<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>NETWORK STRUCTURE AND ENTRY IN THE DEREGULATED AIRLINE INDUSTRY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sub Title</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Author</strong></td>
<td>BERECHMAN, Joseph</td>
</tr>
<tr>
<td></td>
<td>PODDAR, Sougata</td>
</tr>
<tr>
<td></td>
<td>SHY, Oz</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Keio Economic Society, Keio University</td>
</tr>
<tr>
<td><strong>Publication year</strong></td>
<td>1998</td>
</tr>
<tr>
<td><strong>Abstract</strong></td>
<td>We explore how an incumbent airline firm structures its route network under the threat of entry caused by deregulation. We show that entry induces the incumbent firm to strategically alter its flight operating network from a fully connected network (FC) to a hub-and-spoke network (HS). While most of the literature always assumed that the use of a HS network is primarily for the purpose of cost reduction, we show that hub-and-spoke operation can be used as a strategic device by the incumbent firm when it faces a threat of entry in its operating zone.</td>
</tr>
<tr>
<td><strong>Notes</strong></td>
<td>Note</td>
</tr>
<tr>
<td><strong>Genre</strong></td>
<td>Journal Article</td>
</tr>
</tbody>
</table>
NETWORK STRUCTURE AND ENTRY IN THE DEREGULATED AIRLINE INDUSTRY

Joseph Berechman
Department of Economics and Public Policy, Tel Aviv University, Israel

Sougata Poddar
Indira Gandhi Institute of Development Research, Bombay, India

and

Oz Shy
Department of Economics, University of Haifa, Israel

First version received October 1997; final version accepted December 1998

Abstract: We explore how an incumbent airline firm structures its route network under the threat of entry caused by deregulation. We show that entry induces the incumbent firm to strategically alter its flight operating network from a fully connected network (FC) to a hub-and-spoke network (HS). While most of the literature always assumed that the use of a HS network is primarily for the purpose of cost reduction, we show that hub-and-spoke operation can be used as a strategic device by the incumbent firm when it faces a threat of entry in its operating zone.

Key words: Airline Deregulation, Hub-and-Spoke, Networks, Entry
JEL Classification Numbers: L93, L51

“...you may go to heaven or hell when you die, but you'll certainly stop in Atlanta [hub airport] on the way.”
(Folk saying in Florida)

“...you may go to heaven when you die, but at least it's a hell of a lot cheaper than going to Atlanta.”
(A comment on an earlier draft)

Acknowledgement. We thank anonymous referee for most valuable comments on an earlier draft.
1. INTRODUCTION

Among the many changes that took place in the worldwide aviation industry since the process of deregulation started in different regions in different phases, the intensified use of the hub-and-spoke network stands out as one of the most significant one. In a very short period after the deregulation, the hub-and-spoke routing in the US airline market had gone up by about 50 percent (Bailey et al., 1985; McShane and Windle, 1989). The rise in hub-and-spoke operation is also prevalent in Europe and other countries where deregulation took place and today almost all the major airlines all over the world operate a route network which is essentially a central hub oriented. Hence, the natural question arises, why do we observe a significant increase of hub-and-spoke routing network in the present aviation market? In order to find a possible answer to that question, in this paper, we develop a theoretical model which shows an airline firm which operates as a monopoly in a regulated market would find it profitable to switch to a hub-and-spoke network when it faces a potential entrant in its operating zone under a deregulated regime. We prove that switching to a hub-and-spoke network from a fully connected network (or a point-to-point network) is actually a strategic choice employed by an incumbent airline firm when it faces a threat of entry in its operating zone in a deregulated market regime.

In general, the literature on the emergence of the hub-and-spoke route network in aviation can be classified into two main categories depending on the explanations given to this phenomenon. The first is based on cost considerations where the shift to a hub-and-spoke network yields cost savings to the airline firm emanating from passenger density-economies and aircraft-size economies. For example, based on an empirical study of US airlines (1970–1984), McShane and Windle (1989) report that a 1 percent increase in hub-and-spoke routing result in a 0.11 percent decline in unit cost. Bittlingmayer (1990), Hendricks et al. (1995), and Starr and Stinchcombe (1992), assumed the existence of network economies that stem from economies of traffic density and economies of joint production (in a network setting), in order to explain the formation of a hub-and-spoke networks. Morris and Winston (1986) argue that, given demand elasticities, economies of aircraft-size can explain the utilization of a hub-and-spoke network.

The second type of explanation given in the literature for the dominant prevalence of a hub-and-spoke network, pertains to demand conditions, mainly passengers preferences relative to price and frequency of flights. Accordingly, passengers are willing to travel to their final destination via a hub rather than directly, since the increase in frequency of operations (Aircraft Movements, per week) in a hub-and-spoke network, and decline in the airfare for passengers travelling between non-hub locations, more than compensate them for their loss in utility from having to fly via a hub (Barrett, 1990; Morrison and Winston, 1986; Viton, 1986). Berechman and Shy (1996), have presented a theoretical model which shows that for a certain range of passengers demand elasticity with respect
to frequency and willingness to pay for flying directly, an airline firm will operate a hub-and-spoke network (see also Encaoua et al., 1996, for a similar analysis).

In contrast to the above, in this paper, we provide a strategic market explanation for the hub-and-spoke phenomenon which also rationalizes the presently observed, oligopolistic structure of the aviation industry. On the demand side, we have two groups of passengers who are differentiated with respect to their value of time. One group has a high value of time (for example, business travelers), and hence gain a higher utility if they fly directly to their final destination from their city of origin. The other group has a low value of time (for example, leisure travellers), and hence are less concerned whether they fly directly or indirectly to their final destination. Under this demand structure, we show that entry of a new airline induces the incumbent airline firm to strategically alter its operating network structure from a fully connected network (FC) to a hub-and-spoke network (HS). Whereas most of the literature always assumed that the use of a HS network is primarily for the purpose of cost reduction, we demonstrate that hub-and-spoke operation can be used as a strategic device by the incumbent airline when it faces a threat of entry into it operating zone.

In this connection, we would like to mention that recent similar studies by Oum et al. (1995) and Hendricks et al. (1997), which analyze the strategic interactions among the airline firms, where one of the strategic variables is again the choice of network. Oum et al. consider a three-city two stage networking game between two carriers. In the first stage, firms simultaneously select their route structures (i.e. either a fully connected network or a hub-and-spoke network). Each airline firm incurs sunk investment cost of hub development if it chooses a HS network. In the second stage, firms simultaneously set their output levels for a city-pair market by competing in Cournot–Nash fashion. They showed that investment in hubbing makes a firm 'tough' in product market competition, thus, hubbing is a 'top-dog' strategy in the terminology of Fudenberg and Tirole (1984). On the other hand, Hendricks et al. considered an entry game where a regional carrier (the entrant) invades a spoke (i.e. a route between a non-hub and hub city) of the hub-and-spoke system of a national carrier (the incumbent). They showed that if the size of the network is large enough, the hub operator's optimal response to entry in a spoke is not to withdraw its flights from that spoke, even if the regional carriers stays. As a result, regional carriers are forced to exit and entry is deterred. In contrast to this literature, in this paper, we consider a entry game where the entrant wants to operate in a route connecting between two non-hub cities or in other words, wants to enter the "rim" of the hub-and-spoke system (as opposed to the "spoke") of the incumbent firm.1 We conduct our analysis using heterogeneous passengers and characterize a price competition between an incumbent and a potential entrant.

The rest of the paper is organized as follows. In the next section, we describe

1 Indeed, a number of airline companies in the US, such as Southwest Airlines, Reno Air, and Morris Air, have invaded into rims of hub-and-spoke networks.
our model. In section 3, we consider the case of a single monopoly airline. The main analysis is done in section 4, where we consider entry. Finally, section 5 concludes with some discussions.

2. THE MODEL

Consider an economy composed of three cities $A$, $B$, and $C$. Each city pair is connected by an airline route denoted by $i$, $i=1, 2, 3$. Thus, a route $i$ is defined as a link connecting a specific pair of cities. A passenger can be transported by an airline firm either directly from his city of origin to his city of destination, or indirectly through a third city. Such a city will be called a hub.

We distinguish between two basic types of networks that the firm can operate: a Fully-Connected-network (FC) (i.e. a point-to-point network), where all passengers are serviced via direct flights between origin—destination city pairs; and a Hub-and-Spokes network (HS) where all city pairs are linked via a hub. Thus, under the HS network, there are no direct services between city pairs except for those passengers whose final destination or point of departure is the hub city itself. Figures 1a and 1b illustrate a FC network and a HS network with a hub located at city $B$.\(^2\) Thus, there are three routes: route 1 connecting $A$ and $B$, route 2 connecting $B$ to $C$, and route 3 connecting $A$ to $C$. To simplify the notation and the analysis we disregard round trips.

2.1 Passengers

We assume that on each route $i$, $i=1, 2, 3$ there are two types of passengers who are differentiated with respect to their value of time (i) $n>0$ passengers with high value of time are assumed to gain an extra utility of $d$, $(d>0)$, from flying directly to their final destination instead of indirectly through the hub; and (ii) $n>0$ passengers with a low value of time are assumed to be indifferent between flying directly or indirectly to their destination. A good example of the first type of passengers could be business travelers and of the second type could be leisure travelers.

Let $p_i$ denote the airfare on route $i$. The utility of a passenger with a high value of time on route $i$ is affected by whether the flight is direct or an indirect and by the airfare. Formally,

$$
U_{i}^{H} = \begin{cases} 
\beta + d - p_i & \text{if flies directly to destination} \\
\beta - p_i & \text{if flies to destination via a hub} \\
0 & \text{if does not fly at all} 
\end{cases} 
$$

\(^2\) In this paper, we do not discuss the reasons for choosing city $B$ as the location of the hub. This choice can have a mix of economic, historical, and institutional reasons, see Hansen and Kanafani (1990).
The utility of a passenger with low value of on route $i$ is affected by the airfare only. Hence,

$$u_i^L = \begin{cases} 
\beta - p_i & \text{if flies directly or indirectly} \\
0 & \text{if does not fly at all.} 
\end{cases}$$

(2)

where $\beta$, ($\beta > 0$) is the basic value a passenger attaches to the service of being transported from a city of origin to the city of destination.

2.2 The Airline Firm

Let $c$, ($c > 0$), denote the airline’s cost per flight on any route $i$. Observe that this cost is per flight and not per passenger, and will therefore be referred to as the ACM cost (Aircraft Movement Cost). Let $K$ denote the aircraft capacity. That is, $K$ is the maximum number of passengers that can be flown on a single aircraft. We assume that $\beta > 2c/K$ which guarantees that hub-and-spoke services are economically valuable in the sense that a passenger’s valuation for an indirect flight exceeds the per passenger cost of routing the consumer via a hub.

3. A SINGLE MONOPOLY AIRLINE

Consider an airline firm operating under a regulatory regime which permits only one firm to provide services to all cities. This type of regime is observed in a number of countries where only one airline is allowed to provide all domestic flights. We assume that this firm behaves as a monopoly, i.e. that it charges monopoly fares.\(^3\) This analysis will serve as our benchmark case.

\(^3\) One could argue that a single regulated firm is not allowed to charge monopolistic airfare. Here, we associate a single regulated firm with a monopoly since assuming otherwise complicates the analysis in the sense that there is no consensus of what are the objectives of a single regulated airline firm.
3.1 Fully Connected Service Only

A Fully Connected (FC) network is defined here as a network in which traveling from any city of origin into any city of destination consists of a direct flight which does not involve routing via a hub. In reality, save perhaps for low level operations where an airline provides direct services in few markets (routes) only, most airlines maintain many indirect flights.  

On each route, the monopoly airline has two options: (i) to charge a low price \( p_i = \beta \) and serve both types of passengers, or (ii) \( p_i = \beta + d \) and serve only the passengers with high value of time. Let \( \pi_i \) denote profit from the operation on route \( i \), \( i = 1, 2, 3 \), and let \( \pi \) denote the monopoly airline’s profit from the operation of the network. That is, \( \pi = \pi_1 + \pi_2 + \pi_3 \).

When the airfare is low \( (p_i = \beta) \), the number of flights on each route is \( 2n/K \), hence, \( \pi_i = 2n(\beta - c/K) \). When the fare is high \( (p_i = \beta + d) \), the number of flights on each route is \( n/K \), hence, \( \pi_i = n(\beta + d - c/K) \). Comparing the two profits level yield

\[
p_i = \begin{cases} 
\beta + d & \text{if } d > \beta - c/K \\
\beta & \text{if } d \leq \beta - c/K 
\end{cases}
\]

hence, \( \pi^\text{FC} = \begin{cases} 
3n(\beta + d) - 3nc/K & \text{if } d > \beta - c/K \\
6n\beta - 6nc/K & \text{if } d \leq \beta - c/K 
\end{cases} \) \hspace{1cm} (3)

Equation (3) reveals that if the ‘business’ passengers have a very high value of time \( (d \text{ takes a large value}) \), then a monopoly airline operating a FC network will raise the price to extract the entire surplus from these travelers, thereby excluding the passengers with low value of time. In contrast, if \( d \) is low, then a profit-maximizing monopoly airline will adjust to price as to extract the entire surplus from passengers with low value of time; in this case, since the network is FC, passengers with a high value of time gain a surplus equals exactly to \( d \).

3.2 Hub-and-Spoke

Using a HS network, the airline transports route 3 passengers via a hub in city \( B \). In order to determine the airline’s monopoly airfares under the HS network, we need to make the following assumption.

ASSUMPTION 1: Passengers purchasing a ticket from city \( A \) to city \( C \) are flown via the hub located in the city \( B \), can costlessly get off or get on a plane in city \( B \) thereby terminating or starting their journey in city \( B \).

Assumption 1 implies that the airfare the airline can charge passengers on routes 1 and 2 cannot exceed the airfare on route 3, since otherwise, route 1 passengers would purchase a ticket for travelling to city \( C \) and then get off in hub city \( B \). Similarly, under this constraint, route 2 passengers would purchase a ticket for

\footnote{This definition is quite strong, but serves well our analytical investigation. Notice that even the now-established Southwest Airlines provides some level of hub operations, see Jennings (1993).}
traveling on route 3 and board the plane in hub city $B$.

Clearly, assumption 1 is more realistic than assuming the polar situation where passengers are not allowed to embark or disembark at city $B$. However, it should be pointed out that in some cases, airline companies make it difficult for passengers whose destination is city $C$ to disembark in city $B$ by simply shipping their luggage to city $C$.

Under this assumption, since route 3 passengers are flown indirectly via a hub, if the monopoly airline set a high fare, $p_i = \beta + d$, then it excludes all route 3 passengers from the market. In addition, all other passengers with a low value of time are excluded, so only $2n$ passengers will be flying with the airline. In contrast, if it charges a low fare, $p_i = \beta$, all the $6n$ passengers fly with their airline, thereby generating $8n/K$ flights. Summing up, the monopoly's profit maximizing airfare and profit levels under the HS network are

$$p_i = \begin{cases} 
\beta + d & \text{if } d > 2\beta - 3c/K \\
\beta & \text{if } d \leq 2\beta - 3c/K 
\end{cases}$$

hence, $\pi^{HS} = \begin{cases} 
2n(\beta + d) - 2nc/K & \text{if } d > 2\beta - 3c/K \\
6n\beta - 8nc/K & \text{if } d \leq 2\beta - 3c/K 
\end{cases}$ (4)

A line-by-line comparison of (3) and (4) yields

**Proposition 1.** For a monopoly airline, given the assumed constant returns to scale airline technology, the FC network is more profitable to operate than the HS network.

Proposition 1 follows from the assumed constant returns to scale technology, implying that the HS network is not profitable due to the increase in the cost associated with the increase in the number of flights on routes 1 and 2. Clearly, Proposition 1 could be mitigated if we assume increasing returns in the form of excess aircraft capacity. However, since our main purpose is to compare the effect of entry on the choice of network, the present formulation is sufficient.

4. PARTIAL DEREGULATION AND PARTIAL ENTRY

Partial deregulation is defined as the case where entry is permitted in one market (route) only. Good examples are the markets between London and Amsterdam, Dublin and Amsterdam and Dublin and London, which were deregulated even though most other routes originated and terminating at these cities are still

---

5 As many travelers already know, it's usually cheaper to fly to say, Dallas, by buying a ticket to Austin, Texas, with a stopover in Dallas. The flier then simply gets off the airplane in Dallas, saving hundreds of dollars because most airlines charge cheaper fares to less-popular destinations. Also, recently, this kind of arbitrage is reported in the trans-Atlantic routes as well. According to some airline executives, even the Armani suited crowd on the Concorde is trying this trick. They are buying tickets to Brussels from New York, for example, but getting off the plane in London. (New York—London is $4,758 one way: New York—London—Brussels is $3,900). It's direct violation of airline rules, but hard to enforce. “People do it”, says a fares analyst at Virgin Atlantic Airways, who vows that “the whole industry is taking a look at this”—Source: Wall Street Journal Europe, November 29–30, 1996.
We now analyze the strategic use of the network structure by the incumbent airline when entry is allowed into route 3. The incumbent firm is denoted by \( I \) and the potential entrant by \( E \). Under this policy regime, a new entrant can enter in one route only. Given that under HS there is no direct service on route 3, this route is a natural candidate for entry.

We assume that the potential entrant has the same cost and capacity structure as the incumbent airline firm. This rules out any kind of ex ante asymmetry (for example cost or capacity advantage) between the two airline firms.

We define a three stage game between the incumbent and the entering airline firm. At stage I, the incumbent chooses whether to maintain its FC network, or whether to restructure its route network to a HS. In stage II, the incumbent chooses its airfare on each route, \( p_i^I, i = 1, 2, 3 \). In stage III, the entering airline chooses its route 3 airfare, \( p_3^E \). Figure 2 illustrates this game.

Fig. 2. The game between an incumbent and a route 3 entrant.
We look for a subgame-perfect Nash equilibrium for this game. We, therefore, proceed by calculating the equilibrium airfare for the two subgames, the FC subgame and the HS subgame.

4.1 The Fully Connected Subgame

Airfare competition on route 3 reduces the incumbent and the entrant’s airfare to unit cost. Hence, \( p^f_3 = p^e_3 = c/K \). Hence, the incumbent does not earn above normal profit on route 3, that is \( \pi^f_3 = 0 \). Hence,

\[
\pi^{FC} = \begin{cases} 
2n(\beta + d) - 2nc/K & \text{if } d > \beta - c/K \\
4n\beta - 4nc/K & \text{if } d \leq \beta - c/K 
\end{cases}
\]

4.2 The Hub-and-Spokes Subgame

Assumption 1 implies that any price reduction the incumbent makes on route 3, must be made on routes 1 and 2. The most important observation we now make is that the incumbent cannot profitably deter entry on route 3. This follows from the analysis of the previous subgame, where airfare competition on route 3 reduces the incumbent’s profit on route 3 to zero. Therefore, it is profitable for the incumbent to raise route 3 airfare to a level in which the entrant will serve the high value of time passengers, whereas the incumbent serves the low value of time passengers by an indirect flight via the hub. Therefore,

**Proposition 2.** In the HS subgame, in stage II, if the incumbent does not abandon route 3, then its profit-maximizing airfare is: \( p^f_3 = \min\{d + c/K, \beta\} \).

**Proof.** We first must show that at this airfare, the entrant will not find it profitable to undercut the incumbent in stage III by setting \( p^e_3 = p^f_3 - \epsilon \), where \( \epsilon > 0 \) is a small number. This happens when

\[
\pi^e = (p^f_3 + d)n - nc/K \geq 2n(p^f_3 - \epsilon) - 2nc/K, \quad \text{or} \quad p^f_3 \leq d + c/K. 
\]

Equation (6) shows what maximal price set by the incumbent will lead the entrant to increase its own price to \( p^e_3 = p^f_3 + d \) and serve the high value of time consumers, instead of undercutting the incumbent by setting \( p^e_3 = p^f_3 - \epsilon \). Finally, clearly, profit can only be reduced if the incubent sets \( p^f_3 < d + c/K \) and cannot be set above \( \beta \) as no route 3 passenger is willing to pay more than \( \beta \). Q.E.D.

Hence, the profit of the incumbent firm

\[
\pi^{HS} = \begin{cases} 
5n\beta - 6nc/K & \text{if } d > \beta - c/K \\
5n(d + c/K) - 6nc/K & \text{if } d \leq \beta - c/K
\end{cases}
\]

We now state our main proposition.

**Proposition 3.** An incumbent airline firm which faces an entering airline firm on route 3 and which does not abandon route 3 (i.e., some route 3 passengers fly
with the incumbent) will choose to operate a HS network over the FC network whenever \( \frac{1}{2}(3\beta-4c/K) \leq d \leq \frac{1}{2}(3\beta-4c/K) \).

**Proof.** First note that under the FC network, airfare in route 3 drops to unit cost, hence the incumbent makes zero profits from that route. Looking at a line-by-line comparison of (5) and (7) yields the following.

When \( d > (\beta-c/K) \), HS operation is profitable than FC operation if and only if \( d \leq \frac{1}{2}(3\beta-4c/K) \). Notice that \( \frac{1}{2}(3\beta-4c/K) > (\beta-c/K) \) as long as \( \beta > 2c/K \) (by assumption).

On the other hand, when \( d \leq (\beta-c/K) \), HS operation is profitable than FC operation if and only if \( d \geq \frac{1}{2}(4\beta-3c/K) \). Here also, \( \frac{1}{2}(4\beta-3c/K) < (\beta-c/K) \) as long as \( \beta > 2c/K \).

Combining above cases, we get the result. Q.E.D.

Following Proposition 3, we can now state the unique subgame perfect equilibrium for the incumbent to operate HS network, set \( p_i^t = \text{Min}\{d+c/K, \beta\} \), and the entering airline to set

\[
p_i^E = \begin{cases} 
  p_i^t + d & \text{if } p_i^t \leq d + \frac{c}{K} \\
  p_i^t - \epsilon & \text{if } p_i^t > d + \frac{c}{K} 
\end{cases}
\]

Finally, it should be pointed out that in some instances the incumbent will want to abandon route 3 passengers and serve routes 1 or 2 only and charge \( \beta \) or \( \beta + d \). It should be noted that operating a fully connected network when the entrant enters in route 3 is equivalent to abandon route 3 so far as the profit of the incumbent is concerned because both the cases entail zero profit to the incumbent firm from that route. When the incumbent abandons route 3, obviously the FC and the HS network are identical.

**Corollary.** Incumbent airline will abandon route 3 while facing a new entrant in that route and (i) charge a price \( \beta + d \) in routes 1 and 2 if and only if \( d > \frac{1}{2}(3\beta-4c/K) \), (ii) charge a price \( \beta \) in routes 1 and 2 if and only if \( d < \frac{1}{2}(4\beta-3c/K) \).

Intuitively, this means when \( d \) is “high” (actually so high) that it is most profitable to charge \( \beta + d \) and extract all the surplus from route 1 and 2 high time-valued passengers and abandon route 3 and other low time-valued passengers in route 1 and 2. On the other hand when \( d \) is “low” the maximum price that the incumbent firm can charge under a HS operation is low. This price does not generate enough profit to the firm to maintain HS operation and thus it abandons route 3. Notice that under HS operation the incumbent has to fly the route 3 low time-valued passengers twice, which inevitably increases its operating cost. Hence, when \( d \) is “medium” it is most profitable to operate a HS network by serving the
The use of the HS network under entry accommodation serves the incumbent as a mean to differentiate the service provided by the incumbent and the entrant. By shifting to HS network the incumbent can prevent a stiff price competition with the entrant associated with having both airline firms providing a homogeneous.
service under the FC network. The result in proposition 3 stems from the fact that since passengers are heterogeneous with respect to their value of time, the entrant and the incumbent can split the market according to passengers’ value of time. We showed, that there is a significant range of $d$ (the time preference parameter) under which the incumbent will indeed operate a HS network since the two services (direct and indirect) become differentiated. This finding gives us a rationale about the fact of prevalent existence of a hub-and-spoke network in the present aviation industry.

REFERENCES