



Munich Personal RePEc Archive

## **Entry Deterrence and Strategic Alliances**

Gayle, Philip and Xie, Xin

Kansas State University, East Tennessee State University

December 2017

Online at <https://mpra.ub.uni-muenchen.de/83233/>  
MPRA Paper No. 83233, posted 10 Dec 2017 23:37 UTC

# Entry Deterrence and Strategic Alliances

Philip G. Gayle\* and Xin Xie\*\*

This Draft: December 2, 2017  
First Draft: September 20, 2012

Forthcoming in *Economic Inquiry*

## Abstract

Researchers have written extensively on the impact that strategic alliances between airlines have on airfare, but little is known of the market entry deterrent impact of strategic alliances. Using a structural econometric model, this paper examines the market entry deterrent impact of codesharing, a form of strategic alliance, between incumbent carriers in domestic air travel markets. We find evidence of market entry deterrence, but deterrence impact depends on the specific type of codesharing between market incumbents as well as the identity of the potential entrant. We quantify the extent to which market incumbents' codesharing influences potential entrants market entry cost and probability of market entry.

*Keywords:* Entry Deterrence; Strategic Alliances; Dynamic Entry/Exit Model; Airline Competition

*JEL Classification codes:* L13, L93

**Acknowledgements:** For very helpful comments and suggestions we thank editor, Wesley Wilson, three anonymous referees, Dennis Weisman, Myongjin Kim, Anson Ho, Leilei Shen, Philippe Belly, Peri da Silva, Yongmin Chen, Robert McNowen, Charles de Bartolomé, Scott Savage, Jin-Hyuk Kim, Carlos Martins-Filho, George-Levi Gayle, Joseph Cullen, Glenn MacDonald, Maher Said, Bart Hamilton, and participants at the Econometric Society session on “Industrial Organization: Theory and Applications” at the January 2014 Allied Social Science Associations (ASSA) meetings. We also thank Adeel Faheem for excellent research assistance. Any remaining errors are our own.

\*Kansas State University, Department of Economics, 322 Waters Hall, Manhattan, KS 66506; Voice: (785) 532-4581; Fax: (785) 532-6919; email: [gaylep@ksu.edu](mailto:gaylep@ksu.edu);  
**Corresponding author.**

\*\*East Tennessee State University, Department of Economics and Finance, 227 Sam Wilson Hall, Johnson City, TN 37614; Voice: (423) 439-5365; Fax: (423) 439-8583; email: [xiex01@etsu.edu](mailto:xiex01@etsu.edu).

## 1. Introduction

In recent years, strategic alliances between airlines have become increasingly popular. The format of a strategic alliance between airlines can vary from a limited marketing arrangement, for example an arrangement between partner carriers that only makes their frequent-flyer programs reciprocal,<sup>1</sup> to more extensive arrangements that include reciprocal frequent-flyer programs as well as codesharing. Reciprocal frequent-flyer programs effectively allow passengers that hold frequent-flyer membership with one carrier in the alliance to earn and redeem frequent-flyer points across any partner carrier in the alliance. A codeshare arrangement effectively allows each carrier in the alliance to sell tickets for seats on its partners' airplane, i.e., partners essentially share certain facilities, in this case airplanes, that are solely owned by one of the partners.

Researchers have written extensively on the impact that strategic alliances have on airfare [Brueckner and Whalen (2000); Brueckner (2001 and 2003); Bamberger, Carlton and Neumann (2004); Ito and Lee (2007); Gayle (2008 and 2013); Gayle and Brown (2014) among others].<sup>2</sup> However, there is a paucity of work that examines the impact that strategic alliances may have on deterring potential competitors from entering a relevant market. This is a particularly interesting aspect of strategic alliances to study since a substantial amount of these alliances are formed between traditional major/legacy carriers, who may face increasingly stiff competition from the growing prominence of low-cost-carriers (LCCs). Some researchers argue that hub-and-spoke network carriers form and use strategic codeshare alliances to better compete with low-cost-carriers, [Mantovani and Tarola (2007)]. So the following series of relevant questions need careful study. First, does the evidence support the argument that strategic alliances between major airlines, among achieving other goals, serve to deter entry of potential entrants to a relevant market? Second, if an entry-deterrence effect is evident, is there a particular type of practice among

---

<sup>1</sup> Membership in an airline's frequent-flyer program allows the passenger to accumulate points each time the passenger flies on the airline. The frequent-flyer program allows the passenger to be eligible for various rewards once the passenger accumulates points beyond certain pre-determine thresholds. As such, frequent-flyer programs are designed to build customer loyalty to the carrier that offers the program.

<sup>2</sup> Earlier contributions to this literature include: Oum and Park (1997); Park (1997); Park and Zhang (1998); and Park and Zhang (2000).

alliance partners that is most effective at deterring entry? Third, does the market entry-deterrent impact of strategic alliances vary by the identity of potential market entrants?<sup>3</sup>

Chen and Ross (2000) theoretically explore the anticompetitive effect of a particular type of strategic alliance, by which the partner airlines share important facilities such as airplanes, terminals etc. They argue that this type of alliance can forestall a complete and competitive entry by another firm, that is, such alliances can have an entry-deterrent effect. The mechanism through which Chen and Ross envisioned that a strategic alliance may deter a complete and competitive entry is as follows. An incumbent offers to form a strategic alliance with a potential entrant, which takes the form of the incumbent willing to share its facility with the potential entrant in order to discourage the potential entrant from building its own facility and entering on a larger, more competitive scale. In the context of a codeshare alliance, this would translate into the incumbent offering to let a potential entrant sell tickets for seats on the incumbent's plane in order to discourage the potential entrant from putting its own plane on the route. So based on Chen and Ross's argument, entry-deterrent codesharing should primarily take place between a market incumbent and the potential entrant the incumbent is intending to deter.

Lin (2005) uses a theoretical model to show that incumbents can use codeshare alliances as a credible threat to deter the entry of potential entrants who do not have significant cost advantage. The author uses the model to show that, owing to joint profit maximizing behavior between allied airlines, there exists an equilibrium in which the joint profit of two allied airlines is higher than the sum of their individual profits if they were not allied. In addition, this higher joint profit of the allied airlines comes at the expense of lower profit for a new non-allied entrant. This equilibrium implies that if market entry cost is sufficiently high, such that entry in the presence of an alliance between market incumbents is unprofitable for the new non-allied entrant, but profitable if incumbents were not allied, then formation of the alliance can be done to strategically deter entry.<sup>4</sup>

---

<sup>3</sup> In a separate, but related airline entry-deterrence literature, Oum, Zhang and Zhang (1995); Hendricks, Piccione and Tan (1997); Berechman, Poddar and Shy (1998); Aguirregabiria and Ho (2010) among others have argued that hub-and-spoke route networks adopted by many legacy carriers do give these carriers an incentive and the ability to deter entry of other carriers that do not use hub-and-spoke route network, which include many low-cost-carriers. But this literature focuses on the entry deterrence effect of hub-and-spoke networks rather than more specifically on the entry deterrence effect of codeshare alliances.

<sup>4</sup> Lin (2008) extends this model to consider situations in which an incumbent has a relatively large hub-and-spoke network and entry has positive spillover network effects for the incumbent.

Via reduced-form econometric regressions, Goetz and Shapiro (2012) empirically test for the presence of entry-deterrence motives behind codesharing alliances, and find that an incumbent is approximately 25% more likely than average to codeshare when facing the threat of entry by low-cost carriers. However, Goetz and Shapiro (2012) did not investigate whether the entry-deterrence effect they found depends on the type of codesharing (Traditional versus Virtual)<sup>5</sup> employed by incumbent partner airlines. In addition, they did not fully investigate whether the entry-deterrence effect of codesharing depends on the identity of the carrier that is threatening to enter the relevant market.

Previous studies have argued that Southwest Airlines, if not the most formidable LCC in U.S. domestic air travel markets, is certainly among the most formidable LCCs in these markets. As such, many studies have treated Southwest separately than other LCCs, or focused on Southwest as the sole LCC [for example see Morrison (2001), Goolsbee and Syverson (2008), Brueckner, Lee and Singer (2012) among others]. Brueckner, Lee and Singer (2012) find that the presence of potential competition from Southwest reduces fares by 8 percent, while potential competition from other LCCs has no fare effect. Mason and Morrison (2008) find significant differences between low-cost carriers in their business models. Therefore, we are encouraged to investigate whether any possible entry-deterrent effect of codesharing depends on whether the potential entrant is Southwest versus other low-cost carriers.

While Goetz and Shapiro (2012) use a reduced-form regression analysis to empirically test whether domestic codesharing alliances are motivated by an entry-deterrence purpose, to the best of our knowledge, there is no other empirical analysis of this issue. We believe a structural econometric analysis of this issue is needed to take us a step further in examining the evidence on this type of strategic behavior by airlines. Advantages of using a structural econometric model are that: (1) we are able to quantify, in monetary terms, possible market entry barriers associated with codesharing; and (2) we are able to predict the extent to which a potential entrant's market entry probabilities are affected by market incumbent carrier's codesharing.

---

<sup>5</sup> In the Definition and Data section of the paper we define and distinguish Traditional and Virtual codesharing.

Therefore, the main objective of our paper is to use a structural econometric model to investigate: (1) whether codesharing between airlines in domestic air travel markets, a form of strategic alliance, has a deterrent effect on the entry of potential competitors; (2) whether there is a particular type of codesharing among alliance partners that is most effective at deterring entry; and (3) whether the market entry deterrence impact of codesharing varies by the identity of potential market entrants.

To assess the deterrent effect of codesharing on market entry of potential competitors, we proceed as follows. First, we estimate a discrete choice model of air travel demand. Second, for the short-run supply side, we assume that multiproduct airlines set prices for their differentiated products according to a Nash equilibrium price-setting game. The Nash equilibrium price-setting assumption allows us to derive product-specific markups and use them to compute firm-level variable profits, which are subsequently used in a dynamic market entry/exit game. Third, we specify a dynamic market entry/exit game played between airlines in which each airline chooses markets in which to be active during specific time periods in order to maximize its expected discounted stream of profit. Per-period profit comprises variable profit less per-period fixed cost and a one-time entry cost if the airline will serve the relevant market in the next period but not currently serving the market. The dynamic entry/exit game allows us to estimate fixed and entry costs by exploiting previously computed variable profits from the Nash equilibrium price-setting game along with observed data on airlines' decisions to enter and exit certain markets. It is the estimated effect that codesharing between incumbents have on the entry cost of potential entrants that allows us to evaluate whether codesharing has an entry deterrent effect.<sup>6</sup>

We specify entry cost functions such that we can identify whether or not the extent of codesharing by incumbent airlines in a market influences the market entry cost of potential entrants, and whether this influence differs by type of potential entrant. A potential entrant can fall into one of three categories: (1) legacy carriers; (2) Southwest Airlines; or (3) other LCCs. Since the majority of codesharing in U.S. domestic air travel markets occurs between legacy carriers, this implies that our entry cost function

---

<sup>6</sup> For examples of dynamic structural entry deterrence models see Sweeting (2013), Williams (2012), Chicu (2012), and Snider (2009).

specification effectively allows us to explore whether codesharing between legacy carriers differentially influences the market entry of: (1) other legacy carriers; (2) Southwest Airlines; (3) other LCCs; or some subset of the three carrier types.

An important aspect of our analysis is that we follow Ito and Lee (2007) and Gayle (2008) and decompose codesharing into two main types: (1) Traditional Codesharing; and (2) Virtual Codesharing. As such, we are able to investigate whether possible entry deterrent effects of codesharing depend on the type of codesharing.

Our econometric estimates from the entry cost function suggest that more codesharing, both traditional and virtual, between incumbent carriers in a market puts Southwest at a relative disadvantage to enter the market compared to all other potential entrants (legacy carriers and other low-cost carriers). Specifically, each percentage point increase in traditional codeshare products offered by incumbents in a market raises market entry cost for Southwest by 0.3%, but reduces market entry cost by 0.64% and 0.39% for legacy and “other” low-cost carriers respectively. However, each percentage point increase in virtual codeshare products offered by incumbents in a market raises market entry cost for Southwest by 0.08%, but reduces market entry cost by 0.14% and 0.36% for legacy and “other” low-cost carriers respectively. In addition, the model predicts that Southwest’s market entry probabilities increase by a mean 15.81% when the parameters that capture the entry deterrence impact of codesharing are counterfactually set to zero in the model. Therefore, codesharing by market incumbent carriers has a relative market entry deterrent effect on Southwest. Furthermore, the parameter estimates provide evidence that traditional codesharing has a larger impact on Southwest’s market entry cost compared to virtual codesharing.

We argue that the entry deterrent effect is binding for Southwest but not for others due to the evidence that the vast majority of codesharing is done between legacy carriers, and competition between Southwest and legacy carriers is stronger than competition between other low-cost carriers and legacy carriers. For example, as pointed out above, Brueckner, Lee and Singer (2012) provide evidence that incumbent legacy carriers do not cut fares in response to potential competition from other low-cost carriers, but cut fares by 8% in response to potential competition from Southwest.

The remainder of this paper is organized as follows. Next we define and discuss

relevant concepts and terms used throughout this paper, and describe how we construct the dataset of our working sample. Our econometric model is presented in section 3. Section 4 discusses the estimation procedure and summarizes estimation results. Concluding remarks are offered in section 5.

## 2. Definitions and Data

### 2.1 Definitions

A market is defined as a directional pair of origin and destination cities during a particular time period. For example, air travel from New York to Dallas is a different market than air travel from Dallas to New York. Treating markets in a direction-specific manner better enables our model to account for the impact that heterogeneity in demographic, social and economic characteristics across origin cities have on air travel demand.

An itinerary is a detailed plan of a journey from an origin to destination city, so it consists of one or more flight coupons depending on whether or not intermediate stops are required. Each coupon typically represents travel on a particular flight. Each flight has a ticketing carrier and an operating carrier. The ticketing carrier, or sometimes referred to as the marketing carrier, is the airline selling the ticket for the seat, while the operating carrier is the airline whose plane actually transports the passenger. A product is defined as the combination of ticketing carrier, operating carrier(s) and itinerary.

A pure online product has an itinerary whose operating carrier for each flight coupon and ticketing carrier are the same. For example, a two-segment ticket with both segments operated and marketed by United Airlines (UA), i.e. (UA/UA  $\rightarrow$  UA/UA)<sup>7</sup>. A flight is said to be codeshared when the operating and ticketing carriers for that flight differ. A traditional codeshared product is defined as an itinerary that has a single ticketing carrier for the trip, but multiple operating carriers, one of which is the ticketing carrier. For example, a connecting itinerary between Continental Airlines (CO) and Delta Airlines (DL), marketed solely by Delta (CO/DL  $\rightarrow$  DL/DL) is a traditional codeshared product. A

---

<sup>7</sup> The arrow notation divides carrier(s) information from one flight segment to the next. The two-letter code in front of the symbol “/” identify the operating carrier for that segment, while the two-letter code that immediately follows the symbol “/” identify the ticketing carrier.



virtual codeshared product is defined as an itinerary that has the same operating carrier for all trip segments, but this operating carrier differs from the ticketing carrier. For example, a connecting itinerary operated entirely by United Airlines but marketed solely by US Airways (US) (UA/US→ UA/US), is a virtual codeshared product.<sup>8</sup>

## 2.2 Data

We use data from the Airline Origin and Destination Survey (DB1B) collected by the Office of Airline Information of the Bureau of Transportation Statistics. The DB1B survey is a 10% random sample of airline tickets from certified carriers in the United States. A record in this survey represents a ticket. Each ticket contains information on ticketing and operating carriers, origin and destination airports, fare, number of passengers, intermediate airport stops, market miles flown on the trip itinerary, nonstop miles between the origin and destination airports, and number of market coupons. Unfortunately, there is no passenger-specific information in the data, nor is there any information on ticket restrictions such as advance-purchase and length-of-stay requirements.

The data are quarterly, and our study uses data for the entire years of 2005, 2006 and 2007. Following Aguirregabiria and Ho (2012) among others, we select data on air travel between the 65 largest US cities. Some of the cities belong to the same metropolitan area and have multiple airports. Table 1 reports a list of the cities and the relevant airport groupings we use based on common metropolitan areas.

---

<sup>8</sup> Additional discussion and examples of pure online, traditional codeshare and virtual codeshare air travel products can be found in Ito and Lee (2007) and Gayle (2007, 2008 and 2013). In addition, see Gayle and Brown (2014).

**Table 1**  
**Cities, airports and population**

City, State	Airports	City pop.		
		2005	2006	2007
New York-Newark-Jersey	LGA, JFK, EWR	8,726,847	8,764,876	8,826,288
Los Angeles, CA	LAX, BUR	3,794,640	3,777,502	3,778,658
Chicago, IL	ORD, MDW	2,824,584	2,806,391	2,811,035
Dallas, TX <sup>a</sup>	DAL, DFW	2,479,896	2,528,227	2,577,723
Phoenix-Tempe-Mesa, AZ	PHX	2,087,948	2,136,518	2,171,495
Houston, TX	HOU, IAH, EFD	2,076,189	2,169,248	2,206,573
Philadelphia, PA	PHL	1,517,628	1,520,251	1,530,031
San Diego, CA	SAN	1,284,347	1,294,071	1,297,624
San Antonio, TX	SAT	1,258,733	1,292,082	1,323,698
San Jose, CA	SJC	908,870	918,619	931,344
Detroit, MI	DTW	921,149	918,849	917,234
Denver-Aurora, CO	DEN	856,834	869,920	887,796
Indianapolis, IN	IND	789,250	792,619	796,611
Jacksonville, FL	JAX	786,938	798,494	805,325
San Francisco, CA	SFO	777,660	786,149	799,185
Columbus, OH	CMH	738,782	744,473	750,700
Austin, TX	AUS	708,293	730,729	749,120
Memphis, TN	MEM	680,515	682,024	679,404
Minneapolis-St.Paul, MN	MSP	652,481	652,003	656,659
Baltimore, MD	BWI	640,064	640,961	640,150
Charlotte, NC	CLT	631,160	652,202	669,690
El Paso, TX	ELP	587,400	595,980	600,402
Milwaukee, WI	MKE	601,983	602,782	602,656
Seattle, WA	SEA	575,036	582,877	592,647
Boston, MA	BOS	609,690	612,192	622,748

<sup>a</sup> includes Dallas, Arlington, Fort Worth and Plano

**Table 1 Continued**  
**Cities, airports and population**

City, State	Airports	City pop.		
		2005	2006	2007
Louisville, KY	SDF	559,855	559,709	562,632
Washington, DC	DCA, IAD	582,049	583,978	586,409
Nashville, TN	BNA	579,748	586,327	592,503
Las Vegas, NV	LAS	544,806	552,855	559,892
Portland, OR	PDX	534,112	538,091	546,747
Oklahoma City, OK	OKC	532,006	539,001	545,910
Tucson, AZ	TUS	524,830	530,349	536,752
Albuquerque, NM	ABQ	497,543	508,486	517,162
Long Beach, CA	LGB	467,851	463,723	459,925
New Orleans, LA	MSY	455,188	208,548	288,113
Cleveland, OH	CLE	449,188	442,409	438,068
Sacramento, CA	SMF	448,842	449,658	455,760
Kansas City, MO	MCI	463,983	470,076	475,830
Atlanta, GA	ATL	483,108	498,208	519,569
Omaha, NE	OMA	432,148	437,523	442,452
Oakland, CA	OAK	392,112	392,076	397,441
Tulsa, OK	TUL	381,017	382,394	384,592
Miami, FL	MIA	390,768	412,460	424,662
Colorado Springs, CO	COS	393,804	398,778	399,751
Wichita, KS	ICT	354,524	356,592	360,897
St Louis, MO	STL	352,572	353,837	355,663
Santa Ana, CA	SNA	337,121	334,830	335,491
Raleigh-Durham, NC	RDU	553,294	574,065	596,049
Pittsburgh, PA	PIT	316,206	313,306	312,322
Tampa, FL	TPA	325,569	332,604	334,852
Cincinnati, OH	CVG	331,310	332,185	333,321
Ontario, CA	ONT	170,630	170,865	171,603
Buffalo, NY	BUF	277,998	274,740	272,492
Lexington, KY	LEX	278,313	283,324	287,263
Norfolk, VA	ORF	237,487	238,832	236,051

We eliminate tickets with nominal prices cheaper than \$50 and more expensive than \$2000, those with multiple ticketing carriers, and those containing more than 2 intermediate stops. Within each quarter, a given itinerary-airline(s) combination is

repeated many times, each time at a different price, making the dataset extremely large. To make the data more manageable, we collapse the data based on our definition of product (unique itinerary-airline(s) combination) for each quarter. Before collapsing the data, we aggregated the number of passengers and averaged market fare over each defined product. This is the process by which each defined product's quantity and price (subsequently denoted by  $q_j$  and  $p_j$  respectively) are constructed. For example, the nonstop itinerary from New York to Dallas which is operated and ticketed both by United Airlines is repeated 3 times in the data, but with different fares \$150, \$250, and \$100, and number of passengers 5, 8, and 10 that purchase this itinerary at the three distinct prices, respectively. Then we collapse the data to leave only one observation of this product with average market fare of \$166.67, and aggregate number of passengers equal to 23. Products with quantity less than 9 passengers for the entire quarter are dropped from the data.<sup>9</sup> Also, we eliminate monopoly markets, i.e. markets in which only one carrier provides product(s). The collapsed data have 434,329 quarter-specific observations (products) spread across 32,680 quarter-specific origin-destination markets.

From the collapsed dataset, observed product market shares (subsequently denoted by upper case  $S_j$ ) are created by dividing quantity of product  $j$  sold ( $q_j$ ) by the geometric mean of the origin city and destination city populations (subsequently denoted by  $POP$ ), i.e.  $S_j = q_j / POP$ .<sup>10</sup> Other variables that capture air travel product characteristics are created for estimation. One measure of travel convenience of an air travel product is captured by the variable, *Interstops*. *Interstops* counts the number of intermediate stop(s) required by the air travel product in transporting a passenger from the origin to destination city. Our presumption is that the typical passenger dislikes intermediate stops, but the

---

<sup>9</sup> Berry (1992), Aguirregabiria and Ho (2012) among others use a similar, and sometimes more stringent, quantity threshold to help eliminate idiosyncratic product offerings that are not part of the normal set of products offered in a market.

<sup>10</sup>  $POP$  is measured by:  $POP = \sqrt{Origin\ Population \times Destination\ Population}$ . Due to the fact that population magnitudes are significantly larger than quantity sold for any given air travel product, observed product shares, computed as described above, are extremely small numbers. We therefore scale up all product shares in the data by a common factor. The common factor is the largest integer such that the outside good share ( $S_0 = 1 - \sum_{j=1}^J S_j$ ) in each market remains positive. The common factor that satisfies these conditions in the data set is 35.

demand estimation will verify whether this presumption is consistent with consumer choice behavior patterns in the data.

Another measure of the travel convenience of an air travel product is captured by the variable, *Inconvenience*. *Inconvenience* is defined as the flying distance (measured in miles, and represented by the variable *Market miles flown*) required by the travel itinerary in getting passengers from the origin to destination city, divided by the nonstop flying distance between the origin and destination cities. The minimum value for variable *Inconvenience*, which is equal to 1, implies the most travel-convenient itinerary for a given market. Furthermore, for a set of products with equivalent number of intermediate stops, the *Inconvenience* variable is able to distinguish between these products in terms of the “directness” of the routing between the origin and destination since the locations of intermediate stop(s) may differ across these products, which in turn cause the flying distance to differ across these products. Again, our presumption is that a typical passenger prefers more direct routing between their origin and destination, which will be tested by our demand estimation.

We measure the size of an airline's presence at the endpoint cities of a market from different perspectives. The variable *Opres\_out* is a count of the number of different cities that the airline offers nonstop flights to, leaving from the origin city. On the other hand, *Opres\_in* counts the number of different cities that the airline provides nonstop flights from, coming into the origin city of the market. We also construct a destination presence variable *Dpres\_out*, which measures the number of distinct cities that the airline has nonstop flights to, leaving from the destination city.

*Opres\_out* is intended to help explain consumers' choice between airlines at the consumer's origin city. The presumption here is that a consumer is more likely to choose the airline that offers nonstop service to more cities from the consumer's origin city. On the other hand, the *Opres\_in* and *Dpres\_out* may better explain an airline's cost of transporting passengers in a market. The argument is that due to possible economies of passenger-traffic density, an airline's marginal cost of transporting a passenger in a market is lower as the volume of passengers the airline channels through the market increases. An airline with large measures of *Opres\_in* and *Dpres\_out* for a given market, is likely to

channel a large volume of passengers through the market, and therefore is expected to have lower marginal cost of transporting a passenger in the market.

We only identify codeshare products between major carriers, i.e. following much of the literature on airline codesharing, we do not consider products between regional and major carriers as codeshare. For example, a product that involves American Eagle (MQ) and American Airlines (AA), where one of them is the ticketing carrier and the other is an operating carrier, is still considered by us to be pure online since American Eagle is a regional airline that serves for American Airlines.

*Traditional Codeshare* and *Virtual Codeshare* are zero-one dummy variables that take a value of 1 when the itinerary is identified to be traditional codeshared and virtual codeshared respectively. Among the codeshare products in a market, variables *Percent Traditional for Airline* and *Percent Virtual for Airline* measure the percentage of these products of a given codeshare type (Traditional and Virtual respectively) an airline offers for sale to consumers. As such, the measured percentage values in each of these codeshare variables vary across airlines and markets. These two variables are constructed to capture the extent to which each airline engages in codesharing of a given type across markets in the sample.

Summary statistics of the variables used for estimation are presented in Table 2. The variable *Fare* is measured in constant year 1999 dollars. We use the consumer price index to deflate *Fare*.

Table 3 presents a list of ticketing carriers in the dataset according to type of products that each airline provides. The first two columns show that there are 21 airlines involved in pure online products. All airlines in the dataset provide pure online products. The next two columns in Table 3 show that, among all airlines in the dataset, 10 are involved in codeshare products and 7 of these airlines are the ones we classify as legacy carriers. The fifth column in Table 3 reports the percent of codeshare products in the sample that each carrier offers for sale to consumers. The data in this column reveal that the vast majority (approximately 83 percent) of codeshare products are provided by legacy carriers.

**Table 2**  
**Summary Statistics for the Dataset**

Variable	Mean	Std.Dev	Min	Max
Fare <sup>a</sup>	166.35	52.19	45.08	1,522.46
Quantity	149.57	508.25	9	11,643
Opres_out	29.05	28.35	0	177
Opres_in	29.03	28.30	0	177
Dpres_out	29.13	28.47	0	177
Interstops	0.87	0.40	0	2
Market miles flown	1,542.34	695.27	67	4,156
Nonstop miles	1,371.42	648.60	67	2,724
Inconvenience	1.15	0.21	1	2,975
Traditional Codeshare	0.02	0.14	0	1
Virtual Codeshare	0.02	0.14	0	1
Percent Traditional for Airline	3.03	13.94	0	100
Percent Virtual for Airline	3.78	16.44	0	100
Observed Product Shares ( $S_j$ )	0.0067	0.02	5.45E-05	0.97
Number of Products	434,329			
Number of Markets	32,680			

Notes: <sup>a</sup>The variable “Fare” is measured in constant year 1999 dollars. We use the consumer price index to deflate “Fare”.

The last column in Table 3 reports the percent of each carrier’s codeshare products that are codeshared with legacy carriers. Noticeably, almost all of each legacy carrier’s codeshare products are codeshared with other legacy carriers, and moreover, ATA and Southwest Airlines, which are low-cost carriers, do not codeshare with legacy carriers. An exception to this pattern is Frontier Airlines, a low-cost carrier that has 91 percent of its codeshare products codeshared with a legacy carrier (typically with Alaska Airlines). However, the previous column shows that codeshare products offered by Frontier Airlines only account for 0.07 percent of total codeshare products offered. In summary, the data reveal that a substantial amount of codeshare alliances are formed between legacy carriers.

**Table 3**  
**List of Airlines in the Dataset,**  
**by Product type they offer to Consumers**

Airlines Involved in Pure online Products		Airlines that offer Codeshare Products to consumers			
Airlines Name	Code	Airlines Name	Code	Percent of codeshare products in the sample	Percent of each carrier's codeshare products codeshared with legacy carriers
American Airlines Inc.	AA	<b>Legacy Carriers</b>			
Aloha Airlines	AQ	American Airlines Inc.	AA	13.47	98.87
Alaska Airlines Inc.	AS	Alaska Airlines Inc.	AS	7.87	100
JetBlue Airways	B6	Continental Air Lines Inc.	CO	5.76	100
Continental Air Lines Inc.	CO	Delta Air Lines Inc.	DL	4.76	99.88
Independence Air	DH	Northwest Airlines Inc.	NW	10.03	100
Delta Air Lines Inc.	DL	United Air Lines Inc.	UA	28.75	100
Frontier Airlines Inc.	F9	US Airways Inc.	US	12.56	99.82
AirTran Airways	FL	Sub-total		83.20	
Allegiant Air	G4	<b>Low Cost Carriers</b>			
America West Airlines Inc.	HP	Southwest Airlines Co.	WN	9.28	0
Spirit Air Lines	NK	ATA Airlines	TZ	7.45	0
Northwest Airlines Inc.	NW	Frontier Airlines Inc.	F9	0.07	91.67
Skybus Airlines, Inc.	SX	Sub-total		16.80	
Sun Country Airlines	SY	Total		100	
ATA Airlines	TZ				
United Air Lines Inc.	UA				
US Airways Inc.	US				
Southwest Airlines Co.	WN				
ExpressJet Airlines Inc.	XE				
Midwest Airlines	YX				

Notes: The carries we classify as Legacy carriers include: American Airline, Alaska Airlines, Continental Air, Delta Air Lines, Northwest Airlines, United Air Lines, and US Airways.

Table 4 summarizes our data according to the three types of products. Among codeshared products, the number of traditional codeshared products is slightly less than the number of virtual codeshared products, but twice as many passengers travel on virtual codeshared products compared to traditional codeshare products. The descriptive statistics in Table 4 reveal that only 1.25% of total US domestic air travel passengers travel on codeshare products. With such a small percentage of US domestic air travel passengers using codeshare products, it is tempting to use this as justification to not study the effects



of codesharing in US domestic air travel markets. However, it is important to note that while a small percentage of total US domestic air travel passengers travel on codeshare products, the distribution of the percentage of passengers that travel on codeshare products across US domestic air travel markets is skewed. In other words, even though in a majority of markets relatively few passengers use codeshare products, there exists many markets in which a substantial percentage of passengers use codeshare products. Table 5 provides summary statistics evidence revealing that passengers are substantially more exposed to the practice of codesharing in some markets more than others, furthermore, the percentage of passengers using codeshare products in some markets can be substantial.

**Table 4**  
Classification of Cooperative Agreement in Data Set

Classification	Observations/Products		Passengers	
	Frequency	Percent	Frequency	Percent
Pure online	416,537	95.90	64,150,292	98.75
Traditional Codeshare	8,847	2.04	254,065	0.39
Virtual Codeshare	8,945	2.06	558,095	0.86
Total	434,329	100.00	64,962,452	100.00

**Table 5**  
**Passengers use of Codeshare Products Across Markets**

Percentage Interval of Consumers in the market that use Codeshare Products	Number of Market that fall within the Percentage Interval
$0\% < \text{Percent of Codeshare Passengers} \leq 5\%$	6652
$5\% < \text{Percent of Codeshare Passengers} \leq 10\%$	1513
$10\% < \text{Percent of Codeshare Passengers} \leq 20\%$	456
$20\% < \text{Percent of Codeshare Passengers} \leq 50\%$	46

Table 5 reveals that as many as 1,513 markets have between 5% and 10% of the market's passengers traveling on codeshare products, 456 markets have between 10% and 20% of the market's passengers traveling on codeshare products, while 46 markets have between 20% and 50% of the market's passengers traveling on codeshare products. In summary, the descriptive statistics in Table 5 provide sufficient reason for us to better understand the market effects of airline codesharing.

As we explain in subsequent sections of the paper, the short-run demand and supply sides of the model are estimated using the data at the product-market-time period level, while the dynamic entry/exit model is estimated using the data aggregated up to the airline-market-time period level. Since the data contain many more airlines than the dynamic entry/exit model can feasibly handle, at the stage of estimating the dynamic model, we impose additional restrictions to be able to estimate the dynamic model. A restrictive assumption we make is that a set of the airlines in our data can reasonably be lumped into an “Other low-cost carriers” category and treated as if the “Other low-cost carriers” is a single carrier. Similar to many studies in the literature [e.g. Brueckner, Lee and Singer (2012), Morrison (2001) among others], Southwest Airlines is the low-cost carrier that we treat separately than other low-cost carriers. So the “Other low-cost carriers” category includes all low-cost carriers except Southwest Airlines.

By using the number of passengers as a threshold to define whether or not an airline is active in a market, we are able to identify the number of markets that each airline has entered and exited. We define an airline to be active in a directional origin-destination market during a quarter if at least 130 passengers travel on products offered for sale by the airline in this market during the quarter.<sup>11</sup> Each airline's market entry and exit decisions contained in the data are crucial for us to be able to estimate fixed and entry costs, since the dynamic entry/exit model relies on the optimality assumption that potential entrants will only enter a market if the one-time entry cost is less than the expected discounted future stream of profits, and an incumbent will exit a market when per-period fixed cost becomes sufficiently high relative to per-period variable profits such that the expected discounted future stream of profits is non-positive. Therefore, it is useful to get a sense of the extent to which the data contain information relevant for identifying fixed and entry costs from the dynamic model.

Table 6 reports the number of market entry and exit events by airline. The table shows that each airline has several market entry and exit events, but most airlines have more market entry than market exit events, and overall there are substantially more entry

---

<sup>11</sup> Our passenger threshold of 130 for a directional market is equivalent to the 260 for non-directional market used by Aguirregabiria and Ho (2012).

than exit events. This suggests that we might be better able to identify entry cost than fixed cost.

**Table 6**  
**Number of market entry and exit events by airline**

Airlines	Number of market entry events	Number of market exit events
American Airlines Inc.	498	332
Continental Air Lines Inc.	372	303
Delta Air Lines Inc.	348	360
Northwest Airlines Inc.	323	309
United Air Lines Inc.	316	259
US Airways Inc.	655	151
Alaska Airlines Inc.	22	12
Southwest Airlines Co.	262	105
Other low cost carriers	368	625
Overall	3,164	2,456

### 3. Model

#### 3.1 Demand

Similar to Gayle (2013), air travel demand is modeled using a random coefficients logit model. There are  $POP$  potential consumers, who may either buy one of  $J$  differentiated air travel products in a market, indexed by  $j = 1, \dots, J$ , or otherwise choose the outside good ( $j = 0$ , i.e., good 0), e.g. driving, taking a train, or not traveling at all. Each potential consumer, indexed by  $c$ , chooses the travel option that gives him the highest utility, that is, we assume each potential consumer solves the following discrete choice optimization problem:

$$\max_{j \in \{0, \dots, J\}} \{U_{cj} = x_j \phi_c^x + \phi_c^p p_j + \xi_j + \varepsilon_{cj}^d\} \quad (1)$$

where  $U_{cj}$  is the value of travel option  $j$  to consumer  $c$ ;  $x_j$  is a vector of observed non-price characteristics of product  $j$ ; <sup>12</sup>  $\phi_c^x$  is a vector of consumer-specific marginal utilities (assumed to vary randomly across consumers) associated with non-price characteristics in

<sup>12</sup> Non-price product characteristic variables in  $x_j$  include: (1) Opres\_out; (2) Interstops; (3) Inconvenience; (4) Traditional Codeshare; (5) Virtual Codeshare; (6) Percent Traditional for Airline; (7) Percent Virtual for Airline; (8) quarter fixed effects; (9) year fixed effects; (10) ticketing carrier fixed effects; (11) market origin fixed effects; and (12) market destination fixed effects.

$x_j$ ;  $p_j$  is the price the consumer must pay to obtain product  $j$ ;  $\phi_c^p$  is the consumer-specific marginal utility of price, which is assumed to vary randomly across consumers;  $\xi_j$  capture product characteristics that are observed by consumers and airlines, but not observed by us the researchers; and  $\varepsilon_{cj}^d$  is a mean-zero random component of utility.

The random coefficients vary across consumers based on the following specification:

$$\begin{pmatrix} \phi_c^p \\ \phi_c^x \end{pmatrix} = \begin{pmatrix} \phi^p \\ \phi^x \end{pmatrix} + \begin{pmatrix} \phi_p^v & 0 & 0 & 0 & 0 \\ 0 & \phi_1^v & 0 & 0 & 0 \\ 0 & 0 & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & 0 \\ 0 & 0 & 0 & 0 & \phi_L^v \end{pmatrix} \times \begin{pmatrix} v_{cp} \\ v_{c1} \\ \cdot \\ \cdot \\ v_{cL} \end{pmatrix} \quad (2)$$

where  $\phi^p$  is the mean (across consumers) marginal utility of price;  $\phi^x$  is a vector of mean marginal utilities for respective non-price product characteristics;  $\phi^v = (\phi_p^v, \phi_1^v, \dots, \phi_L^v)$  is a set of parameters that measure variation across consumers in random taste shocks for respective product characteristics; and  $v_c = (v_{cp}, v_{c1}, \dots, v_{cL})$  is a set of consumer  $c$ 's random taste shocks for respective product characteristics. We assume that  $v_c$  follows a standard normal probability distribution across consumers.

Following much of the literature on discrete choice demand model [see Nevo (2000)], we assume that  $\varepsilon_{cj}^d$  in equation (1) is governed by an independent and identically distributed extreme value probability density. As such, the probability that product  $j$  is chosen, or equivalently the predicted market share of product  $j$  is:

$$s_j(x_j, p_j, \xi_j; \phi^x, \phi^p, \phi^v) = \int \frac{\exp(\delta_j + \mu_{cj})}{1 + \sum_k^J \exp(\delta_k + \mu_{ck})} dG(v) \quad (3)$$

where  $\delta_j = x_j \phi^x + \phi^p p_j + \xi_j$  is the mean utility obtained across consumers who choose product  $j$ ;  $\mu_{cj} = \phi_p^v p_j v_{cp} + \sum_{l=1}^L \phi_l^v x_{jl} v_{cl}$  is a consumer-specific deviation from the mean utility level; and  $G(\cdot)$  is the standard normal distribution function for the taste shocks. Since there is no closed-form solution for the integral in equation (3), this integral is approximated numerically using random draws from  $G(v)$ .<sup>13</sup>

<sup>13</sup> We use 200 random draws from  $G(\cdot)$  for the numerical approximation of  $s_j(\cdot)$ .

The quantity demand for product  $j$  is simply specified to equal to the probability that product  $j$  is chosen times the total number of potential consumers,  $POP$ :

$$d_j = s_j(p, x, \xi; \Phi^d) \times POP \quad (4)$$

where  $\Phi^d = (\phi^p, \phi^x, \phi^v)$  is the vector of demand parameters to be estimated.

### 3.2 Supply

The ticketing carrier of a codeshare product markets and sets the final price for the round-trip ticket and compensates the operating carrier for operating services provided. Unfortunately for researchers, partner airlines do not publicize details of how they compensate each other on their codeshare flights. Therefore, our challenge as researchers is to specify a modeling approach that captures our basic understanding of what is commonly known about how a codeshare agreement works without imposing too much structure on a contracting process about which we have few facts. As such, we follow the modeling approach outlined in Chen and Gayle (2007) and Gayle (2013).

Chen and Gayle (2007) and Gayle (2013) suggest that for modeling purposes a codeshare agreement can be thought of as a privately negotiated pricing contract between partners  $(w, \Gamma)$ , where  $w$  is a per-passenger price the ticketing carrier pays over to an operating carrier for transporting the passenger, while  $\Gamma$  represents a potential lump-sum transfer between partners that determines how the joint surplus is distributed. For the purposes of this paper we do not need to econometrically identify an equilibrium value of  $\Gamma$ , but in describing the dynamic part of the model, we do show where  $\Gamma$  enters the model.

Suppose the final price of a codeshare product is determined within a sequential price-setting game, where in the first stage of the sequential process the operating carrier sets price,  $w$ , for transporting a passenger using its own plane(s), and privately makes this price known to its partner ticketing carrier. In the second stage, conditional on the agreed upon price  $w$  for services supplied by the operating carrier, the ticketing carrier sets the final round-trip price  $p$  for the codeshare product. The final subgame in this sequential price-setting game is played between ticketing carriers, and produces the final ticket prices observed by consumers.

Each ticketing carrier  $i$  solves the following profit maximization problem:

$$\begin{aligned} & \max_{p_{jmt} \forall j \in B_{imt}} VP_{imt} \\ & = \max_{p_{jmt} \forall j \in B_{imt}} \left[ \sum_{j \in B_{imt}} (p_{jmt} - mc_{jmt}) q_{jmt} \right] \end{aligned} \quad (5)$$

where  $VP_{imt}$  is the variable profit carrier  $i$  obtains in market  $m$  during period  $t$  by offering the set of products  $B_{imt}$  to consumers,  $q_{jmt}$  is the quantity of tickets for product  $j$  sold in market  $m$ ,  $p_{jmt}$  is the price of product  $j$ , and  $mc_{jmt}$  is the effective marginal cost incurred by ticketing carrier  $i$  from offering product  $j$ .

Let  $f = 1, \dots, F$  index the corresponding operating carriers. If product  $j$  is a traditional codeshare product, then  $mc_{jmt} = c_{jmt}^i + w_{jmt}^f$ , where  $c_{jmt}^i$  is the marginal cost that ticketing carrier  $i$  incurs by using its own plane to provide transportation services on some segment(s) of the trip needed for product  $j$ , while  $w_{jmt}^f$  is the price ticketing carrier  $i$  pays to operating carrier  $f$  for its transportation services on the remaining trip segment(s). If instead product  $j$  is a virtual codeshare product, then  $mc_{jmt} = w_{jmt}^f$ , where  $w_{jmt}^f$  is the price the ticketing carrier pays to operating carrier  $f$  for its exclusive transportation services in the provision of product  $j$ .<sup>14</sup> Last, if product  $j$  is a pure online product, then  $mc_{jmt} = c_{jmt}^i$ . In the case of a pure online product, the ticketing carrier is also the sole operating carrier of product  $j$ , i.e.,  $i = f$ .

In equilibrium, the amount of product  $j$  an airline sells is equal to the quantity demanded, that is,  $q_{jmt} = d_{jmt} = s_{jmt}(p, x, \xi; \Phi^d) \times POP$ . The optimization problem in (5) yields the following set of  $J$  first-order conditions – one for each of the  $J$  products in the market:

$$\sum_{k \in B_i} (p_k - mc_k) \frac{\partial s_k}{\partial p_j} + s_j = 0 \text{ for all } j = 1, \dots, J \quad (6)$$

We have dropped the market and time subscripts in equation (6) only to avoid a clutter of notation. The set of first-order conditions can be represented in matrix notation as follows:

---

<sup>14</sup> The implicit assumption here is that the ticketing carrier of a virtual codeshare product only incurs fixed expenses in marketing the product to potential passengers.

$$(\Omega.*\Delta) \times (p - mc) + s = 0 \quad (7)$$

where  $p$ ,  $mc$ , and  $s$  are  $J \times 1$  vectors of product prices, marginal costs, and predicted product shares respectively,  $\Omega$  is a  $J \times J$  matrix of appropriately positioned zeros and ones that capture ticketing carriers' "ownership" structure of the  $J$  products in a market,  $.*$  is the operator for element-by-element matrix multiplication, and  $\Delta$  is a  $J \times J$  matrix of own and cross-price effects, where element  $\Delta_{jk} = \frac{\partial s_k}{\partial p_j}$ . Since for purposes of the model the ticketing carrier is considered the "owner" of a product, in the discussion that follows, "airline" is synonymous with ticketing carrier.

Equation (7) can be re-arranged to yield a vector of product markups:

$$mkup(x, \xi; \Phi^d) = p - mc = -(\Omega.*\Delta)^{-1} \times s \quad (8)$$

Based on equations (5) and (8), and with estimates of demand parameters in hand,  $\widehat{\Phi}^d$ , firm-level variable profit can be computed by:

$$VP_{imt} = \sum_{j \in B_{imt}} mkup_{jmt}(x, \xi; \widehat{\Phi}^d) q_{jmt} \quad (9)$$

### 3.3 Dynamic Entry/Exit Game

In the dynamic entry/exit game, each airline chooses markets in which to be active during specific time periods. An airline being active in a market means that the airline actually sells products to consumers in the market even though a subset of those products may use the operating services of the airline's codeshare partner carriers. Each airline optimally makes this decision in order to maximize its expected discounted stream of profit:

$$E_t \left( \sum_{r=0}^{\infty} \beta^r \Pi_{im,t+r} \right) \quad (10)$$

where  $\beta \in (0,1)$  is the discount factor, and  $\Pi_{im,t+r}$  is the per-period profit of airline  $i$  in origin-destination market  $m$ . Airline  $i$ 's per-period profit is:

$$\Pi_{imt} = a_{im,t-1} VP_{imt} - a_{imt} F_{imt} \quad (11)$$

where  $VP_{imt}$  represents the variable profit of airline  $i$  in origin-destination market  $m$  during period  $t$  that is computed from the previously discussed differentiated products Nash price-

setting game;  $a_{im,t-1}$  is a zero-one indicator that equals 1 only if airline  $i$  had made the decision in period  $t-1$  to be active in market  $m$  during period  $t$ , therefore  $a_{imt} = 1$  only if airline  $i$  makes decision in period  $t$  to be active in market  $m$  during period  $t+1$ ; and  $F_{imt}$  is the sum of fixed and entry costs of airline  $i$  in market  $m$  during period  $t$ .

It is important to note that the time subscript on indicator variable  $a_{imt}$  identifies the period in which the airline makes a decision regarding being active or not in a market, but the decision does not become effective until the subsequent period. In other words, an airline that is active in a market during period  $t$  and earning variable profit  $VP_{imt}$  is a consequence of the airline making this decision in period  $t-1$  to be active in period  $t$ . This is commonly referred to as a time-to-build assumption since it is assumed that once a decision is made to become active in a market, it will take a full period for the airline to implement the necessary plans for actual operations. Note however, that the time subscript on  $VP_{imt}$  identifies the period in which the variable profit is earned, and the time subscript on  $F_{imt}$  identifies the period in which the relevant costs are incurred and paid.

Let  $F_{imt}$  be specified as:

$$F_{imt} = FC_{imt} + \epsilon_{imt}^{FC} + (1 - a_{im,t-1})[EC_{imt} + \epsilon_{mt}^{Trad} + \epsilon_{mt}^{Virt} + \epsilon_{imt}^{EC}] \quad (12)$$

where  $FC_{imt}$  represents the deterministic part of per-period fixed cost of operating flights in origin-destination market  $m$ . The component  $\epsilon_{imt}^{FC}$  represents a private firm-idiosyncratic shock to airline  $i$ 's fixed cost. The fixed cost  $FC_{imt} + \epsilon_{imt}^{FC}$  is paid now only if the airline decides to be active in market  $m$  next period, i.e., if  $a_{imt} = 1$ .

The entry cost  $EC_{imt} + \epsilon_{mt}^{Trad} + \epsilon_{mt}^{Virt} + \epsilon_{imt}^{EC}$  has four components;  $EC_{imt}$  is a deterministic component, while  $\epsilon_{mt}^{Trad}$ ,  $\epsilon_{mt}^{Virt}$ , and  $\epsilon_{imt}^{EC}$  represent shocks to entry cost. Shocks  $\epsilon_{mt}^{Trad}$  and  $\epsilon_{mt}^{Virt}$  only vary by market and time and are observed by firms, but not by us the researchers, while  $\epsilon_{imt}^{EC}$  represents a private firm-idiosyncratic shock to airline  $i$ 's entry cost. The entry cost is paid only when the airline is not active in market  $m$  at period  $t$  but it decides to be active in the market next period, i.e., if  $a_{im,t-1} = 0$  and  $a_{imt} = 1$ .

Let the composite private firm-idiosyncratic shock to airline  $i$ 's fixed and entry costs be denoted by  $\varepsilon_{imt}$ . Based on equation (12),  $\varepsilon_{imt} = \epsilon_{imt}^{FC} + (1 - a_{im,t-1})\epsilon_{imt}^{EC}$ . We assume that the composite private information shock,  $\varepsilon_{imt}$ , is independently and identically



distributed over firms, markets and time, and has a type 1 extreme value probability distribution function.

The deterministic portions of fixed and entry costs are specified as:

$$FC_{imt} = \theta_0^{FC} + \theta_1^{FC} Pres_{imt} \quad (13)$$

$$\begin{aligned} EC_{imt} = & \theta_0^{EC} + \theta_1^{EC} Pres_{imt} \\ & + \theta_2^{EC} Percent\_Trad_{mt} \\ & + \theta_3^{EC} Percent\_Virtual_{mt} \\ & + \theta_4^{EC} Percent\_Trad_{mt} \times Southwest \\ & + \theta_5^{EC} Percent\_Virtual_{mt} \times Southwest \\ & + \theta_6^{EC} Percent\_Trad_{mt} \times Other\_lcc \\ & + \theta_7^{EC} Percent\_Virtual_{mt} \times Other\_lcc \end{aligned} \quad (14)$$

where  $Pres_{imt}$  is the mean across size-of-presence variables  $Opres\_in$  and  $Dpres\_out$  for airline  $i$  at the endpoint cities of market  $m$ ; <sup>15</sup>  $Percent\_Trad_{mt}$  is the percent of products in market  $m$  during period  $t$  that are traditional codeshare;  $Percent\_Virtual_{mt}$  is the percent of products in market  $m$  during period  $t$  that are virtual codeshare;  $Southwest$  is a zero-one dummy variable that equals to one only if the airline is Southwest;  $Other\_lcc$  is a zero-one dummy variable that equals to one for low-cost carriers other than Southwest; and  $\{\theta_0^{FC}, \theta_1^{FC}, \theta_0^{EC}, \theta_1^{EC}, \theta_2^{EC}, \theta_3^{EC}, \theta_4^{EC}, \theta_5^{EC}, \theta_6^{EC}, \theta_7^{EC}\}$  is the set of structural parameters to be estimated.

$\theta_0^{FC}$  measures mean (across airlines, markets and time) fixed cost, while  $\theta_1^{FC}$  measures the effect that size of an airline's city presence has on fixed cost. The mean recurrent fixed cost parameter  $\theta_0^{FC}$  may comprise fixed expenses incurred by a ticketing carrier when the carrier markets a codeshare product to potential consumers. In our previous discussion we define  $(w, \Gamma)$  as a privately negotiated codeshare contract between partner carriers, where  $w$  is a per-passenger price the ticketing carrier pays an operating carrier for transporting the passenger, while  $\Gamma$  represents a potential lump-sum transfer between partners that determines how the joint surplus is distributed. It was shown that  $w$  enters the effective marginal cost of the ticketing carrier. However, the lump-sum transfer

---

<sup>15</sup> As we previously defined in the section, *Definitions and Data*,  $Opres\_in$  is a variable that counts the number of different cities that the airline provides nonstop flights from, going into the origin city of the market, while variable  $Dpres\_out$  counts the number of distinct cities that the airline has nonstop flights to, leaving from the destination city.

between partners,  $\Gamma$ , is nested in  $\theta_0^{FC}$ , but we do not attempt to separately identify  $\Gamma$  since knowing its value is not essential for the purposes of our paper.

$\theta_0^{EC}$  measures mean (across airlines, markets and time) entry cost – we also allow mean entry cost to differ by the three carrier-types we consider (*Legacy*, *Southwest* and *Other low cost carriers*), in which case  $\theta_0^{EC}$  would be a vector containing three parameters. The coefficient on  $Pres_{imt}$  is  $\theta_1^{EC}$ , which measures the effect that size of an airline's city presence has on entry cost. Parameter  $\theta_2^{EC}$  measures the impact that traditional codesharing between incumbent airlines have on market entry costs of legacy carriers that are potential entrants to the relevant market, while  $\theta_2^{EC} + \theta_4^{EC}$  captures the impact on Southwest's market entry cost, and  $\theta_2^{EC} + \theta_6^{EC}$  captures the impact on other low-cost carrier's market entry cost. Hence, testing whether  $\theta_4^{EC}$  and  $\theta_6^{EC}$  are positive tells us whether traditional codesharing among incumbents increases the barrier to entry for Southwest or other low-cost carriers, respectively. In the case of virtual codesharing, parameter  $\theta_3^{EC}$  measures the impact that virtual codesharing between incumbent airlines have on market entry costs of legacy carriers that are potential entrants to the relevant market, while  $\theta_3^{EC} + \theta_5^{EC}$  captures the impact on Southwest's market entry cost, and  $\theta_3^{EC} + \theta_7^{EC}$  captures the impact on other low-cost carrier's market entry cost. Hence, testing whether  $\theta_5^{EC}$  and  $\theta_7^{EC}$  are positive tells us whether virtual codesharing among incumbents increases the barrier to entry for Southwest or other low-cost carriers, respectively.

$Percent\_Trad_{mt}$  and  $Percent\_Virtual_{mt}$  measure the extent of codesharing that takes place in a market. For the sake of not making the model overly complex and difficult to estimate, we chose not to explicitly model airlines' optimizing decision of whether or not to codeshare in a market. However, it is reasonable to conjecture that airlines' optimizing decision of whether or not to codeshare is influenced by the effective cost an airline faces to use its own planes to begin providing service in the market (part of its market entry cost). This further suggests that shocks to market entry cost that are unobserved to us,  $\epsilon_{mt}^{Trad}$  and  $\epsilon_{mt}^{Virt}$ , are likely to influence  $Percent\_Trad_{mt}$  and  $Percent\_Virtual_{mt}$  respectively. As such, we formally specify the following equations:

$$Percent\_Trad_{mt} = Z_{mt}\gamma + \epsilon_{mt}^{Trad} \quad (15)$$

$$Percent\_Virtual_{mt} = Z_{mt}\lambda + \epsilon_{mt}^{Virt} \quad (16)$$

where  $Z_{mt}$  is a matrix of variables that influence the extent of traditional and virtual codesharing that takes place in a market;  $\gamma$  and  $\lambda$  are vectors of parameters associated with these variables in equations (15) and (16) respectively; while  $\epsilon_{mt}^{Trad}$  and  $\epsilon_{mt}^{Virt}$  are assumed to be independently and identically distributed normal random variables with mean zero and standard deviations  $\sigma^{Trad}$  and  $\sigma^{Virt}$  respectively. Therefore, the model accounts for the endogeneity of variables  $Percent\_Trad_{mt}$  and  $Percent\_Virtual_{mt}$  in the entry cost function.

The variables we include in  $Z_{mt}$  are: (1) the geometric mean of the origin city and destination city populations ( $POP$ ), which is a measure of market size; (2) nonstop flight distance between the origin and destination; (3) one-period lag of the Herfindahl-Hirschman Index (HHI) computed based on the relative sizes of airlines' presence at the market endpoint cities, where an airline's size of city presence is measured by the previously defined variables,  $Opres\_in$  and  $Dpres\_out$ ; <sup>16</sup> (4) origin city fixed effects; (5) destination city fixed effects; and (6) quarter fixed effects.

Our specified equations do not include a firm-specific component of fixed cost and entry cost for two reasons. First, estimation of the dynamic model is computationally quite intensive, and convergence is difficult to achieve when the number of parameters being optimized over is large. Even with the model restricted to 10 parameters and four quarters of data, optimization took approximately seven days of continuously running the computer program. Second, even without firm-specific parameters, the fixed and entry cost functions do capture some heterogeneity across firms via the firm-specific variable  $Pres_{imt}$ .

### ***Reducing the dimensionality of the dynamic game***

From the previously discussed Nash price-setting game, firm-level variable profit is:  $VP_{imt}(x, \xi; \Phi^d) = \sum_{j \in B_{imt}} mkup_{jmt}(x, \xi; \Phi^d) * q_{jmt}$ . Let

$$R_{imt}^* = a_{im,t-1} VP_{imt} \quad (18)$$

Note that  $(x, \xi)$  are state variables that are needed in the dynamic entry/exit game. As

---

<sup>16</sup>  $Opres\_in$  is a variable that counts the number of different cities that the airline provides nonstop flights from, going into the origin city of the market, while variable  $Dpres\_out$  counts the number of distinct cities that the airline has nonstop flights to, leaving from the destination city.

pointed out and discussed in Aguirregabiria and Ho (2012),  $R_{imt}^*$  aggregates these state variables in an economically meaningful way so that these state variables can enter the dynamic game through  $R_{imt}^*$ . Therefore, as a simplifying assumption, Aguirregabiria and Ho (2012) recommend treating  $R_{imt}^*$  as a firm-specific state variable, rather than treating  $x$  and  $\xi$  as separate state variables. Though restrictive, this simplifying assumption substantially reduces the dimensionality of the state space to make estimation feasible. The payoff-relevant information of firm  $i$  in market  $m$  is:

$$y_{imt} \equiv \{s_{imt}, R_{imt}^*, Pres_{imt}, Percent\_Trad_{mt}, Percent\_Virtual_{mt}, \epsilon_{mt}^{Trad}, \epsilon_{mt}^{Virt}\} \quad (19)$$

where  $s_{imt} = a_{im,t-1}$ .

Each airline has its own vector of state variables,  $y_{imt}$ , and airlines take into account these variables when making decisions. It might seem that each airline does not take into account the strategies that other airlines adopt. However, an airline's vector of state variables,  $y_{imt}$ , depends on previous period entry and exit decisions of other airlines. For example, the variable profit state variable,  $R_{imt}^*$ , depends on competition from other incumbents currently in the market, which implies that this state variable depends on the previous period's entry/exit decisions of other airlines. Accordingly, our entry/exit model incorporates dynamic strategic interactions among airlines.

Let  $\sigma \equiv \{\sigma_{im}(y_{imt}, \epsilon_{imt}), i = 1, 2, \dots, N; m = 1, 2, \dots, M\}$  be a set of strategy functions, one for each airline.  $\sigma$  is a Markov Perfect Equilibrium (MPE) if the profile of strategies in  $\sigma$  maximizes the expected value of airline  $i$  at every state  $(y_{imt}, \epsilon_{imt})$  given the opponent's strategy.

### ***Value Function and Bellman Equation***

For notational convenience, we drop the market subscript. Let  $V_i^\sigma(y_t, \epsilon_{it})$  be the value function for airline  $i$  given that the other airlines behave according to their respective strategies in  $\sigma$ . The value function is the unique solution to the Bellman equation:

$$\begin{aligned}
& V_i^\sigma(y_t, \varepsilon_{it}) \\
& = \text{Max}_{a_{it} \in \{0,1\}} \{ \Pi_{it}^\sigma(a_{it}, y_t) - \varepsilon_{it} * a_{it} \\
& \quad + \beta \int V_i^\sigma(y_{t+1}, \varepsilon_{it+1}) dG_i(\varepsilon_{it+1}) F_i^\sigma(y_{t+1} | a_{it}, y_t) \}
\end{aligned} \tag{20}$$

where  $\Pi_{it}^\sigma(a_{it}, y_t)$  and  $F_i^\sigma(y_{t+1} | a_{it}, y_t)$  are the expected one-period profit and expected transition of state variables, respectively, for airline  $i$  given the strategies of the other airlines. The profile of strategies in  $\sigma$  is a MPE if, for every airline  $i$  and every state  $(y_t, \varepsilon_{it})$ , we have:

$$\begin{aligned}
& \sigma_i(y_t, \varepsilon_{it}) \\
& = \text{argmax}_{a_{it}} \{ \Pi_{it}^\sigma(a_{it}, y_t) - \varepsilon_{it} * a_{it} \\
& \quad + \beta \int V_i^\sigma(y_{t+1}, \varepsilon_{it+1}) dG_i(\varepsilon_{it+1}) F_i^\sigma(y_{t+1} | a_{it}, y_t) \}
\end{aligned} \tag{21}$$

The transition rules we use for state variables are described in Appendix A. In Appendix B we illustrate that the MPE can also be represented as a vector of conditional choice probabilities (CCPs) that solves the fixed point problem  $\mathbf{P} = \psi(\mathbf{P}, \theta)$ , where  $\mathbf{P} = \{P_i(\mathbf{y}): \text{for every firm and state } (i, \mathbf{y})\}$ .  $\mathbf{P} = \psi(\mathbf{P}, \theta)$  is a vector of best response probability mapping, where  $\psi(\cdot)$  is the CDF of the type 1 extreme value distribution.

## 4. Estimation and Results

### 4.1 Estimation of demand

The demand model is estimated using Generalized Methods of Moments (GMM). Following Berry (1994), Berry, Levinsohn, and Pakes (1995), and Nevo (2000), we can solve for  $\xi_j$  as a function of demand parameters and the data, where  $\xi_j = \delta_j - x_j \phi^x - \phi^p p_j$ .  $\xi_j$  is the error term used to formulate the GMM optimization problem:

$$\min_{\phi^x, \phi^p, \phi^v} \xi' Z^d W Z^{d'} \xi \tag{22}$$

where  $Z^d$  is the matrix of instruments that are assumed orthogonal to the error vector  $\xi$ , while  $W$  is the standard weighting matrix,  $W = \left[ \frac{1}{n} Z^{d'} \xi \xi' Z^d \right]^{-1}$ . Since parameters  $\phi^p$  and  $\phi^x$  enter the error term linearly, we can restructure the GMM

optimization problem in (22) such that the search to minimize the objective function,  $\xi'Z^dWZ^{d'}\xi$ , is done exclusively over parameter vector  $\phi^v$ , i.e., the GMM optimization problem reduces to  $\min_{\phi^v} \xi'Z^dWZ^{d'}\xi$ . Once the optimization over  $\phi^v$  is complete, we can recover estimates of  $\phi^p$  and  $\phi^x$ .<sup>17</sup>

*Percent Traditional for Airline* and *Percent Virtual for Airline* are two of the non-price product characteristic variables in  $x_j$ . Among the codeshare products in a market, recall that variables *Percent Traditional for Airline* and *Percent Virtual for Airline* measure the percentage of these products of a given codeshare type an airline offers for sale to consumers. Since airlines optimally choose the extent to which to codeshare with others in a market, it is possible that these codeshare variables are correlated with shocks to demand captured in  $\xi_j$ , making *Percent Traditional for Airline* and *Percent Virtual for Airline* endogenous in the demand model. In addition, it is well-known that  $p_j$  is correlated with  $\xi_j$ . Therefore, our estimation of the demand model takes into account the endogeneity of  $p_j$ , *Percent Traditional for Airline* and *Percent Virtual for Airline*. Specifically, instruments for these three variables are included in  $Z^d$ .

### ***Instruments for endogenous variables in demand model***

To obtain a set of valid instruments for price, we exploit the fact that the menu of products offered by airlines in a market is predetermined at the time of shocks to demand. Furthermore, the non-price characteristics of an airline's products are primarily determined by the route network structure of the airline, and unlike price, this network structure is not routinely and easily changed during a short period of time, which mitigates the influence of demand shocks on the menu of products offered and their associated non-price characteristics. As such, one set of product price instruments we use are: (1) the squared deviation of a product's itinerary distance from the average itinerary distance of competing products offered by other airlines; and (2) the number of competing products offered by other airlines, where these competing products have number of intermediate stops equivalent to the product in question. The rationale for these instruments is that they are

---

<sup>17</sup> For details of this estimation algorithm of a random coefficients logit model, see Nevo (2000).

measures of the degree of competition that a product faces, which affects the size of a product's markup.

Similar to Villas-Boas (2007), we also exploit the time dimension of the data to construct another set of instruments for price. Since jet fuel price<sup>18</sup> vary over the time span of the data, these changes in jet fuel price are likely to affect airlines' marginal cost differently because airlines differ in the intensity with which they use fuel owing to differences in their route network structure and the size distribution of their aircraft fleet. In addition, it is reasonable to assume that airlines do not routinely change their aircraft fleet with each change in jet fuel price. Given that an airline's marginal cost is correlated with its product price, and we assume that shocks to jet fuel price are uncorrelated with  $\xi_j$ , then the following are another set of valid instruments we use for air travel product price: (1) itinerary distance flown; (2) interaction of jet fuel price with itinerary distance flown; and (3) interaction of jet fuel price with operating carrier dummies.

For the variables *Percent Traditional for Airline* and *Percent Virtual for Airline*, we adopt two instruments: (i) one-period lag of the squared deviation of an airline's size presence at the market endpoint cities from the average size presence of other airlines at the market endpoints; and (ii) the interaction of (i) with nonstop flight distance. The size of an airline's presence at the market endpoints is computed by averaging across variables *Opres\_in* and *Dpres\_out*, which are variables we defined in the *Definitions and Data* section. An airline's measures of *Opres\_in* and *Dpres\_out* at the endpoints of a market are more determined by the airline's extended route network structure rather than features of the given origin-destination market. Therefore, it is reasonable to assume that *Opres\_in* and *Dpres\_out* are uncorrelated with  $\xi_j$ . In addition, lower presence for an airline at the endpoints of a market makes it more likely that the airline will codeshare with others that are already serving the market. So *Opres\_in* and *Dpres\_out* are in principle good instruments for *Percent Traditional for Airline* and *Percent Virtual for Airline*. Last, we allow the influence of an airline's size of presence at the market endpoints on its extent of market codesharing to depend on the nonstop flight distance of the market. This explains the rationale for instrument (ii).

---

<sup>18</sup> The jet fuel price we use is U.S. Gulf Kerosene-Type Jet Fuel Spot Price FOB from the U.S. Energy Information Administration.

## 4.2 Results from demand estimation

We begin by estimating a standard logit specification of the demand model, which is more restrictive than the random coefficients logit demand model outlined previously in the sense that the standard logit model does not allow marginal utilities for product characteristics to vary across consumers. Table 7 reports both Ordinary Least Square (OLS) and Two-stage Least Squares (2SLS) estimates of the standard logit model. First, focusing on coefficient estimates for the variable, *Fare* ( $p_j$ ), we find that even though the sign of the OLS and 2SLS coefficient estimates on *Fare* is consistent with intuition, there is a large difference in the size of the two coefficient estimates. The OLS versus the 2SLS coefficient estimates on variables *Percent Traditional for Airline* and *Percent Virtual for Airline* are also contrasting in magnitudes. This preliminary evidence suggests that estimates in the OLS regression are biased and inconsistent and thus instruments are needed for these potentially endogenous variables.

To formally confirm that variables *Fare*, *Percent Traditional for Airline* and *Percent Virtual for Airline* are endogenous, we perform a Hausman exogeneity test. The result of the Hausman test shown in Table 7 easily rejects the exogeneity of these three variables at conventional levels of statistical significance. To evaluate whether the instruments have statistically significant explanatory power of variations in the endogenous variables, we estimate first-stage reduced-form regressions for each of the endogenous variables. When *Fare* is the dependent variable in the reduced-form regression, R-squared is 0.321, but when *Percent Traditional for Airline* and *Percent Virtual for Airline* are dependent variables, the R-squared values are respectively 0.331 and 0.327. *F*-tests of the joint statistical significance of the instruments in these first-stage reduced-form regressions yield *F*-statistic values of  $F(46, 434144) = 1717.50$ ,  $F(46, 434144) = 115.70$ , and  $F(46, 434144) = 157.91$  for the *Fare*, *Percent Traditional for Airline* and *Percent Virtual for Airline* regressions, respectively. In each case the *p*-value for the *F*-statistic is 0.000, suggesting that the instruments do have statistically significant explanatory power of variations in each endogenous variable.



**Table 7**  
**Demand Estimation**

	Standard Logit Model				Random Coefficients Logit Model	
	OLS		2SLS		GMM	
	Estimates	Robust Std. Error	Estimates	Robust Std. Error	Estimates	Robust Std. Error
<b>Means (<math>\phi^p</math> and <math>\phi^x</math>)</b>						
Fare (in thousand \$)	-0.9290***	0.0343	-8.2570***	0.1669	-19.12***	0.7401
Opres_out	0.3973***	0.0082	0.1841***	0.0182	0.473***	0.0314
Interstops	-1.4588***	0.0060	-1.471***	0.0084	-1.495***	0.0219
Inconvenience	-0.9491***	0.0064	-0.9265***	0.0106	-0.760***	0.0186
Traditional Codeshare	-0.6827***	0.0116	-5.1325***	0.1559	-6.849***	0.2346
Virtual Codeshare	-0.8867***	0.0127	-0.7232***	0.0923	-1.122***	0.1344
Percent Traditional for Airline <sup>a</sup>	0.1325***	0.0144	9.0277***	0.3084	12.35***	0.4624
Percent Virtual for Airline <sup>a</sup>	-0.1147***	0.0115	-1.0889***	0.1683	-0.643***	0.2426
Spring	0.1111***	0.0045	0.1570***	0.0067	0.413***	0.0137
Summer	0.0826***	0.0045	0.1244***	0.0066	0.327***	0.0129
Fall	0.0638***	0.0045	0.0829***	0.0066	0.222***	0.0106
Constant	-3.0333***	0.0219	-2.0197***	0.0376	-4.164***	0.1135
Ticketing carrier fixed effects		YES		YES		YES
Year fixed effects		YES		YES		YES
Market Origin fixed effects		YES		YES		YES
Market Destination fixed effects		YES		YES		YES
<b>Taste variation (<math>\phi^v</math>)</b>						
Constant	-	-	-	-	3.225***	0.0844
Price (in thousand \$)	-	-	-	-	0.346	9.847
Interstops	-	-	-	-	0.099	0.2096
R-squared		0.4417		-		
Value of GMM objective function		-		-		51365.004
<b>Test of endogeneity</b>						
Ho: variables are exogenous						
Durbin-Wu-Hausman test:	$\chi^2(3) = 5293.29***$		Prob_Value = 0.0000			
Robust regression F-test:	$F(3, 434184) = 1818.82***$		Prob_Value = 0.0000			

\*\*\* indicates statistical significance at 1%. <sup>a</sup> variable is measured in values between zero and one when variable is used in demand estimation. For the Standard Logit Model, the well-known linear equation used for estimating the parameters is:  $\ln(S_j) - \ln(S_0) = x_j\phi^x + \phi^p p_j + \xi_j$ , where  $S_j = q_j / POP$  is the observed share of product  $j$ ,  $S_0 = 1 - \sum_{j=1}^J S_j$  is the observed share of the outside good, and  $\xi_j$  is the error term of the equation.

The following discussion of demand regression results in Table 7 focuses on the less restrictive random coefficients logit model. The upper panel of the table reports the mean marginal (dis)utilities for each product characteristic ( $\phi^p$  and  $\phi^x$ ), while the lower panel of the table reports the parameter estimates that measure taste variation across consumers for respective product characteristics ( $\phi^v$ ).

As expected, the coefficient estimate on *Fare* is negative, implying that higher prices are associated with lower levels of utility, *ceteris paribus*. In other words, all else equal, passengers prefer cheaper air travel products.

The coefficient estimate on *Opres\_out* is positive. This result is consistent with our priors, and suggests that travelers prefer to fly with airlines, all else equal, that offer services to more destinations from the travelers' origin city. This estimated effect is possibly in part due to the benefits of frequent-flyer programs. Travelers are more likely to hold frequent-flyer membership with the airline they believe they are most likely to use in the future, and it is reasonable for a passenger to conjecture that they will most often use the airline that offers service to a relatively large number of destinations from the passenger's origin city. Once the passenger becomes invested in the airline's frequent-flyer program, this helps reinforce the passenger's loyalty to the airline.

The coefficient estimate on *Interstops* is negative, implying that consumers most prefer nonstop flights between their origin and destination compared to travel itineraries that require intermediate stops. This is reasonable since passengers should prefer the most convenient travel itinerary from origin to destination. In addition, the coefficient estimate on *Inconvenience* is negative. This intuitively makes sense as well since, for any given number of intermediate stops, passengers prefer the most direct routing to the destination.

The coefficient estimate on *Interstops* divided by the coefficient estimate on the airfare variable, multiplied by 1000, is 78.19, which suggests that, on average, consumers are willing to pay up to \$78.19 per intermediate stop to avoid air travel products with intermediate stops. For example, if the price of an air travel product that requires one intermediate stop between an origin and destination is \$200, then the coefficient estimates suggest that a typical consumer is willing to purchase a nonstop product in this market at a price of \$278 or lower, *ceteris paribus*.

The *Traditional Codeshare* dummy variable has a negative coefficient estimate, implying that a traditional codeshare product makes passengers' utility lower relative to a pure online product. A likely reason is that the flight itinerary for a pure online product is typically very streamlined because an airline can better organize its own flights and schedules to minimize layover time, as well as efficiently organize its own gates at airports. Even though codeshare partners try to streamline flights across carriers to minimize layover times and facilitate smoother connections, the negative coefficient estimate on the *Traditional Codeshare* variable suggests that this process has not achieved parity with pure online products [Gayle (2013)]. The ratio of the coefficient estimates on the *Traditional Codeshare* and air fare variables suggest that a typical consumer is willing to pay up to \$358.26 ( $= \frac{6.85}{19.12} \times 1000$ ) extra to obtain a pure online product compared to an otherwise equivalent traditional codeshare product, *ceteris paribus*.

The *Virtual Codeshare* dummy variable has a negative coefficient estimate as well. This result suggests that passengers perceive virtual codeshare products as inferior substitutes to pure online products. For the itineraries that include virtual segments, first-class upgrades using accumulated frequent-flyer miles are not usually available [Ito and Lee (2007)]. This could explain why passengers perceive virtual codeshare products as inferior to pure online products. The ratio of the coefficient estimates on the *Virtual Codeshare* and air fare variables suggest that a typical consumer is willing to pay up to \$58.58 ( $= \frac{1.12}{19.12} \times 1000$ ) extra to obtain a pure online product compared to an otherwise equivalent virtual codeshare product, *ceteris paribus*.

So while both types of codeshare products are inferior to pure online products, the evidence suggests that consumers perceive virtual codeshare products as less inferior than traditional codeshare products. This consumer preference comparison across the two types of codeshare products make sense since virtual codeshare products do not require a passenger to switch operating airlines on the flight schedule, while traditional codeshare products do.

Note that the coefficient estimate on *Percent Traditional for Airline* is positive, suggesting that consumers tend to choose the airlines that offer more traditional codeshare products in a market. This result is consistent with the argument that airline codesharing

has a demand-increasing effect [Gayle and Brown (2012)]. A rationale for the demand-increasing effect is due to the fact that codeshare partners typically make their frequent-flyer programs reciprocal, thus allowing travelers holding frequent-flyer membership with one partner carrier to accumulate frequent-flyer points when flying with any partner carrier in the alliance. Thus the new opportunities for travelers to accumulate frequent-flyer points across partner carriers can increase demand for the codeshare partners' products. This result does not contradict with the previously discussed result suggesting that traditional codeshare products are less preferred by consumers compared to pure online products. The reason is that the coefficient estimate on the *Percent Traditional for Airline* does not capture a consumer preference comparison between traditional codeshare products and pure online products, instead the coefficient estimate capture the impact of an airline providing more traditional codeshare products compared to competing airlines.

Interestingly, the coefficient estimate on *Percent Virtual for Airline* is negative, suggesting that, unlike traditional codesharing, virtual codesharing does not have a demand-increasing effect. The evidence therefore suggests that airlines use these two types of codesharing practices to achieve different objectives. To the best of our knowledge, this paper is the first to provide rigorous formal evidence suggesting that traditional codesharing has a demand-increasing effect, but virtual codesharing does not.

The demand model yields a mean own-price elasticity estimate of -3.07. As pointed out by Oum, Gillen and Noble (1986) and Brander and Zhang (1990), a reasonable range for own-price elasticity in the airline industry is from -1.2 to -2.0. Peters (2006) study of the airline industry yields own-price elasticity estimates ranging from -3.2 to -3.6. Berry and Jia (2010) find own-price elasticity estimates ranging from -1.89 to -2.10 in their year 2006 sample, while Gayle and Wu (2015) find own-price elasticity estimates ranging from -1.65 to -2.39 in their year 2010 sample. Therefore, we are satisfied that the elasticity estimates generated from our model are reasonable and consistent with evidence in the existing literature.

As revealed by equation (8), the demand parameter estimates in Table 7 can be combined with the short-run supply-side Nash equilibrium price-setting assumption to compute product markups. Overall, mean price is \$166.35, while computed mean product markup is \$60.18. Since price minus markup yields marginal cost, then mean product

marginal cost is \$106.17. Therefore, the demand estimates combined with the short-run supply-side Nash equilibrium price-setting assumption suggest that marginal cost is, on average, approximately 64% ( $\approx \frac{\$106.17}{\$166.35} \times 100$ ) of product price.

Financial data reported by airlines are usually categorized by accounting concepts rather than economic concepts. For example, the Bureau of Transportation Statistics compiles financial data from US airlines in the "Form 41 Financial Schedule" database, and this database reports various accounting decompositions of operating expenses and operating revenues. However, nowhere in the airline database will you see data on economically relevant concepts such as marginal cost. In other words, financial data reported by airlines are typically not in a format that is directly comparable to the type of economically relevant estimates generated by the econometric model. Nevertheless, it is still worthwhile an attempt to piece together certain reported financial data, and use them to roughly validate some estimates from the econometric model.

A financial variable reported by airlines in the operating revenues section of the "Form 41 Financial Schedule" is "Transport-Related Revenues". Another useful financial variable reported by airlines in the operating expenses section of the database is "Flying Operations Expenses", which according to definitions by the Bureau of Transportation Statistics, these are "expenses incurred directly in the in-flight operation of aircraft and expenses related to the holding of aircraft and aircraft operational personnel in readiness for assignment for an in-flight status." It is reasonable to presume that number of in-flight passengers positively correlates with size of aircraft and cost of in-flight operations. Therefore, we believe that the reported "Flying Operations Expenses" is correlated with the number of in-flight passengers that need to be served, and therefore is related to in-flight variable expenses of an airline. As such, an airline's "Flying Operations Expenses" as a percentage of its "Transport-Related Revenues" should reasonably approximate what marginal cost as a percentage of price ought to be. Interestingly, for the set of airlines in our data sample, these reported financial data reveal that "Flying Operations Expenses", on average, is approximately 51% of "Transport-Related Revenues".<sup>19</sup> Therefore, it is reassuring that our

---

<sup>19</sup> These financial data are drawn for the four quarters of 2007 from the "Air Carrier Financial: Schedule P-1.2" file located in the Air Carrier Financial Reports (Form 41 Financial Data) database. This database is published and maintained by the Bureau of Transportation Statistics.

model estimates suggest that marginal cost is approximately 64% of product price. Furthermore, the supply equation in (8) suggests that if the marginal costs generated by the model are reasonable, then the product markups generated by the model should also be reasonable.

We also use the demand estimates along with equations (8) and (9) to compute quarterly market-level variable profits by airline. Equation (9) illustrates that variable profit is computed as product markup multiplied by product quantity sold. As we stated previously in the data section of the paper, the original database, before any cleaning, is only a 10% random sample of air travel tickets sold. Therefore, even though our model seems to produce reasonable levels of product markups, due to the data only capturing at most 10% of the actual number of travel tickets sold, then the magnitudes of our variable profit estimates are at most roughly 10% of actual variable profits. Variable profits are measured in constant year 1999 dollars. Overall, an airline's mean quarterly variable profit in a directional origin-destination market is \$53,984.43, while the median is \$17,270.3.

### 4.3 Estimation of Dynamic Model

The likelihood function for the dynamic model is given by,

$$L(\theta, \gamma, \lambda) = \prod_{m=1}^M \prod_{i=1}^N \prod_{t=1}^T P(\mathbf{a}_{mt} | \tilde{Z}_{imt}^P, \tilde{e}_{imt}^P, \epsilon_{mt}^{Trad}, \epsilon_{mt}^{Virt}, \theta) f(\epsilon_{mt}^{Trad} | Z_{mt}, \gamma) f(\epsilon_{mt}^{Virt} | Z_{mt}, \lambda) \quad (23)$$

where  $\mathbf{a}_{mt} = (a_{1mt}, a_{2mt}, \dots, a_{Nmt})$  is the vector of market participation actions taken by airlines in period  $t$ . Note that the likelihood function is comprised of three parts. The first part,  $P(\mathbf{a}_{mt} | \tilde{Z}_{imt}^P, \tilde{e}_{imt}^P, \epsilon_{mt}^{Trad}, \epsilon_{mt}^{Virt}, \theta)$  computes the conditional likelihood of observing the logit choice probabilities of airlines being active in markets across the sample during the time span of the data. To obtain the full unconditional likelihood, we multiply the conditional likelihood by the probabilities of observing specific values of  $\epsilon_{mt}^{Trad}$  and  $\epsilon_{mt}^{Virt}$ , where  $\epsilon_{mt}^{Trad} = Percent\_Trad_{mt} - Z_{mt}\gamma$  and  $\epsilon_{mt}^{Virt} = Percent\_Virtual_{mt} - Z_{mt}\lambda$  based on equations (15) and (16). Since we assume that  $\epsilon_{mt}^{Trad}$  and  $\epsilon_{mt}^{Virt}$  are normally distributed random variables with zero means and standard deviations  $\sigma^{Trad}$  and  $\sigma^{Virt}$  respectively, then  $f(\cdot)$  is the normal probability density function.

While joint estimation of the full set of parameters  $(\theta, \gamma, \lambda)$  is desirable due to potential efficiency gains, such joint estimation is extremely computationally demanding in this dynamic model. Fortunately, a convenient feature of the likelihood function above is that each of the three vectors of parameters in  $(\theta, \gamma, \lambda)$  is identified by separate parts of the likelihood function. Specifically,  $P(\mathbf{a}_{mt} | \tilde{Z}_{imt}^P, \tilde{e}_{imt}^P, \epsilon_{mt}^{Trad}, \epsilon_{mt}^{Virt}, \theta)$  is the part that identifies parameters in vector  $\theta$ , while  $f(\epsilon_{mt}^{Trad} | Z_{mt}, \gamma)$  and  $f(\epsilon_{mt}^{Virt} | Z_{mt}, \lambda)$  are the parts that identify parameter vectors  $\gamma$  and  $\lambda$  respectively. This implies that parameter vectors  $\gamma$  and  $\lambda$  can be separately estimated in a first step using likelihood functions  $\prod_{m=1}^M \prod_{t=1}^T f(\epsilon_{mt}^{Trad} | Z_{mt}, \gamma)$  and  $\prod_{m=1}^M \prod_{t=1}^T f(\epsilon_{mt}^{Virt} | Z_{mt}, \lambda)$  respectively. Given estimates  $\hat{\gamma}$  and  $\hat{\lambda}$ , we can compute  $f(\epsilon_{mt}^{Trad} | Z_{mt}, \hat{\gamma})$  and  $f(\epsilon_{mt}^{Virt} | Z_{mt}, \hat{\lambda})$  and use them to construct the relevant parts of  $L(\theta, \hat{\gamma}, \hat{\lambda})$  in order to estimate  $\hat{\theta}$  in a second step.

Based on the discussion above, we use the following pseudo log likelihood function to estimate parameters in vector  $\theta$ :

$$\begin{aligned}
Q(\theta, \mathbf{P}, \hat{\gamma}, \hat{\lambda}) = & \sum_{m=1}^M \sum_{i=1}^N \sum_{t=1}^T \{ a_{imt} \ln[\psi(\tilde{Z}_{imt}^P \times \theta + \tilde{e}_{imt}^P)] \\
& + (1 - a_{imt}) \ln[\psi(-\tilde{Z}_{imt}^P \times \theta - \tilde{e}_{imt}^P)] \\
& + \ln[f(\epsilon_{mt}^{Trad} | Z_{mt}, \hat{\gamma})] + \ln[f(\epsilon_{mt}^{Virt} | Z_{mt}, \hat{\lambda})] \}
\end{aligned} \tag{24}$$

where  $Q(\theta, \mathbf{P}, \hat{\gamma}, \hat{\lambda})$  is called a ‘‘pseudo’’ log likelihood function because airlines’ conditional choice probabilities (CCPs) in  $\psi(\cdot)$  are arbitrary and do not represent the equilibrium probabilities associated with  $\theta$ , where  $\theta$  is the vector of parameters in the fixed and entry cost functions previously specified in equations (13) and (14). Since the focus now is describing how  $\theta$  is estimated, in what follows we drop  $\hat{\gamma}$  and  $\hat{\lambda}$  when discussing ‘‘pseudo’’ log likelihood function  $Q(\cdot)$  only for notational convenience.

We begin by implementing the Pseudo Maximum Likelihood (PML) estimation procedure [Aguirregabiria and Ho (2012)]. The PML requires two steps. In step 1, we estimate relevant state transition equations. Appendix A describes transition rules used for state variables. In addition, nonparametric estimates of the choice probabilities  $\widehat{P}_0$  are computed in step1. These nonparametric probability estimates, along with state variables and estimated state transition probabilities, are used to compute  $\tilde{Z}_{imt}^{\widehat{P}_0}$  and  $\tilde{e}_{imt}^{\widehat{P}_0}$  as described

in Appendix B. Using  $\tilde{Z}_{imt}^{\widehat{P}_0}$  and  $\tilde{z}_{imt}^{\widehat{P}_0}$ , we are able to construct the pseudo log likelihood function,  $Q(\theta, \widehat{P}_0)$ . In step 2 of the PML estimation algorithm, the vector of parameters  $\hat{\theta}_{PML}$  is estimated by:

$$\hat{\theta}_{PML} = \arg \max_{\theta} Q(\theta, \widehat{P}_0) \quad (25)$$

This PML algorithm is simple and does not require solving for an equilibrium in the dynamic game, and thus substantially reduces computational burden. However, the two-step pseudo maximum likelihood estimator  $\hat{\theta}_{PML}$  can have a large finite sample bias [Aguirregabiria and Mira (2007)]. To achieve consistency of the parameter estimates, we follow Aguirregabiria and Mira (2002, 2007) and use as a starting point the PML parameter estimates along with the non-parametric estimates of the choice probabilities to implement the Nested Pseudo Likelihood (NPL) estimation algorithm. To assess robustness of parameter convergence in our application of the NPL estimation algorithm, we have tried starting the algorithm at several distinct initial sets of  $\theta$  and find that the NPL algorithm converged to qualitatively similar  $\theta$  on each run of the estimation algorithm. In subsequent discussion of parameter estimates from the dynamic model, we piece together certain official financial data reported by airlines, and use these data where possible to roughly validate some estimates from the dynamic model. We describe the NPL estimation algorithm in Appendix C.<sup>20</sup>

### ***Results from first-stage estimation of parameter vectors $\gamma$ and $\lambda$***

Table 8 reports the estimation results for first-stage estimation of parameter vectors  $\gamma$  and  $\lambda$ . The results suggest that more concentrated airline presence at the market endpoints (measured by variable *Lag HHI of Presence*), and longer distance between market endpoints (measured by variable *Nonstop Flight Distance*) seem to incentivize relatively higher levels of traditional codesharing, but lower levels of virtual codesharing. At a minimum we can infer from these results that airlines' choice of the most prevalent

---

<sup>20</sup> While the demand model is estimated using all three years in the data set (2005, 2006 and 2007), due to significant computational burden, we find that the dynamic entry/exit model can only feasibly be estimated using, at most, four quarters of the data. We only use data in year 2005 when estimating the dynamic entry/exit model. Even with just four quarters of data, the computer code for the dynamic entry/exit model took more than seven days of continuous running before convergence is achieved.



type of codesharing to employ in a market depends in part on certain market characteristics. Last, results of  $F$ -tests shown in the table suggest that all regressors as a group do explain variations in  $Percent\_Trad_{mt}$  and  $Percent\_Virtual_{mt}$ .

**Table 8**  
**Estimation of Linear Equations**  
**for Percent Codeshare Variables**

Variables	Dependent Variable: <i>Percent_Traditional</i>		Dependent Variable: <i>Percent_Virtual</i>	
	Coefficient Estimates ( $\gamma$ )	Standard Error	Coefficient Estimates ( $\lambda$ )	Standard Error
POP	-2.84E-08	2.32E-07	1.37E-07	2.35E-07
Nonstop flight distance	0.0016***	7.68E-05	-0.0012***	7.79E-05
Lag HHI of Presence	0.9831**	0.4001	-3.6714***	0.4056
Constant	-1.5868***	0.4384	2.6997***	0.4444
Origin fixed effects	YES		YES	
Destination fixed effects	YES		YES	
Quarter fixed effects	YES		YES	
R-squared	0.2421		0.2943	
F-test	29.60	Prob>F = 0.000	38.63	Prob>F = 0.000

\*\*\* indicates statistical significance at 1%

\*\* indicates statistical significance at 5%

Equations are estimated using ordinary least squares.

#### 4.4 Results from the dynamic model

Table 9 reports estimates of parameters in the fixed and entry cost functions from the dynamic model. The quarterly discount factor,  $\beta$ , is fixed at 0.99 (that implies an annual discount factor of 0.96). All the estimated fixed and entry cost parameters are measured in ten thousands of annual 1999 dollars. Due to previously discussed properties of our data sample, the reader is reminded that the magnitudes of our computed variable profits that feed into the dynamic model are at most roughly 10% of actual magnitudes, which in turn implies that the magnitudes of our fixed and entry cost estimates are at most 10% of actual magnitudes.

Point estimates of parameters in the fixed cost functions are unreasonably small

and imprecisely estimated. As such, we cannot draw reliable inferences about the size of fixed cost. Fortunately, based on the objectives of our study we are most interested in parameter estimates in the entry cost function, which is where we now focus the remainder of the discussion.

**Table 9**  
**Estimates of Parameters in Fixed and Entry Cost Functions**

Variables	Parameter Estimates ( $\theta$ ) (In ten thousand \$)	Standard Errors <sup>a</sup>
<b>Fixed cost (quarterly):</b>		
Mean fixed cost	5.00E-10	0.0145
Size of Presence at market endpoints	7.06E-11	1.01E-04
<b>Entry costs:</b>		
Mean entry cost for Legacy carriers	3.0755***	0.0153
Mean entry cost for Southwest	3.3499***	0.0164
Mean entry cost for Other LCCs	2.7467***	0.0135
Size of Presence at market endpoints	-0.0072***	9.85E-05
Traditional Codesharing	-0.0197***	8.67E-04
Virtual Codesharing	-0.0042***	5.42E-04
Traditional Codesharing $\times$ Southwest	0.0295***	0.0011
Virtual Codesharing $\times$ Southwest	0.0069***	0.0016
Traditional Codesharing $\times$ Other LCCs	0.0090***	7.18E-04
Virtual Codesharing $\times$ Other LCCs	-0.0058***	6.53E-04

\*\*\* indicates statistical significance at 1%

<sup>a</sup> Standard errors are computed via bootstrapping. The bootstrapping procedure is described in Appendix D.

It is difficult to obtain data separate from those used in this study to validate market entry cost estimates generated by the model. One reason for this difficulty is that the cost an airline faces to enter a market may in part be in terms of opportunity cost, i.e., the revenue forgone by not being able to use aircrafts in an alternate market. With this caveat in mind, we now discuss parameter estimates in the entry cost function.

Since the median variable profit is \$17,270.3, then for a given directional origin-destination market an airline generates, on average, less than \$17,270.3 profit each quarter.

Estimates from Table 9 show that the average estimated cost to enter an origin-destination market is approximately \$30,574, which is at least 1.77 times ( $=\$30,574/\$17,270.30$ ) as large as quarterly profit. It is notable from the estimates that mean entry cost differs by the carrier categories considered. Southwest has the highest mean market entry cost followed by legacy carriers and other low-cost-carriers, \$33,499, \$30,755 and \$27,467 respectively. Furthermore, the pairwise difference between any two of these three mean market entry costs is statistically significant at conventional levels of statistical significance. The decision of market entry is forward-looking, and our estimates suggest that it will take an airline slightly more than one quarter of profits to recoup the one-time sunk entry cost investment.

The estimated entry cost coefficient on “Size of Presence at market endpoints” is negative and statistically significant at conventional levels of statistical significance, suggesting that an airline’s market entry cost decreases with the size of the airline’s presence at the endpoint cities of the market. In other words, larger endpoint city presence makes it easier for the airline to actually start servicing the route. This result is consistent with how the literature believes airline markets work [see Berry (1992); Goolsbee and Syverson (2008); Gayle and Wu (2013) among others].

The coefficient estimates on traditional and virtual codesharing variables are negative and statistically significant. Based on our previous discussion of the interpretation of parameters in the entry cost function (equation (14)), the coefficient estimates on these two codeshare variables essentially capture the influence of codesharing on the market entry cost of potential entrants that are legacy carriers. Therefore, these coefficient estimates suggest that an increase in the extent of codesharing by incumbent carriers in a market reduces the market entry cost of potential entrants that are legacy carriers.

Recall that our descriptive statistics in Table 3 show that: (1) the vast majority of codeshare products are provided by legacy carriers; and (2) almost all of each legacy carrier’s codeshare products are codeshared with other legacy carriers. Therefore, the econometric evidence in Table 9 suggesting that more codesharing in a market makes it less costly for potential entrant legacy carriers to enter the market may in part be driven by the Chen and Ross (2000) argument, which is that incumbents may offer to share their facility (in our context, predominantly airplane seats owned by legacy carriers) with some

potential entrants (apparently other legacy carriers) in order to discourage the potential entrant from entering on a larger, and more competitive, scale by exclusively using its own plane on the full route. In other words, entry may be encouraged, as reflected by the lower entry cost, in a way that limits the scale of entry.

A key result is that the coefficient estimates on the two interaction variables between codesharing and the Southwest dummy variable are both positive and statistically significant. In addition, the total effect (rather than the relative effect) of each type of codesharing on Southwest's market entry cost is positive. As previously discussed, the total effect of each type of codesharing is captured by the sum of the coefficients on the relevant codeshare variable and its interaction with the Southwest dummy variable. As such, the total effect of traditional codesharing on Southwest's market entry cost is 0.0098 ( $= 0.0295 - 0.0197$ ), while the total effect of virtual codesharing on Southwest's market entry cost is 0.0027 ( $= 0.0069 - 0.0042$ ). So, even though each type of incumbent codesharing has a positive effect on Southwest's market entry cost, the parameter estimates provide evidence that traditional codesharing has a larger impact on Southwest's market entry cost compared to virtual codesharing. In contrast, the total effect of each type of codesharing on other LCC's market entry cost is negative. The total effect of traditional codesharing on other LCC's market entry cost is -0.0107 ( $= 0.0090 - 0.0197$ ), while the total effect of virtual codesharing on other LCC's market entry cost is -0.01 ( $= -0.0058 - 0.0042$ ). In summary, the coefficient estimates suggest that incumbent codesharing raises the market entry cost for Southwest, but reduces market entry cost for all other carriers (legacy and other LCC). In other words, market incumbent codesharing puts Southwest at a relative disadvantage to enter the market compared to potential entrants.

A useful feature of the structural econometric model is that the model allows us to monetize the extent to which codesharing by market incumbent carriers influences market entry barriers faced by potential entrants. Parameter estimates in the entry cost function suggest that each percentage point increase in traditional codeshare products offered by incumbents in a market raises market entry cost for Southwest by 0.3% ( $= \frac{\$295 - \$197}{\$33,499} \times 100$ ), while each percentage point increase in virtual codeshare products raises market entry cost for Southwest by 0.08% ( $= \frac{\$69 - \$42}{\$33,499} \times 100$ ). In contrast, each percentage point

increase in traditional codeshare products offered by incumbents in a market reduces market entry cost by 0.64% ( $= \frac{\$197}{\$30,755} \times 100$ ) for potential entrant legacy carriers, and by 0.39% ( $= \frac{\$197-\$90}{\$27,467} \times 100$ ) for potential entrants that are other LCCs. Similarly each percentage point increase in virtual codeshare products in a market reduces market entry cost by 0.14% ( $= \frac{\$42}{\$30,755} \times 100$ ) for potential entrant legacy carriers, and by 0.36% ( $= \frac{\$42+\$58}{\$27,467} \times 100$ ) for potential entrants that are other LCCs.

We argue above that a possible reason why potential entrant legacy carriers find it less costly to enter markets with more codesharing is due to the fact that the incumbents that codeshare are typically legacy carriers, and legacy carriers typically codeshare with other legacy carriers. However, what is the rationale for the econometric results suggesting that codesharing between legacy carriers make it more difficult for Southwest, but less difficult for other LCCs to enter a market? Codesharing alliances between legacy carriers prompt consumer loyalty through the carriers' reciprocal frequent-flyer programs and more available product options, which requires Southwest to exert more effort to secure its consumer base. For other LCCs, perhaps a large set of consumers served by them does not have significant overlap with the set of consumers served by legacy carriers, and therefore the two carrier types only weakly compete with each other. However, Southwest has a unique crossover strategic position in which it effectively competes with both legacy and other LCC carriers' for their respective consumer bases. As such, a relative entry-deterrent effect for Southwest translates into a relative ease of market entry for both legacy and other LCC carriers.

### ***Counterfactual Experiment***

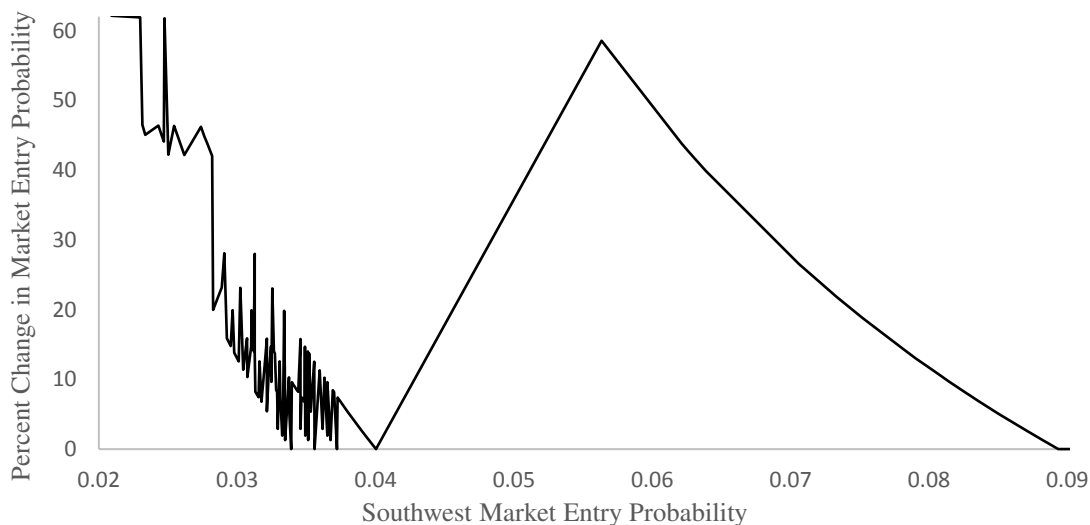
Another useful feature of the structural econometric model is that we can use it to perform counterfactual experiments. Details on the general technical procedure of how we use the dynamic entry model to perform counterfactual experiment are laid out in Aguirregabiria and Ho (2012, pp. 170 -171). However, the simple intuitive idea of the counterfactual experiment is the following. We first generate market entry probabilities from the model using the factual set of parameter estimates reported in Table 9. These

market entry probabilities can be referred to as the factual market entry probabilities. Next, we generate market entry probabilities from the model using a counterfactual set of parameters, which can be referred to as the counterfactual market entry probabilities. Since the objective is to analyze the impact of market incumbent carrier's codesharing on the probability that Southwest enters the market, the distinction between the factual and counterfactual sets of parameters is that the coefficient estimates on all codeshare variables are set equal to 0 in the counterfactual set of parameters. When the entry cost function parameter estimates associated with codesharing are counterfactually set to zero in the model, this effectively lowers (increases) the market entry cost for Southwest (other carriers). A comparison between Southwest's factual market entry probabilities with its counterfactual market entry probabilities reveals the impact that codesharing has on the probability that Southwest enters various markets.

The factual market entry probabilities for Southwest range from 0.02 to 0.089 for markets in which codesharing impacts Southwest's probability of entering. Our model predicts that Southwest's market entry probabilities increase by a mean 15.81% when the parameters that capture the entry deterrence impact of codesharing are counterfactually set to zero in the model. If we focus on the subset of markets in which Southwest's market entry probabilities range from 0.02 to 0.05, then its market entry probabilities are predicted to increase by a mean 17.15% when the parameters that capture the entry deterrence impact of codesharing are counterfactually set to zero in the model. So the markets in which Southwest have lower entry probabilities tend to have larger predicted percent increases in its entry probabilities. These mean percent increases in Southwest's market entry probabilities are statistically significant at the 1% level.

Figure 1 illustrates the distribution of the market entry-deterrent impact of codesharing over the sample range of market entry probabilities for Southwest. Specifically, the figure plots the relationship between Southwest's market entry probability and the predicted percent change in Southwest's market entry probability due to counterfactual removal of the entry-deterrent impact of incumbent codesharing. First, it is evident from Figure 1 that counterfactual removal of the entry-deterrent impact of incumbent codesharing increases the probability that Southwest enters a market over the sample range of Southwest's market entry probabilities. However, the extent of the

percentage increase in Southwest's market entry probability is not uniform across its market entry probabilities. It is evident that the predicted percent increases in entry probabilities have a bimodal feature over our sample range of Southwest's market entry probabilities. In particular, the entry-deterrent impact of incumbent codesharing on Southwest's market entry probabilities is distinctively larger at two different levels of its market entry probabilities; specifically, counterfactual removal of the entry-deterrent impact of incumbent codesharing yields an approximate 60% predicted increase at market entry probability levels 0.02 and 0.056 respectively.



**Figure 1:** Relationship between Southwest's Market Entry Probability and Predicted Percent Change in its Market Entry Probability due to Counterfactual Removal of Entry-Deterrent Impact of Codesharing by Market Incumbents

In summary, the empirical analysis suggests that codesharing between market incumbent airlines results in a relative increase in Southwest's market entry cost and a decrease in the probability that Southwest will enter the relevant market, which can be interpreted as an entry deterring effect to Southwest. Importantly, we also find that the entry deterrent impact of incumbent codesharing is not linear with respect to the likelihood/probability of Southwest's market entry.

## 5. Concluding Remarks

The main objective of our paper is to use a structural econometric model to investigate: (1) whether codesharing between airlines in domestic air travel markets, a form of strategic alliance, has a deterrent effect on the entry of potential competitors; (2) whether there is a particular type of codesharing among alliance partners that is most effective at deterring entry; and (3) whether the market entry deterrence impact of codesharing varies by the identity of potential market entrants. Advantages of using a structural econometric model are that: (1) we are able to quantify, in monetary terms, possible market entry barriers associated with codesharing; and (2) we are able to predict the extent to which a potential entrant's market entry probabilities are affected by market incumbent carrier's codesharing.

We find that more codesharing, both traditional and virtual, between incumbent carriers in a market puts Southwest at a relative disadvantage to enter the market compared to all other potential entrants (legacy carriers and other low-cost carriers). Specifically, each percentage point increase in traditional codeshare products offered by incumbents in a market raises market entry cost for Southwest by 0.3%, but reduces market entry cost by 0.64% and 0.39% for legacy and "other" low-cost carriers respectively. However, each percentage point increase in virtual codeshare products offered by incumbents in a market raises market entry cost for Southwest by 0.08%, but reduces market entry cost by 0.14% and 0.36% for legacy and "other" low-cost carriers respectively. In addition, the model predicts that Southwest's market entry probabilities increase by a mean 15.81% when the parameters that capture the entry deterrence impact of codesharing are counterfactually set to zero in the model. Therefore, codesharing by market incumbent carriers has a relative market entry deterrent effect on Southwest. Furthermore, the parameter estimates provide evidence that traditional codesharing has a larger impact on Southwest's market entry cost compared to virtual codesharing.

We argue that the entry deterrent effect is binding for Southwest but not for others due to the evidence that the vast majority of codesharing is done between legacy carriers, and competition between Southwest and legacy carriers is stronger than competition between other low-cost carriers and legacy carriers. Consistent with this argument, previous work provides evidence that incumbent legacy carriers do not cut fares in response



to potential competition from other low-cost carriers, but cut fares in response to potential competition from Southwest [Brueckner, Lee and Singer (2012)].

We also find that an airline's market entry cost decreases with the size of the airline's presence at the endpoint cities of the market. This finding is consistent with findings in Aguirregabiria and Ho (2012), and may be due to economies of scale and scope by concentrating most operations in a hub airport.

The focus of our study is on U.S. domestic air travel markets, however future work may investigate whether results similar to ours exist for codesharing in international air travel markets.

## Appendix A: Transition Rules for State Variables

The state variables we observe are:  $\{S_{imt}, R_{imt}^*, Pres_{imt}, Percent\_Trad_{mt}, Percent\_Virtual_{mt}\}$ . We utilize Vector autoregressions (VAR) to model transition rules for the state variables  $R_{imt}^*$ ,  $Pres_{imt}$ ,  $Percent\_Trad_{mt}$ , and  $Percent\_Virtual_{mt}$ .

Let  $Z_{mt}^{var} = \{1, R_{1mt}^*, \dots, R_{9mt}^*, Pres_{1mt}, \dots, Pres_{9mt}, Percent\_Trad_{mt}, Percent\_Virtual_{mt}\}$ .

Transition rules for the state variables are as follows:

$$S_{im,t+1} = a_{it} \quad (A1)$$

$$R_{im,t+1}^* = a_{imt}(\alpha^R Z_{mt}^{var} + \zeta_{imt}^R) \quad (A2)$$

$$Pres_{im,t+1} = \alpha^{Pres} Z_{mt}^{var} + \zeta_{imt}^{Pres} \quad (A3)$$

$$Percent\_Trad_{m,t+1} = \alpha^{Percent\_Trad} Z_{mt}^{var} + \zeta_{mt}^{Percent\_Trad} \quad (A4)$$

$$Percent\_Virtual_{m,t+1} = \alpha^{Percent\_Virtual} Z_{mt}^{var} + \zeta_{mt}^{Percent\_Virtual} \quad (A5)$$

where  $\zeta_{imt}^R$ ,  $\zeta_{imt}^{Pres}$ ,  $\zeta_{mt}^{Percent\_Trad}$ , and  $\zeta_{mt}^{Percent\_Virtual}$  are assumed to be normally distributed.

The joint transition probabilities of the state variables are determined by:

$$F_i^\sigma(y_{t+1}|a_{it}, y_t) = \begin{cases} 1\{S_{i,t+1} = 1\} * Pr_R * Pr_{Pres} * Pr_{Percent\_Trad} * Pr_{Percent\_Virtual} * Pr_{comp} \\ 1\{S_{i,t+1} = 0\} * Pr_R' * Pr_{pres} * Pr_{Percent\_Trad} * Pr_{Percent\_Virtual} * Pr_{comp} \end{cases} \quad (A6)$$

where

$$\Pr_R = F_R(R_{it+1}|R_{it}) * \prod_{j \neq i} F_R(R_{jt+1}|R_{jt}) \quad (\text{A7})$$

$$\Pr_{Pres} = F_{Pres}(Pres_{it+1}|Pres_{it}) * \prod_{j \neq i} F_{Pres}(Pres_{jt+1}|Pres_{jt}) \quad (\text{A8})$$

$$\Pr_{Percent\_Trad} = F_{Percent\_Trad}(Percent\_Trad_{t+1}|Percent\_Trad_t) \quad (\text{A9})$$

$$\Pr_{Percent\_Virtual} = F_{Percent\_Virtual}(Percent\_Virtual_{t+1}|Percent\_Virtual_t) \quad (\text{A10})$$

$$\Pr_R' = 1\{R_{i,t+1} = 0\} * \prod_{j \neq i} F_R(R_{jt+1}|R_{jt}) \quad (\text{A11})$$

$$\Pr_{comp} = \prod_{j \neq i} Pr(s_{jt+1} = \sigma_j(y_{jt}, \varepsilon_{jt})|y_{jt}) \quad (\text{A12})$$

## Appendix B: Representation of Markov Perfect Equilibrium (MPE) using Conditional Choice Probabilities (CCPs)

Recall that expected one-period profit function,  $\Pi_{imt}(a_{it}, y_t)$ , is specified as:

$$\Pi_{imt}(a_{it}, y_t) = R_{imt}^* - a_{imt}(FC_i + (1 - s_{imt})EC_i), \quad (\text{B1})$$

where parametric specifications for  $FC_i$  and  $EC_i$  were previously given in equations (13) and (14). Based on equation (B1):

$$\Pi_{imt}(0, y_t) = R_{imt}^* \quad (\text{B2})$$

and

$$\Pi_{imt}(1, y_t) = R_{imt}^* - FC_i - (1 - s_{imt})EC_i \quad (\text{B3})$$

Let

$$z_{imt}(0, y_t) = \{R_{imt}^*, 0, 0, 0, 0, 0, 0, 0, 0, 0\} \quad (\text{B4})$$

and

$$\begin{aligned} & z_{imt}(1, y_t) \\ &= \left\{ R_{imt}^*, \quad -1, \quad -Pres_{imt}, \right. \\ & \quad -1, \quad -(1 - s_{it})Pres_{imt}, \\ & \quad -(1 - s_{it})Percent\_Trad_{mt}, \\ & \quad -(1 - s_{it})Percent\_Virtual_{mt}, \\ & \quad -(1 - s_{it})Percent\_Trad_{mt} \times Southwest, \\ & \quad -(1 - s_{it})Percent\_Virtual_{mt} \times Southwest, \\ & \quad -(1 - s_{it})Percent\_Trad_{mt} \times Other\_lcc, \\ & \quad \left. -(1 - s_{it})Percent\_Virtual_{mt} \times Other\_lcc \right\} \end{aligned} \quad (\text{B5})$$

and

$$\theta = \{1, \theta_0^{FC}, \theta_1^{FC}, \theta_0^{EC}, \theta_1^{EC}, \theta_2^{EC}, \theta_3^{EC}, \theta_4^{EC}, \theta_5^{EC}, \theta_6^{EC}, \theta_7^{EC}\} \quad (B6)$$

Therefore, we can re-write:

$$\Pi_{imt}(0, y_t) = z_{imt}(0, y_t) \times \theta \quad (B7)$$

and

$$\Pi_{imt}(1, y_t) = z_{imt}(1, y_t) \times \theta \quad (B8)$$

As discussed in Aguirregabiria and Ho (2012), the MPE can be represented as a vector of conditional choice probabilities (CCPs),  $\mathbf{P}$ .  $\mathbf{P} = \{P_i(y): \text{for every firm and state } (i, y)\}$  that solves fixed point problem  $\mathbf{P} = \psi(\mathbf{P}, \theta)$  is a vector of best response mapping:

$$\left\{ \psi \left( \tilde{Z}_i^P(y) \frac{\theta}{\sigma_\varepsilon} + \tilde{e}_i^P(y) \right) : \text{for every firm and state } (i, y) \right\} \quad (B9)$$

where in our study  $\psi(\cdot)$  is the CDF of the type 1 extreme value distribution, and

$$\tilde{Z}_i^P(y) = Z_i(1, y) - Z_i(0, y) + \beta [F_{i,y}^P(1) - F_{i,y}^P(0)] \times w_{z,i}^P, \quad (B10)$$

$$\tilde{e}_i^P(y) = \beta [F_{i,y}^P(1) - F_{i,y}^P(0)] \times w_{e,i}^P, \quad (B11)$$

$$w_{z,i}^P = (1 - \beta * \overline{F_{i,y}^P})^{-1} \times \{P_i(y) * Z_i(1, y) + [1 - P_i(y)] * Z_i(0, y)\}, \quad (B12)$$

$$w_{e,i}^P = (1 - \beta * \overline{F_{i,y}^P})^{-1} \times [P_i(y) * e_i^P] \quad (B13)$$

and

$$\overline{F_{i,y}^P} = [(P_i(y) \times \mathbf{1}'_M) * F_{i,y}^P(1) + ((1 - P_i(y)) \times \mathbf{1}'_M) * F_{i,y}^P(0)]. \quad (B14)$$

where  $F_{i,y}^P(0)$  and  $F_{i,y}^P(1)$  are state transition probability matrices for  $a_{it} = 0$  and  $a_{it} = 1$  respectively;  $w_{z,i}^P$  and  $w_{e,i}^P$  are vectors of valuations that depend on CCPs and transition probabilities, but not on the dynamic parameters being estimated. Since  $\varepsilon_{imt}$  is assumed type 1 extreme value distributed,  $e_i^P$  is a function vector equal to  $e_i^P = \gamma - \ln(P_i(y))$  where  $\gamma = 0.5772$  is Euler's constant.

## Appendix C: Implementing the Nested Pseudo Likelihood (NPL) Estimator

Nadaraya-Watson kernel nonparametric regression is used to get nonparametric estimate. Given the PML estimator,  $\hat{\theta}_{PML}$ , and the initial nonparametric estimate of CCPs,  $\widehat{P}_0$ , we construct a new estimator of CCPs,  $\widehat{P}_1$ , using the best response CCPs equation  $\widehat{P}_1 = \psi(y, \widehat{P}_0, \hat{\theta}_{PML})$ . Then we redo the maximization of the pseudo likelihood function to obtain a new estimate of  $\theta$  using  $\widehat{P}_1$ , instead of  $\widehat{P}_0$ , in the pseudo log likelihood function, that is, we solve  $\hat{\theta}_2 = \arg \max_{\theta} Q(\theta, \widehat{P}_1)$ . The process is repeated  $K$  times, and the  $K^{th}$  estimates of  $\theta$  and  $P$  are obtained by  $\hat{\theta}_K = \arg \max_{\theta} Q(\theta, \widehat{P}_{K-1})$  and  $\widehat{P}_K = \psi(y, \widehat{P}_{K-1}, \hat{\theta}_K)$  respectively. The algorithm is terminated on the  $K^{th}$  iteration only if the CCP vector  $\widehat{P}_K$  is “close” to  $\widehat{P}_{K-1}$  based on a stipulated tolerance level. Based on this algorithm, an NPL fixed point is defined as a pair  $(\hat{\theta}_K, \widehat{P}_{K-1})$ . In our estimation, the algorithm is terminated when  $K=5$ . Aguirregabiria and Mira (2002, 2007) argue that this NPL estimation algorithm can reduce significantly the finite sample bias of the two-step PML estimator.

## Appendix D: Bootstrapping Standard Errors for Parameter Estimates in the Dynamic Entry Model

First, we assume that parameter estimates from the first-stage regressions are normally distributed with means equal to the point estimates of the parameters, and their variances and covariances equal to the estimated variance-covariance matrix for these parameters. This assumption allows us to generate normal random draws of the parameter estimates from the first-stage regressions. The first-stage regressions are: (1) the demand model; and (2) the traditional codeshare, and virtual codeshare linear regressions (these are equations (15) and (16) in the paper). Specifically, we generate 35 random draws from a multivariate normal distribution. The multivariate normal distribution we use has means equal to the existing first-stage parameter estimates, and variance/covariance equal to the variance/covariance matrix of the first-stage parameter estimates. A single draw from the multivariate normal distribution produces a set of first-stage parameter estimates. As such, we effectively generate 35 sets of first-stage parameter estimates from this random draw process.

For each draw of the first-stage parameter estimates, we re-estimate the dynamic entry model to obtain a set of dynamic parameter estimates associated with each of the 35 sets of first-stage parameter estimates, respectively. In other words, the dynamic entry model is re-estimated 35 times, where each estimation uses a different set of first-stage parameter estimates. This procedure is extremely computationally intensive since a single estimation of the dynamic entry model can take several weeks of continuous computer running to achieve convergence of the estimation algorithm.

Once we have 35 different sets of dynamic parameter estimates, we then use simple descriptive statistics to compute the standard errors across the 35 data points for each structural parameter estimate. This produces a bootstrap standard error for each dynamic parameter estimate.

## References

- Aguirregabiria, Victor and Chun-Yu Ho. 2012. "A Dynamic Oligopoly Game of the US Airline Industry: Estimation and Policy Experiments," *Journal of Econometrics*, Vol. 168, pp. 156-173.
- Aguirregabiria, Victor and Chun-Yu Ho. 2010. "A Dynamic game of Airline Network Competition: Hub-and-Spoke Networks and Entry Deterrence," *International Journal of Industrial Organization*, Vol. 28, pp. 377-382.
- Aguirregabiria, Victor and Pedro Mira. 2002. "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70, pp. 1519-1543.
- Aguirregabiria, Victor and Pedro Mira. 2007. "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75, pp. 1-53.
- Bamberger, G., D. Carlton and L. Neumann. 2004. "An Empirical Investigation of the Competitive Effects of Domestic Airline Alliances," *Journal of Law and Economics*, Vol. XLVII, pp. 195-222.
- Berechman, J., Poddar, S. and O. Shy. 1998. "Network Structure and Entry in the Deregulated Airline Industry," *Keio Economic Studies*, 35, pp. 77-82.
- Berry, Steven. 1990. "Airport Presence as Product Differentiation," *American Economic Review*, Vol. 80: 394-399.
- Berry, Steven. 1992. "Estimation of a Model of Entry in the Airline Industry", *Econometrica*, 60 (4): 889-918.
- Berry, Steven. 1994. "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242-262.
- Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium," *Econometrica*, Vol. 63, 841-990.
- Berry, S., and P. Jia. 2010. "Tracing the Woes: An Empirical Analysis of the Airline Industry," *American Economic Journal: Microeconomics*, Vol. 2(3): 1-43.
- Brander, James A., and Anming Zhang. 1990. "Market Conduct in the Airline Industry: An Empirical Investigation," *RAND Journal of Economics*, Vol. 21, 567-583.
- Brueckner, J. 2001. "The Economics of International Codesharing: An Analysis of Airline Alliances," *International Journal of Industrial Organization*, Vol. 19, 1475-1498.
- Brueckner, J. 2003. "International Airfares in the Age of Alliances," *Review of Economics and Statistics*, Vol. 85, 105-118.
- Brueckner, Jan K., Darin Lee and Ethan Singer. 2012. "Airline Competition and Domestic U.S. Airfares: A Comprehensive Reappraisal," in *Economics of Transportation*.
- Brueckner, Jan K., Darin Lee and Ethan Singer. 2011. "Alliances, Codesharing, Antitrust Immunity and International Airfares: Do Previous Patterns Persist?" *Journal of Competition Law & Economics*, 7(3), 573-602.
- Brueckner, Jan K and W. Tom Whalen. 2000. The Price Effects of International Airline Alliances. *Journal of Law and Economics* 43:503-45.
- Chicu, Mark 2013. "Dynamic Investment and Deterrence in the U.S. Cement Industry," Manuscript, Available at SSRN: <http://ssrn.com/abstract=2589580> or <http://dx.doi.org/10.2139/ssrn.2589580>.
- Chen, Yongmin, and Philip G. Gayle 2007. "Vertical Contracting Between Airlines: An Equilibrium Analysis of Codeshare Alliances," *International Journal of Industrial Organization*, Vol. 25, Issue 5, pp. 1046-1060.

- Chen, Z. and Ross, T. 2000. "Strategic Alliances, Shared Facilities, and Entry Deterrence," *RAND Journal of Economics*, Vol. 31(2), 326-344.
- Daraban, Bogdan and Gary M. Fournier. 2008. "Incumbent Responses to Low-Cost Airline Entry and Exit: A Spatial Autoregressive Panel Data Analysis", *Research in Transportation Economics*, Vol 24, Issue 1, pp. 15-24.
- Gayle, Philip G. 2013. "On the Efficiency of Codeshare Contracts Between Airlines: Is Double Marginalization Eliminated?" *American Economic Journal: Microeconomics*, Vol. 5, Issue 4, pp. 244-273.
- Gayle, Philip G. 2008. "An Empirical Analysis of the Competitive Effects of the Delta/Continental/Northwest Codeshare Alliance," *Journal of Law and Economics*, Vol.51, pp. 743-766.
- Gayle, Philip G. 2007. "Is Virtual Codesharing A Market Segmenting Mechanism Employed by Airlines?" *Economics Letters*, Vol. 95, No. 1, pp. 17-24.
- Gayle, Philip G. and Dave Brown 2014. "Airline Strategic Alliances in Overlapping Markets: Should Policymakers be Concerned?" *Economics of Transportation*, Volume 3, Issue 4, December 2014, Pages 243–256.
- Gayle, Philip G. and Chi-Yin Wu 2013. "A Re-examination of Incumbents' Response to the Threat of Entry: Evidence from the Airline Industry" *Economics of Transportation*, 2 (2013), pp. 119-130.
- Gayle, Philip G. and Chi-Yin Wu 2015. "On the Extent to which the Presence of Intermediate-stop(s) Air Travel Products Influences the Pricing of Nonstop Air Travel Products" forthcoming in *Review of Network Economics*, Volume 13, Issue 3, Pages 355–395, June 2015.
- Goetz, Christopher F., and Adam H. Shapiro. 2012. "Strategic Alliance as a Response to the Threat of Entry: Evidence from Airline Codesharing", *International Journal of Industrial Organization*, Elsevier, vol. 30(6), pages 735-747.
- Goolsbee, Austan and Chad Syverson. 2008. "How Do Incumbents Respond to The Threat of Entry? Evidence from The Major Airlines," *The Quarterly Journal of Economics*, Vol. 123, No. 4: 1611–1633.
- Hendricks, K., Piccione, M. and G. Tan. 1997. "Entry and Exit in hub-and-spoke Networks," *RAND Journal of Economics*, Vol. 28, Issue 2, pp. 291-303.
- Ito, H., and D. Lee. 2007. "Domestic Codesharing, Alliances and Airfares in the U.S. Airline Industry," *Journal of Law and Economics*, Vol. 50, pp. 355-380.
- Lin, Ming Hsin. 2005. "Alliances and Entry in a Simple Airline Network," *Economics Bulletin*, Vol.12, No. 2, pp. 1-11.
- Lin, Ming Hsin. 2008. "Airline Alliances and Entry Deterrence," *Transportation Research Part E*, Vol. 44, pp. 637-652.
- Mantovani, Andrea and Ornella Tarola. 2007. "Did the Entry of Low-cost Companies Foster the Growth of Strategic Alliances in the Airline Industry?" *Rivista di Politica Economica*, Vol. 97, Issue 1-2, pp. 189-220.
- Mason, Keith J. and William G. Morrison. 2008. "Towards A Means of Consistently Comparing Airline Business Models with an Application to the 'Low Cost' Airline Sector," *Research in Transportation Economics* 24, 75-84.
- Morrison, S. A. 2001. "Actual, Adjacent, and Potential Competition: Estimating the Full Effect of Southwest Airlines," *Journal of Transport Economics and Policy* 35 (2), 239-256.

- Nevo, Aviv. 2000. "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9 (4): 513-548.
- Oum, Tae, David W. Gillen, and S. E. Noble. 1986. "Demand for Fareclass and Pricing in Airline Markets," *Logistics and Transportation Review*, Vol. 22, 195-222.
- Oum, T.H. and J.H. Park 1997. "Airline Alliances: Current Status, Policy Issues, and Future Directions," *Journal of Air Transport Management*, Vol. 3, pp. 133-144.
- Oum, T.H., Zhang, A. and Y. Zhang 1995. "Airline network rivalry," *Canadian Journal of Economics*, Vol. 28, Issue 4a, pp. 836-857.
- Park, J.H. 1997. "Strategic Alliance Modeling and Empirical Analysis," Ph.D. thesis, Faculty of Commerce and Business Administration, *University of British Columbia*.
- Park, J.H. and Zhang 1998. "Airline Alliances and Partner Firms' Outputs," *Transportation Research*, Vol. 34, pp. 245-255.
- Park, J.H. and Zhang 2000. "An Empirical Analysis of Global Airline Alliances: Cases in North American Markets," *Review of Industrial Organization*, Vol. 16, pp. 367-384.
- Peters, C. 2006. "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry," *Journal of Law and Economics*, Vol. XLIX, 627-649.
- Snider, C. 2009. "Predatory Incentives and Predation Policy: The American Airlines Case," Working Paper, Department of Economics, UCLA.
- Sweeting, A. 2013. "A Model of Dynamic Limit Pricing with an Application to the Airline Industry," Working Paper, Department of Economics, Duke University.
- Villas-Boas, Sofia Berto. 2007. "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data," *Review of Economic Studies*, 74 (2): 625-652.
- Williams, J. 2012. "Capacity Investment, Exclusionary Behavior, and Welfare: A Dynamic Model of Competition in the Airline Industry," Working Paper, University of Georgia.