International Journal of Industrial Organization 28 (2010) 377–382

Contents lists available at ScienceDirect

International Journal of Industrial Organization

journal homepage: www.elsevier.com/locate/ijio

A dynamic game of airline network competition: Hub-and-spoke networks and entry deterrence

Victor Aguirregabiria [a,*], Chun-Yu Ho [b]

[a] University of Toronto, Canada
[b] Georgia Institute of Technology, United States

ARTICLE INFO

Available online 6 March 2010

JEL classification:
C10
C35
C63
C73
L10
L13
L93

Keywords:
Dynamic games
Airline networks
Hub-and-spoke
Entry deterrence
Supermodularity

ABSTRACT

In a hub-and-spoke network, the profit function of an airline is supermodular with respect to the airline's own entry decisions for different city-pairs. This source of complementarity implies that a hub-and-spoke network can be an effective strategy for deterring the entry of competitors. This paper presents an empirical dynamic game of airline network competition that incorporates this entry deterrence motive for using hub-and-spoke networks. We summarize the results of the estimation of the model, with special attention to empirical evidence regarding the entry deterrence motive.

1. Introduction

An airline's network is the set of city-pairs that the airline connects via non-stop flights. The choice of network structure is one of the most important strategic decisions of an airline. Two network structures that have received particular attention in studies of the airline industry are *hub-and-spoke networks* and *point-to-point networks*. In a *hub-and-spoke network*, an airline concentrates most of its operations in one airport, called the hub. All other cities in the network (i.e., the spokes) are connected to the hub by non-stop flights such that travelers between two spoke cities must take a connecting flight to the hub. In contrast, in a point-to-point network, all cities are connected with each other through non-stop flights. Pure hub-and-spoke and pure point-to-point networks are very rare. They represent the two extreme cases of the degree of concentration of an airline's operations in a few airports. Table 1 presents concentration ratios based on airline networks including the 55 largest US cities.[1] While most airlines have some degree of concentration of their activity in a few airports, there is also very significant heterogeneity in their concentration ratios.

The relationship between network structure and airlines' operating costs has received significant attention among IO economists in both theoretical and empirical work. Different studies have shown how a hub-and-spoke network can exploit significant economies of scope at the airport level and economies of traffic density.[2] An argument for the use

* Corresponding author. Address: 150 St. George Street. Toronto, ON, Canada M5S 3G7. Tel.: +416 9784358.
E-mail address: victor.aguirregabiria@utoronto.ca (V. Aguirregabiria).

0167-7187/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijindorg.2010.03.003

[1] Table 1 presents two concentration ratios, *CR1* and *CR2*. Let n be the total number of non-stop connections of an airline. Let $n^{(1)}$ be the number of non-stop connections of the airline at the airport where it has the most connections. Similarly, let $n^{(k)}$ be the number of non-stop connections of the airline in its *k-th* largest airport, excluding connections with any of its $k-1$ largest airports. Then, for any integer K greater than or equal to one, the index CRK is defined as $(\Sigma_{k=1}^{K} n^{(k)})/n$. The following are some examples of the values of these concentration ratios: for a pure hub-and-spoke network, $CR1 = CR2 = 1$; for a pure point-to-point network, $CR1 = 2/n$; for a network with two hubs, $CR1 > 1/2$ and $CR2 = 1$.

[2] A hub-and-spoke network can fully connect C cities using the minimum number of direct connections, $C-1$. Therefore, it minimizes fixed costs associated with establishing non-stop connections. Furthermore, the routes that an airline has utilizing the same airport may share some common operating costs. An airline can benefit from these economies of scope by concentrating its operations in a few airports. Last but not least, larger planes are more efficient on a per-seat basis, and airlines can exploit these cost savings by seating passengers in a single plane who have different final destinations or who come from different points of origin (i.e., economies of traffic density). See Caves et al. (1984), Berry (1990), Brueckner and Spiller (1991), and Brueckner (2004), among others.

Table 1
Measures of 'Hubbing' in the US airline industry: year 2004.

Airline (code)	1st largest hub (# connections)	Concentration ratio CR1	2nd largest hub (# connections)	Concentration ratio CR2
Southwest (WN)	Las Vegas (35)	9.3	Phoenix (33)	18.2
American (AA)	Dallas (52)	22.3	Chicago (46)	42.0
United (UA)	Chicago (50)	25.1	Denver (41)	45.7
Delta (DL)	Atlanta (53)	26.7	Cincinnati (42)	48.0
Continental (CO)	Houston (52)	36.6	New York (45)	68.3
Northwest (NW)	Minneapolis (47)	25.6	Detroit (43)	49.2
US Airways (US)	Charlotte (35)	23.3	Philadelphia (33)	45.3
America West (HP)	Phoenix (40)	35.4	Las Vegas (28)	60.2
Alaska (AS)	Seattle (18)	56.2	Portland (10)	87.5
ATA (TZ)	Chicago (16)	48.4	Indianapolis (6)	66.6
JetBlue (B6)	New York (13)	59.0	Long Beach (4)	77.3
Frontier (F9)	Denver (27)	56.2	Los Angeles (5)	66.6
AirTran (FL)	Atlanta (24)	68.5	Dallas (4)	80.0
Trans States (AX)	St Louis (18)	62.0	Pittsburgh (7)	93.9
Reno Air (QX)	Portland (8)	53.3	Denver (7)	100.0
Sun Country (SY)	Minneapolis (11)	100.0	(0)	100.0

Source: DB1B Database form the US Bureau of Transportation. Year 2004.

of hub-and-spoke networks that has received almost no empirical attention is the one that postulates that some airlines can use hub-and-spoke networks as a strategy to deter the entry of competitors. This argument was first established by Hendricks et al. (1997) using a sequential game of entry between an incumbent hub-and-spoke carrier and a point-to-point regional carrier.[3] In a hub-and-spoke network, the profit function of an airline is supermodular with respect to its entry decisions for different city-pairs. This complementarity implies that a hub-and-spoke airline may be willing to operate non-stop flights for a city-pair even when profits are negative because operating between that city-pair can generate positive profits connected with other routes. Potential entrants are aware of this, and therefore, it may deter entry. This argument for entry deterrence does not suffer from several limitations that hinder other more standard arguments of predatory conduct. In particular, it does not require a *sacrifice* on the part of the incumbent (i.e., a reduction in current profits) that will be compensated for in the future only if competitors do not enter the market.[4] Furthermore, it is not subject to well-known criticisms of some arguments and models of spatial entry deterrence (see Judd, 1985). [5]

Despite these attractive features of the Hendricks–Piccione–Tan entry deterrence argument, there are no previous studies that empirically explore this entry deterrence motive in airlines' use of hub-and-spoke networks. Part of the reason for this lack of empirical evidence is the absence of structural models of dynamic network competition that incorporate this hypothesis and that are flexible and realistic enough to be estimated with actual data. The objective of this paper is to present a dynamic game of airline network competition that incorporates the strategic entry deterrence motive of a hub-and-spoke network and that can be estimated using publicly available data from the US Bureau of Transportation. We describe how the estimated model can be used to test for strategic entry deterrence. In a companion paper (Aguirregabiria and Ho, 2009), we estimate this model and use it to measure the contribution of demand, cost, and strategic factors to explaining hub-and-spoke networks. Here, we summarize the main empirical results of that paper, with particular attention to the empirical evidence regarding the entry deterrence motive.

2. Model

2.1. Basic framework

The industry is configured by N airline companies and C cities. The *network* of an airline consists of the set of city-pairs that the airline connects with non-stop flights. Entry/exit for a city-pair is not *directional* — i.e., if an airline operates non-stop flights from city A to city B, then it should operate flights from B to A. Therefore, there are $M \equiv C(C-1)/2$ markets or *city-pairs*. We index time using t, markets using m, and airlines using i. An airline's *network* can be represented by a vector $\mathbf{x}_{it} \equiv \{x_{imt}: m=1,2,...,M\}$, where $x_{imt} \in \{0,1\}$ is the binary indicator of the event "airline i operates non-stop flights for city-pair m at period t". The whole industry network is represented by the vector $\mathbf{x}_t \equiv \{\mathbf{x}_{it}: i=1,2,...,N\} \in X$, where $X \equiv \{0,1\}^{NM}$. Travelers are concerned about *routes*. A *route* is a directional round-trip between two cities, including possible stops. A network implicitly describes all of the routes for which an airline provides flights, either with stops or non-stop. In principle, we can construct routes with many stops. However, we consider only routes with zero, one, or two stops.[6] $L(\mathbf{x}_{it})$ is the set of these routes associated with network $\mathbf{x}_{it}$. We index routes using r.

Every period (quarter) t, airlines take as given the current industry network $\mathbf{x}_t$ and choose prices for all of the routes where they operate flights either non-stop or with stops.[7] Price competition is static and determines variable profits for each airline and route.[8] Airlines also decide their networks for the next quarter, $\mathbf{x}_{it+1}$. We assume that it takes one quarter to build the inputs that are needed to start operating non-stop flights between a city-pair. Similarly, we assume that it takes one quarter to scrap the inputs to exit from servicing a city-pair. Fixed costs and entry

[3] See also Oum et al. (1995) and Hendricks et al. (1999).

[4] It is difficult to generate this type of predatory conduct as a stationary Markov perfect equilibrium. Furthermore, in antitrust cases, it is typically quite difficult to find convincing empirical evidence regarding the *sacrifice* component of the argument. See the papers by Kim (2009) and Snider (2009) that deal with this issue in the context of the US v. American Airlines case.

[5] Judd (1985) notes that some models of entry and spatial location that generate entry deterrence as a subgame perfect equilibrium include strong assumptions regarding firms' level of commitment. Those papers assume that entry and location decisions are completely irreversible, with no possibility of exit or relocation. Judd shows that when there is strong enough substitutability among the stores of the same firm, allowing for exit may result in non-successful spatial preemption by the incumbent. Potential entrants know that the incumbent firm may prefer to have a monopoly in a single location rather than being a monopolist in one location and a duopolist in another nearby location. Therefore, spatial preemption and entry deterrence by the incumbent do not constitute a credible strategy. This type of argument does not apply to a hub-and-spoke airline because the profits from different city-pairs (different "stores") are not substitutes but are rather complements. This complementarity makes entry deterrence a credible strategy in equilibrium.

[6] Routes with more than two stops represent less than 1% of all of the air tickets in the US Origin and Destination (DB1B) database.

[7] The DB1B database has quarterly frequency.

[8] Intertemporal price discrimination and plane capacity constraints can generate dynamic (forward-looking) pricing strategies at the level of individual flights (i.e., specific flight number and day). However, that type of pricing dynamics is short-run and flight-specific, and it plays a very minor role in the dynamics of network structure. For simplicity's sake, this model ignores dynamic pricing.

costs are paid at quarter t, but entry and exit decisions are not effective until quarter $t+1$.[9] The airline's total profit function is:

$$\Pi_{it} = \sum_{r \in L(\mathbf{x}_{it})} R_{ir}(\mathbf{x}_t) - \sum_{m=1}^{M} x_{imt+1}(FC_{imt} + (1-x_{imt})SC_{imt}) \quad (1)$$

$R_{ir}(\mathbf{x}_t)$is the variable profit of airline i in the Nash–Bertrand equilibrium for route r. FC_{imt} represents the fixed cost for airline i in market m and quarter t. SC_{imt} is a start-up cost or entry cost at the city-pair level — i.e., an additional fixed cost that should be paid if airline i was not active in market m at period t and decides to start operations there in period $t+1$. The equilibrium of the dynamic game implies a Markov transition probability for the industry network, $Pr(\mathbf{x}_{t+1}|\mathbf{x}_t)$, and its corresponding ergodic probability distribution, $p^*(\mathbf{x}_t)$.[10]

2.2. Demand and price competition

For notational simplicity, we omit the time subindex t from the description of the static model of demand and price competition. Let H_r be the number of potential travelers in route r. For a given route, there are two forms of product differentiation: the airline (i) and the indicator for non-stop flights (n). Travelers decide which product (i,n) to purchase, if any. The indirect utility for a consumer on route r who purchases product (i,n) is $b_{irn}-p_{irn}+v_{irn}$, where p_{irn} is the price, b_{irn} is the *quality* of the product, and v_{irn} is a consumer-specific component that captures consumer heterogeneity in preferences. Travelers can choose an *outside alternative* of not traveling by air. The quality and price of the outside alternative are normalized to zero. In this paper, we ignore hub-size effects on demand and variable costs and consider a simple specification for product quality: [11] $b_{irn}=\alpha_i^{(0)}+\alpha_i^{(1)}n$, where the parameter $\alpha_i^{(0)}$ represents the quality of a flight with stops, and $\alpha_i^{(0)}+\alpha_i^{(1)}$ represents the quality of a non-stop flight. We assume that v_{irn} are independent Type I extreme value random variables. Therefore, the aggregate demand of product (i,n) on route r is $q_{irn}=\exp\{b_{irn}-p_{irn}\}/[1+\sum_k \exp\{b_k-p_k\}]$, where the sum $\sum_k$ is over all products available for route r. The variable profit of airline i on route r is $R_{ir}=(p_{ir0}-c_{ir0})q_{ir0}+(p_{ir1}-c_{ir1})q_{ir1}$, where c_{irn} is the constant marginal cost of product (i,r,n). The specification of this marginal cost is similar to that of product quality: $c_{irn}=\omega_i^{(0)}+\omega_i^{(1)}n$, where $\omega_i^{(0)}$ and $\omega_i^{(0)}+\omega_i^{(1)}$ represent the marginal costs of flights with stops and non-stop flights, respectively. Nash–Bertrand equilibrium prices depend on the quality and marginal costs of all the products that are active on the same route.

[9] Note that, given our assumption that entry/exit decisions are made one period ahead (i.e., time-to-build), the assumption regarding the timing of the payment of the entry cost and fixed cost is really innocuous. Let F_{imt} and SC_{imt} be the fixed cost and the entry cost in our model, respectively, under our assumption that these costs are paid in period t for operation during period $t+1$. Suppose that these costs were not actually paid in period t but were sometime instead paid between periods t and $t+1$ —i.e., the fixed cost is paid at $t+d_{FC}$ and the entry cost is paid at $t+d_{SC}$—for some values $d_{FC}\in[0,1]$ and $d_{SC}\in[0,1]$ that are unknown to us as researchers. This implies the following relationship between our "structural parameters", F_{imt} and SC_{imt}, and the actual values of the costs, F_{imt}^* and SC_{imt}^*: $F_{imt}=\beta^{d_{FC}}F_{imt}^*$, and $SC_{imt}=\beta^{d_{SC}}SC_{imt}^*$, where β is the discount factor. That is, our "structural parameters" are discounted values of the actual current values of these costs. It is clear that not knowing the actual timing of the payments implies that we can only set-identify F_{imt}^* and SC_{imt}^* from estimates of F_{imt} and SC_{imt}: i.e., all that we know is that $F_{imt}^*\in[F_{imt},F_{imt}/\beta]$ and $SC_{imt}^*\in[SC_{imt}, SC_{imt}/\beta]$. Nevertheless, for most of the relevant empirical questions, all that we need to know about the fixed cost and entry cost are the discounted values F_{imt} and SC_{imt} (i.e., $\beta^{d_{FC}}F_{imt}^*$ and $\beta^{d_{SC}}SC_{imt}^*$). We do not need to know the current values F_{imt}^* and SC_{imt}^*. Furthermore, given the quarterly frequency of our data and the value of the discount factor, $\beta=0.99$, the difference between these values is smaller than 1%.

[10] In Aguirregabiria and Ho (2009), we see a more general version of the model that includes exogenous state variables, $\mathbf{z}_t$, that affect demand and costs. In that model, the dynamics of the industry can be described using the endogenous Markov transition probability $Pr(\mathbf{x}_{t+1}|\mathbf{x}_t,\mathbf{z}_t)$ and the exogenous transition probability of $\mathbf{z}_t$.

[11] Aguirregabiria and Ho (2009) consider a richer specification of demand and variable costs that includes a nested logit structure for travelers' idiosyncratic preferences, permanent airline and city heterogeneity, and hub-size effects on both demand and variable costs.

2.3. Supermodularity of variable profit

Let $TR_i \equiv \sum_{r \in L(\mathbf{x}_i)} R_{ir}(\mathbf{x})$ be the total variable profit function. For the main purpose of this paper, it is important to note that for an airline with a hub-and-spoke network this function is supermodular with respect to the airline's own entry decisions for different city-pairs. To illustrate this point, consider an industry with three cities, A, B, and C. There are three city-pairs (AB, AC, and BC), and an airline's network is described in terms of three binary indicators of non-stop flights: x_{AB}, x_{AC}, and x_{BC} — we omit the airline subindex for notational convenience. For the sake of simplicity, suppose that the three city-pairs are equivalent in terms of the variable profits that they generate. Let R^{ns} be the variable profit on one route if the airline operates only non-stop flights. Similarly, R^s is the variable profit if the airline operates only flights with stops, and R^{ns+s} is the profit when there are both non-stop flights and fights with stops. Consumer substitution between non-stop and stop flights on the same route implies that $R^{ns+s} \leq (R^{ns}+R^s)$. The total variable profit function is as follows:

$$\begin{aligned} TR(x_{AB},x_{AC},x_{BC}) = {} & (x_{AB}+x_{AC}+x_{BC})R^{ns} + (x_{AC}x_{BC}+x_{AB}x_{BC}+x_{AB}x_{AC})R^s \\ & + 3x_{AB}x_{AC}x_{BC}\left(R^{ns+s}-R^{ns}-R^s\right) \end{aligned} \quad (2)$$

Suppose that the airline has a hub-and-spoke network with the hub at city A — i.e., $x_{AB}=1$, $x_{AC}=1$, and $x_{BC}=0$. The profit of this airline is $TR\ (1,1,0)=2R^{ns}+R^s$. The profit of operating in only one city-pair is $TR\ (1,0,0)=R^{ns}$. Therefore, for a hub-and-spoke network the variable profit function is supermodular — i.e., $[TR(1,1,0)-TR_i(0,1,0)]-[TR_i\ (1,0,0)-TR_i(0,0,0)]=R^s>0$. This implies that the airline is willing to operate non-stop flights in a city-pair (say, AB) even if profits from that route are negative as long as this negative profit is more than compensated for the profit from the route B to C with a stop at A. It is straightforward to show that the degree of supermodularity in the variable profit function increases with the number of spoke cities in the hub-and-spoke network. That is, the larger the hub, the stronger the supermodularity and the more likely it is that an airline using a hub-and-spoke network will be willing to operate some spoke routes with negative profits. This is known by potential entrants into the spoke route and can deter them.

In contrast, in a point-to-point network, either there is no supermodularity or it is significantly weaker than in a hub-and-spoke network. The variable profit of a point-to-point network for this airline is $TR(1,1,1)=3R^{ns+s}$, and we have that $[TR(1,1,1)-TR_i(0,1,1)]-[TR_i\ (1,0,1)-TR_i(0,0,1)]=3(R^{ns+s}-R^{ns}-R^s)+R^s$. The term $3(R^{ns+s}-R^{ns}-R^s)$ is negative because of the substitutability between non-stop flights and flights with stops within the same route. Therefore, it is clear that the complementarity between the entry decisions is weaker than in a hub-and-spoke network.

2.4. Fixed costs and start-up costs[12]

The structure of the fixed cost is $FC_{imt}=(\gamma_i^{FC(0)}-\gamma_i^{FC(1)}HUB_{imt})+\varepsilon_{imt}^{FC}$, where $\gamma_i^{FC(0)}\geq 0$ and $\gamma_i^{FC(1)}\geq 0$ are parameters. The component ε_{imt}^{FC} is private information that the airline possesses on its own cost. This private information shock is assumed to be independently and identically distributed over firms and over time with zero mean. HUB_{imt} represents the hub-size of airline i in the airports of city-pair m as measured by the total number of cities that airline i connects with non-stop flights from the origin and destination airports in city-pair m: $HUB_{imt}\equiv\sum_{m' \in C_m} x_{im't}$, where C_m is the set of markets with a common city with market m. The parameter $\gamma_i^{FC(0)}$ represents airline i's fixed cost in a market where it does not have any other connections.

[12] As in the case of variable profits, Aguirregabiria and Ho (2009) consider a richer specification of fixed costs and entry costs that includes permanent heterogeneity for both airlines and cities.

The parameter $\gamma_i^{FC(1)}$ measures how airline i's fixed costs decline with its hub-size in the city-pair. The start-up cost SC_{imt} has the same structure as the fixed cost: $SC_{imt}=(\delta_i^{SC(0)}-\delta_i^{SC(1)}HUB_{imt})+\varepsilon_{imt}^{SC}$, where $\delta_i^{SC(0)}\geq 0$ and $\delta_i^{SC(1)}\geq 0$ are the parameters.

When $\gamma_i^{FC(1)}$ or $\delta_i^{SC(1)}$ are strictly positive, profits for different city-pairs are interconnected through the hub-size effects. This is the other source of complementarity between an airline's entry decisions for different city-pairs.

2.5. Markov perfect equilibrium

Airlines maximize intertemporal profits. They are forward-looking and take into account the implications of their entry and exit decisions for future profits and for the expected future reactions of their competitors. We assume that airlines' strategies depend only on payoff-relevant state variables — i.e., the Markov perfect equilibrium (MPE) assumption. An airline's payoff-relevant information at quarter t is $\{\mathbf{x}_t,\varepsilon_{it}\}$, where ε_{it} is the vector of airline-specific private information shocks $\{\varepsilon_{imt}^{FC}, \varepsilon_{imt}^{SC}: m=1,2,...,M\}$. Let $\sigma\equiv\{\sigma_i(\mathbf{x}_t,\varepsilon_{it}): i=1,2,...,N\}$ be a set of strategy functions, one for each airline, such that σ_i is a function from $X\times\mathbb{R}^{2M}$ into $\{0,1\}^M$. A MPE in this game is a set of strategy functions such that each airline's strategy maximizes its value for each possible state $(\mathbf{x}_t,\varepsilon_{it})$ and taking as given other airlines' strategies.

Let $V_i^{\sigma}(\mathbf{x}_t,\varepsilon_{it})$ represent the value function for airline i given that the other airlines behave according to their respective strategies in σ, and given that airline i uses his best response/strategy. By the principle of optimality, this value function is implicitly defined as the unique solution to the following Bellman equation:

$$V_i^{\sigma}(\mathbf{x}_t,\varepsilon_{it}) = \max_{\mathbf{x}_{it+1}} \Big\{\Pi_i(\mathbf{x}_t,\mathbf{x}_{it+1})-\varepsilon_{it}(\mathbf{x}_{it+1}) + \beta E\left[V_i^{\sigma}(\mathbf{x}_{t+1},\varepsilon_{it+1})\,|\,\mathbf{x}_t,\mathbf{x}_{it+1}\right]\Big\} \tag{3}$$

where $\beta\in(0,1)$ is the discount factor; $\Pi_i(\mathbf{x}_t,\mathbf{x}_{it+1})$ represents the common-knowledge part of the profit function, i.e., $\Pi_i(\mathbf{x}_t,\mathbf{x}_{it+1})\equiv\sum_{r\in L(\mathbf{x}_{it})}R_{ir}(\mathbf{x}_t)-{}_{m=1}^{M}x_{imt+1}\left(\left(\gamma_i^{FC(0)}-\gamma_i^{FC(1)}HUB_{imt}\right)+(1-x_{imt})(\delta_i^{SC(0)}-\delta_i^{SC(1)}HUB_{imt})\right)$; and the term $\varepsilon_{it}(\mathbf{x}_{it+1})$ contains the private information part of the profit function, i.e., $\varepsilon_{it}(\mathbf{x}_{it+1})\equiv{}_{m=1}^{M}x_{imt+1}(\varepsilon_{imt}^{FC}+(1-x_{imt})\varepsilon_{imt}^{SC})$. A set of strategies σ is a MPE if, for every airline i and every state $(\mathbf{x}_t,\varepsilon_{it})$, we have that:

$$\sigma_i(\mathbf{x}_t,\varepsilon_{it}) = \operatorname*{argmax}_{\mathbf{x}_{it+1}} \Big\{\Pi_i(\mathbf{x}_t,\mathbf{x}_{it+1})-\varepsilon_{it}(\mathbf{x}_{it+1}) + \beta E\left[V_i^{\sigma}(\mathbf{x}_{t+1},\varepsilon_{it+1})\,|\,\mathbf{x}_t,\mathbf{x}_{it+1}\right]\Big\}. \tag{4}$$

That is, every airline's strategy is a best response to the other airlines' strategies.

An equilibrium in this dynamic game provides a description of the joint dynamics of price, quantities, and airlines' incumbent status for every route between the C cities of the industry. To compute the equilibrium and perform the structural estimation of the model, one may define an MPE in terms of airlines' *conditional choice probabilities* (CCPs). Define the choice probability $P_i(\mathbf{x}_{it+1}|\mathbf{x}_t)\equiv\int 1\{\sigma_i(\mathbf{x}_t,\varepsilon_{it})=\mathbf{x}_{it+1}\}dG_{\varepsilon}(\varepsilon_{it})$, where $1\{.\}$ is the indicator function, and G_{ε} is the CDF of ε_{it}. $P_i(\mathbf{x}_{it+1}|\mathbf{x}_t)$ is the probability that airline i operates a network $\mathbf{x}_{it+1}$ at period $t+1$ given that the industry network at period t is $\mathbf{x}_t$. Let $\mathbf{P}$ be the vector of CCPs associated with σ, i.e., $\mathbf{P}=\{P_i(\mathbf{x}_{it+1}|\mathbf{x}_t):\ i=1,2,...,N;\ \mathbf{x}_{it+1}\in\{0,1\}^M;\ \mathbf{x}_t\in X\}$. Following Aguirregabiria and Mira (2007), a MPE in this dynamic game can be described as a vector of probabilities $\mathbf{P}$ that solves the fixed point problem $\mathbf{P}=\Psi(\mathbf{P})$, where $\Psi(\mathbf{P})$ is a vector-valued best-response probability function.

2.6. Entry deterrence and hub-and-spoke networks

The vector of structural parameters of the model, θ, includes parameters in demand, $\{\alpha_i^{(0)},\alpha_i^{(1)}\}$, variable costs, $\{\omega_i^{(0)},\omega_i^{(1)}\}$, fixed costs, $\{\gamma_i^{FC(0)},\gamma_i^{FC(1)}\}$, and entry costs, $\{\delta_i^{SC(0)},\delta_i^{SC(1)}\}$. Aguirregabiria and Ho (2009) show that these parameters are identified using data on prices and quantities at the airline-route level (to identify demand and variable cost parameters) and longitudinal data on airline networks (to identify fixed costs and entry costs).

Given consistent estimates of the vector of structural parameters θ and of the equilibrium in the data as represented by the vector of choice probabilities $\mathbf{P}$, we are interested in measuring the role of hub-and-spoke networks as a credible strategy for deterring the entry of point-to-point carriers. This entry deterrence argument is based on the supermodularity (complementarity) of the total variable profit function of a hub-and-spoke airline. The elimination of this supermodularity should also remove this potential source of entry deterrence. More specifically, if this supermodularity generates entry deterrence, then in eliminating it for a certain airline, we should find that the airline has both a lower tendency toward 'hubbing' (i.e., a lower concentration of its operations in a few airports as measured by concentration ratios *CR1* and *CR2*) and a lower number of city-pairs for which it operates as a monopolist.

To implement this type of comparative statics exercise, we need to define a *counterfactual scenario* wherein we eliminate this source supermodularity from the variable profit of an airline. We consider the following approach. Suppose that we describe an airline as a group of local managers, one for each city-pair m. The double index (i,m) represents the local manager of airline i in market m. This local manager decides whether to operate non-stop flights in city-pair m. In our model, the decision-making of the airline is centralized. Therefore, the model assumes that all local managers of an airline internalize the complementarities between their entry–exit decisions.[13] To eliminate supermodularity from an airline's variable profits, we consider the counterfactual scenario wherein the local managers of an airline are concerned with the maximization of its own city-pair profit, which includes only the variable profit from non-stop flights between two cities. This hypothetical local manager ignores that his city-pair is a segment in many other routes and that the operation of his city-pair can generate additional profits associated with these other routes.

To illustrate this, consider the example in Section 2.3 of an industry with three cities and a hub-and-spoke airline with a hub at city A and spokes at cities B and C. The total variable profit of this airline is $TR=(x_{AB}+x_{AC})R^{ns}+x_{AB}x_{AC}\,R^{s}$, which is a supermodular function in (x_{AB}, x_{AC}). However, in the counterfactual scenario, the local manager in city-pair AB is only concerned with its local variable profit $x_{AB}R^{ns}$ (and the local manager AC is only concerned with profit $x_{AC}R^{ns}$). Therefore, these "uncoordinated" local managers do not take into account the complementarity of their decisions in terms of the total profit of the airline. In this counterfactual model, we expect that the network of the airline will present a lower degree of 'hubbing'. Furthermore, if the entry deterrence motive is significant in the factual equilibrium, we expect that in the counterfactual, the airline will be a monopolist in a smaller number of markets.

3. Empirical evidence

Aguirregabiria and Ho (2009) estimate the dynamic game described above using data from the Airline Origin and Destination Survey (DB1B) for the 55 largest cities in the US. Here we summarize the main empirical results presented in that paper, paying particular attention to those results related to strategic entry deterrence.

[13] The estimated model in Aguirregabiria and Ho (2009) considers a certain degree of decentralization in local managers' decision-making. However, in that model, local managers still internalize the complementarities of their operation decisions.

Table 2
Entry deterrence motive of hub-and-spoke networks.

Carrier	Observed CR2	Counterfactual CR2	Observed # city-pairs where airline is a monopolist	Counterfactual change in # city-pairs where airline is a monopolist	Counterfactual change in total # monopoly city-pairs in the industry
Southwest	18.2	16.5	151	−2	−2
American	42.0	24.5	31	−5	−4
United	45.7	30.3	16	−4	−3
Delta	48.0	22.1	57	−36	−10
Continental	68.3	42.8	27	−7	−5
Northwest	49.2	23.2	66	−41	−14
US Airways	45.3	35.2	8	−2	−2

Based on estimation results in Aguirregabiria and Ho (2009).

3.1. Economies of scope in fixed costs and start-up costs

The estimates of fixed costs and start-up costs are sizeable. The mean value of fixed costs (averaged over airlines and markets) is $119,000 or 75% of city-pair quarterly variable profit. Start-up costs are on average equal to $298,000, which accounts for 187% of city-pair quarterly variable profit. Although we find very significant airline heterogeneity in terms of estimated fixed and start-up costs, the average airline has significant incentives to reduce these costs by decreasing the number of non-stop connections using a hub-and-spoke network. These incentives are even larger when we take into account economies of scope at the airport level. The estimated values of the parameters that measure these economies of scope are $\gamma^{FC(1)} =$ \$1020 and $\delta^{SC(1)} =$ \$9260. That is, one additional non-stop connection (i.e., a unit increase in hub-size HUB_{imt}) reduces city-pair fixed costs by $1020 and start-up costs by $9260 . Counterfactual experiments show that these economies of scope for fixed costs and particularly for start-up costs play an important role in explaining the propensity of *legacy carriers* to use hub-and-spoke networks.

3.2. Strategic entry deterrence

To measure the entry deterrence motive for developing hub-and-spoke networks, we implement the counterfactual experiment described at the end of previous section.[14] Table 2 presents a summary of the counterfactual experiments. Each row represents a different experiment — i.e., an experiment where we eliminate supermodularity from the variable profit of a single airline, leaving the other airlines unchanged, and calculate a new equilibrium for the dynamic game. Note that the Bertrand model of price competition is exactly the same in the factual and counterfactual scenarios. Therefore, if the products associated with a route and the hub sizes of the entrants were the same in the factual and counterfactual models, then equilibrium prices for that route would be the same under the two scenarios. Of course, we generally find that prices are not the same in the two scenarios because the number of entrants and their respective hub sizes are different.

For all of the legacy airlines (i.e., all of the airlines in Table 2 except Southwest), eliminating supermodularity from variable profits implies a significant reduction in the concentration ratio *CR2*, which measures the degree of ‘hubbing’. Interestingly, there is also an important reduction in the number of city-pairs where the airline operates as a monopolist.[15] Reductions in *CR2* and in the number of monopoly markets are particularly important for Northwest and Delta. It seems that the entry deterrence motive for using a hub-and-spoke network plays an important role for these airlines. Northwest and Delta are the airlines that, after Southwest, operate as monopolists for the largest number of city-pairs and have largest hub sizes. Also, these airlines tend to operate as monopolies in city-pairs with relatively small market size. According to our estimates, the large number of monopoly markets enjoyed by these carriers cannot be explained by more efficient exogenous costs — i.e., by lower values for the parameters $\omega_i^{(0)}$, $\omega_i^{(1)}$, $\gamma_i^{FC(0)}$, and $\delta_i^{SC(0)}$. Instead, it is mainly explained by lower values for the endogenous part of the costs — i.e., $-\gamma_i^{FC(1)} HUB_{im}$ and $-\delta_i^{SC(1)} HUB_{im}$.

In a certain sense, Southwest is the opposite of Northwest and Delta with regard to these issues. Southwest is by far the airline with the smallest contribution of the entry deterrence motive. According to the estimated model, Southwest enjoys a large number of monopoly markets not because entry deterrence but because it has much lower values than any other airline for the exogenous part of the cost ($\omega_i^{(0)}$, $\omega_i^{(1)}$, $\gamma_i^{FC(0)}$, and $\delta_i^{SC(0)}$). This cost efficiency allows the airline to operate in small markets.

The last column in Table 2 presents the change from the factual to the counterfactual scenario in the total number of city-pairs for which an airline has a monopoly. It is interesting to note that in the experiments using Northwest and Delta, the decline in the total number of monopoly markets is much smaller than the reduction in the number of monopoly markets enjoyed by these airlines. That is, other airlines replace either Northwest or Delta as monopolists. Perhaps not surprisingly, Southwest is the “replacing monopolist” in a significant proportion of these cases. These results seem fully consistent with the entry deterrence argument. Also note that in those city-pairs where a monopoly by Northwest or Delta is replaced by a monopoly by Southwest, prices decline significantly because our estimates show that the marginal costs (and the quality) of Southwest are significantly lower than those of any other airline.

Nevertheless, the net welfare effect of this type of entry deterrence behavior is ambiguous. On the one hand, this strategy restricts the entry of carriers that have ex-ante lower operating costs, such as Southwest Airlines. On the other hand, the estimated model shows that hub-and-spoke networks exploit economies of scope and density that generate ex-post cost reductions that are very significant. Southwest's low-cost strategy has been shown to be an effective way to compete with large hub-and-spoke carriers. However, the estimated model shows that this does not mean that the Pareto optimal structure of the industry would have many “Southwest-like” point-to-point carriers. This industry structure would not exploit the very significant cost savings associated with hub-and-spoke networks. In this context, it will be of great interest to consider and evaluate policies that try to reduce the entry deterrence effect of hub-and-spoke networks but that can maintain most of the cost savings associated with these networks.

Acknowledgments

We would like to thank the editor and an anonymous referee for their comments. This research has also benefited from helpful conversations with Volker Nocke, Daniel Yi Xu, and participants at the

[14] To deal with the problem of multiple equilibria in the counterfactual scenario, we implement the method in Aguirregabiria (2009).

[15] We should clarify our use of the term *monopolist* in Table 2. What we mean is that the airline is the only carrier that operates non-stop flights for the city-pair. Of course, this does not mean that the airline is really a monopolist with regard to the routes between the two cities because there may be other airlines that operate flights with stops that connect the two cities. Nevertheless, given that consumers prefer non-stop flights to flights with stops, it is clear that being the only carrier operating non-stop flights ensures significant market power.

2009 EARIE Conference in Ljubljana. The first author would like to acknowledge the University of Toronto for funding this research through a Connaught Research Grant.

References

Aguirregabiria, V., and C-Y Ho (2009): "A Dynamic Oligopoly Game of the US Airline Industry: Estimation and Policy Experiments," manuscript. Department of Economics. University of Toronto.

Aguirregabiria, V., Mira, P., 2007. Sequential estimation of dynamic discrete games. Econometrica 75, 1–53.

Aguirregabiria, V. (2009): "A Method for Implementing Counterfactual Experiments in Models with Multiple Equilibria," manuscript. Department of Economics. University of Toronto.

Berry, S., 1990. Airport presence as product differentiation. American Economic Review 80, 394–399.

Brueckner, J., 2004. Network structure and airline scheduling. Journal of Industrial Economics 52, 291–312.

Brueckner, J., Spiller, P., 1991. Competition and mergers in airline networks. International Journal of Industrial Organization 9, 323–342.

Caves, D., Christensen, L., Tretheway, M., 1984. Economies of density versus economies of scale: why trunk and local service airline costs differ. Rand Journal of Economics 15, 471–489.

Hendricks, K., Piccione, M., Tan, G., 1997. Entry and exit in hub-spoke networks. Rand Journal of Economics 28 (2), 291–303.

Hendricks, K., Piccione, M., Tan, G., 1999. Equilibria in networks. Econometrica 67, 1407–1434.

Judd, K., 1985. Credible spatial preemption. The RAND Journal of Economics 16 (2), 153–166.

Kim, S.-H., 2009. Predatory reputation in US airline markets. International Journal of Industrial Organization 27, 592–604.

Oum, T., Zhang, A., Zhang, Y., 1995. Airline network rivalry. Canadian Journal of Economics 28, 836–857.

Snider, C., 2009. "Predatory Incentives and Predation Policy: The American Airlines Case," manuscript. Department of Economics, UCLA.