still allow some AB customers with sufficiently high valuations to be served.

4. Conclusion

In this paper, we show that hubbing provides airlines with increased flexibility in adjusting their capacity allocation across markets as new information about demand conditions becomes available. This argument can easily be extended to encompass even *anticipated* demand shifts, such as seasonal effects. Hubbing, thus, may also reflect an effort to combine markets with non-perfectly positively correlated anticipated demand variations.

[9] Unless, of course, demands in the various markets are perfectly positively correlated.

Assessing the importance of demand uncertainty in explaining the wide adoption of the h&s network structure would require: i) evaluating the importance of demand uncertainty in various markets, and ii) determining to what extend the random demand components are correlated across markets. Some preliminary discussions with airline officials seem to confirm both the presence of a market-idiosyncratic uncertainty as well as the presence of medium run inflexibility in re-affecting capacity across markets. Of course the issue begs for a more rigorous empirical analysis.

Finally, let us point out that the interaction between demand uncertainty and such factors as cost considerations, consumers preferences and strategic interactions - all left aside in this paper- may yield important new insights. The analysis of these interactions features high in our research agenda.

REFERENCES:

Bailey E.E., D.R. Graham and D.P. Kaplan (1985), *Deregulating the Airlines*, Cambridge, Mass.: MIT Press.

Caves D.W., L.R. Christensen and M.W. Tretheway (1984), "Economies of Density versus Economies of Scale: why Trunk and Local Service Airline Costs Differ," *RAND Journal of Economics,* Vol. 15, no 4, 471-486.

Hendricks K., M. Piccione and G. Tan (1995), "The Economics of Hubs: The Case of Monopoly," *Review of Economics Studies*, Vol. 62, 83-99.

Levine M (1987), "Airline Competition in Deregulated Markets : Theory, Firm Strategy, and Public Policy." *Yale Journal of Regulation*, Vol. 8, 393-494.

Oum T.A., A. Zhang and Y. Zhang (1995), "Airline Network Rivalry," *Canadian Journal of Economics*, Vol. 28, 836-857.

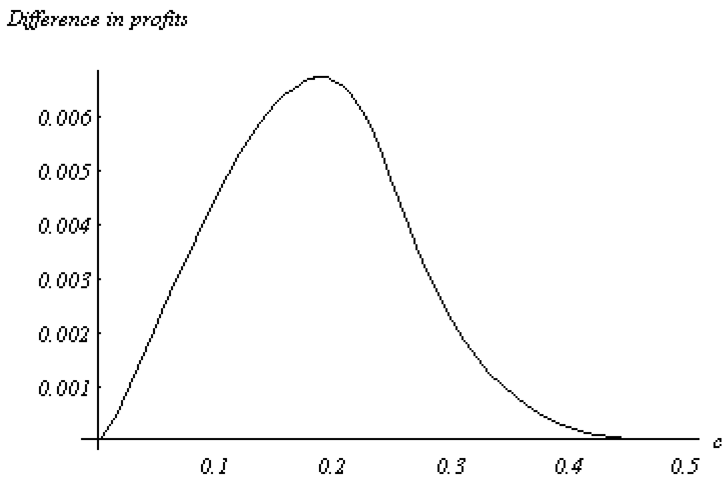
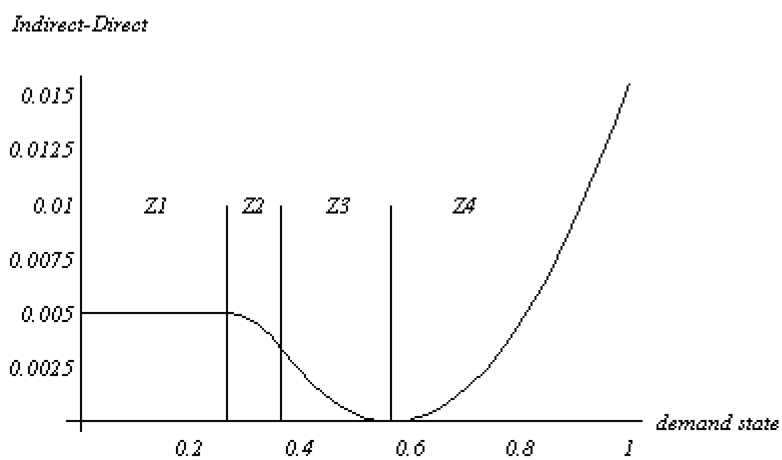


Figure 2. Difference in the firm expected profit of the two network structures (I-D) as a function of *c*.



Network Structure	Capacities	Effects of Capacity Constraints
AB is served directly	$K_H^{D*} = 0.45, K_{AB}^{D*} = 0.1837$	• Capacity constraint is binding for $\alpha_{AB} > 0.3674$
AB is served through the hub	$K_H^I = 0.45 + 0.1837 = 0.6337$	• Capacity constraints is binding for $\alpha_{AB} > 0.2674$

Figure 3. Difference in the firm *ex post* profit (I-D) as a function of α_{AB} for c=0.1.