Airline strategic alliances in overlapping Markets: Should policymakers be concerned?

Gayle, Philip and Brown, Dave

Kansas State University, Pennsylvania State University

January 2015

Airline Strategic Alliances in Overlapping Markets: Should Policymakers be Concerned?

Philip G. Gayle*
Kansas State University

Dave Brown†
Pennsylvania State University

This version: January, 2015
First version: June, 2008

Forthcoming in *Economics of Transportation*

Abstract

When there is significant overlap in potential partner airlines' route networks, policymakers have expressed concern that an alliance between such airlines may facilitate collusion on price and/or service levels in the partners' overlapping markets. The contribution of our paper is to put together a structural econometric model that is able to explicitly disentangle the demand and supply effects associated with an alliance between such airlines. The estimates from our structural econometric model do identify demand-increasing effects associated with the Delta/Continental/Northwest alliance, but statistically reject collusive behavior between the partners.

JEL Classification: L40, L13, L93

Keywords: Codeshare Alliance; Collusion; Airline Competition; Discrete Choice Demand Model; Nested Logit.

Acknowledgment: We thank editor Erik Verhoef and two anonymous referees for very helpful comments and suggestions. We also thank Ariel Pakes, Joseph Cullen, Andrew Ching, Greg Lewis, Carlene Belford, Volodymyr Bilotkach, Karunakaran Sudhir, and participants at the 7th International Industrial Organization Conference for their invaluable suggestions and comments. Any remaining errors are our own.

*Corresponding author: Department of Economics, 320 Waters Hall, Kansas State University, Manhattan, KS, 66506, USA; Voice: (785) 532-4581; Fax: (785) 532-6919; E-mail: gaylep@ksu.edu.

†Department of Economics, 607 Kern Graduate Building, Pennsylvania State University, University Park, PA 16802, USA, E-mail: drb39@psu.edu.

1 Introduction

Policymakers have expressed skepticism when reviewing airlines' application to form a codeshare alliance in the event that such an alliance involves potential partners that have significant overlap in their route networks. The heart of the concern is that these potential partners are direct competitors in the segments of their networks that overlap, and an alliance between them, which often requires broad discussions between partners to make their interline[1] service seamless, could facilitate collusion on prices and/or service levels in the partners' overlapping markets. Before ultimately approving the Delta/Continental/Northwest alliance, which was formed in June 2003, the U.S. Department of Transportation (DOT) expressed these concerns.[2] The DOT's review of this proposed alliance points out that the three airlines' service overlap in 3,214 markets accounting for approximately 58 million annual passengers, which is in contrast to the next largest alliance between United Airlines and US Airways with overlapping service in only 543 markets accounting for 15.1 million annual passengers. So unlike much of the literature that focuses on international airline alliances [Brueckner, Lee and Singer (2011); Brueckner and Proost (2010); Brueckner (2003); Brueckner and Whalen (2000); Bilotkach (2007); Lederman (2007) among others], we focus on a U.S. domestic alliance [Ito and Lee (2007); Bamberger, Carlton and Neumann (2004); Gayle (2008)].

Using a reduced-form econometric model similar to that in Bamberger, Carlton and Neumann (2004), Gayle (2008) has shed some light on price effects associated with the Delta/Continental/Northwest codeshare alliance. In particular, Gayle (2008) finds that the alliance is associated with a marginal price increase, which by itself points to possible collusive effects. But a marginal price increase is also consistent with increased demand and there is good reason to believe that an alliance has a demand-increasing effect associated with it. For example, passengers that are members of an airline's frequent-flyer program may cumulatively earn and redeem frequent-flyer miles across any partner in the alliance. The new opportunities for passengers to earn and redeem miles will likely increase demand for the alliance partners' products. In the case of enhancements to international frequent-flyer partnerships, Lederman (2007) provides reduced-form econometric evidence suggesting that enhancements to international frequent-flyer partnerships are associated with increases in domestic airline demand.

[1] Interline means that at some point in the trip when passengers change planes they also change airlines.

[2] See "Termination of review under 49U.S.C. § 41720 of Delta/Northwest/Continental Agreements," published by *Office of the Secretary, Department of Transportation*, January 2003.

To better understand the market effects associated with an alliance, both from the demand and supply sides of a market, it is important to go beyond the reduced-form analyses that currently exist in the literature. As such, the main contribution of our present paper is to specify and estimate a structural econometric model that allows us to disentangle demand changes from possible changes in airline pricing behavior that are associated with a codeshare alliance. The empirical separation of demand changes from airline pricing behavior changes allows us to: (1) statistically test whether a codeshare alliance is associated with a demand-increasing effect; and (2) statistically test whether a codeshare alliance is associated with collusive pricing behavior in the partners' overlapping markets, as feared by policymakers.

Our key findings are as follows: First, the econometric estimates for the air travel demand equation suggest that the Delta/Continental/Northwest codeshare alliance has a demand-increasing effect associated with it. Importantly, the demand-increasing effect is only evident in markets that the partners have a substantial joint passenger share (greater than 49%) prior to implementation of the alliance. Since a relatively larger proportion of passengers in a market are more likely to have frequent-flyer membership with at least one of the three carriers in markets that the carriers jointly dominate prior to the alliance, this finding is consistent with the argument that these frequent-flyer passengers will increase their demand for the alliance partners' products given that the alliance creates new opportunities for passengers to accumulate and redeem frequent-flyer points across partner carriers.

Second, a statistical non-nested test applied to air travel supply model selection suggests that Bertrand Nash pricing behavior, rather than collusive pricing behavior, between the three airlines better fit the data in markets where the three airlines codeshare together. To the best of our knowledge, this is the first paper to explicitly test and statistically reject that collusive pricing behavior is associated with a codeshare alliance.

The rest of the paper is organized as follows: In the next section we make some key definitions which build the foundation for important issues we subsequently model, analyze, and discuss. In section 3 we discuss characteristics of our data. We present the structural econometric model in section 4, while estimation strategy is discussed in section 5. Results are presented and discussed in section 6. Concluding remarks are offered in section 7.

2 Definitions

A *market* is defined as directional round-trip air travel between an origin and a destination airport during a particular period. The assumption that markets are directional implies that a round-trip air travel from Atlanta to Detroit is a distinct market than round-trip air travel from Detroit to Atlanta. Furthermore, this directional assumption allows for the possibility that origin city characteristics may influence market demand [see Gayle (2007a, 2007b, 2013), Berry, Carnall and Spiller (2006)].

A flight *itinerary* is defined as a specific sequence of airport stops in traveling from the origin to destination airport. An air travel *product* is defined as a unique combination of airline(s) and flight itinerary. Following Ito and Lee (2007), a *pure online product* means that the same airline markets and operates all segments of a round-trip. For example, three separate pure online products are: (1) a non-stop round-trip from Atlanta to Detroit marketed and operated by Delta Air Lines; (2) a round-trip from Atlanta to Detroit with one stop in Minneapolis marketed and operated by Delta Air Lines; and (3) a non-stop round-trip from Atlanta to Detroit marketed and operated by Northwest Air Lines. Note that all three products are in the same market - Atlanta to Detroit.

A *codeshare agreement* effectively allows one carrier (called the "*ticketing carrier*" or "*marketing carrier*") to sell seats on its partners' plane as if these seats are owned by the carrier selling the seats. The carrier whose plane that actually transports the passenger is referred to as the "*operating carrier*". For example, Northwest may sell tickets for a subset of seats on a Delta operated flight between Atlanta and Detroit as if the plane were owned by Northwest. Thus, a passenger that uses a codeshare itinerary may have bought the round-trip ticket from Northwest, but actually flies on a plane operated by Delta.

The literature on domestic airline alliances has identified two main types of codeshare itineraries: (1) *traditional* codeshare; and (2) *virtual* codeshare.[3] *Traditional* codeshare itineraries combine interline operating services of partner carriers on a given route, where one of these operating carriers is the sole ticketing carrier for the entire trip. An example of a traditional codeshare product is a trip from Atlanta to Detroit with one stop in Minneapolis, where the Atlanta to Minneapolis segment of the trip is operated by Delta, the Minneapolis to Detroit segment of the trip is operated by Northwest, but the ticket for the entire trip is marketed by Northwest. Brueckner and Whalen

[3] See Ito and Lee (2007) and Gayle (2008) for discussions of the main types of codeshare products in the U.S. domestic market.

(2000), Brueckner (2003), Ito and Lee (2007) and Gayle (2008) find evidence that traditional codesharing tends to lower rather than raise prices. An often cited reason for this price-decreasing effect of traditional codesharing is that this type of codesharing eliminates double markup that would otherwise persist when carriers are unaffiliated.[4]

Owing to the existing robust empirical evidence of a price-decreasing effect associated with traditional codesharing, this type of codesharing is not the focus of our present analysis. The type of codesharing we focus on in this research is referred to as *virtual* codeshare. A passenger using a virtual codeshare itinerary remains on a single operating carrier's plane(s) for the entire round-trip, but the ticket for the trip was marketed and sold by a partner ticketing carrier. Thus a key distinction between virtual codeshare and traditional codeshare is that traditional codeshare requires the passenger to travel on different operating carriers' planes (interline air travel) on a multi-segment route, while virtual codeshare does not involve interline air travel even when the passenger changes planes on a multi-segment route. We focus on virtual codesharing because Gayle (2008) finds that this is the only type of codesharing that is associated with price increases.

Figure 1 gives an example where two airlines' route networks overlap and the airlines may virtual codeshare together in the origin-destination market. The figure shows that Northwest and Delta both operate non-stop flights in the Atlanta to Detroit market. If they virtual codeshare together in this market, then a subset of the passengers on the Delta plane would have bought their tickets from Northwest, while a subset of the passengers on the Northwest plane would have bought their tickets from Delta.

[4] See Gayle (2013) for an empirical investigation of situations in which double markup may persist for traditional codeshare products. Chen and Gayle (2007) provides an analogous theoretical analysis of this issue.

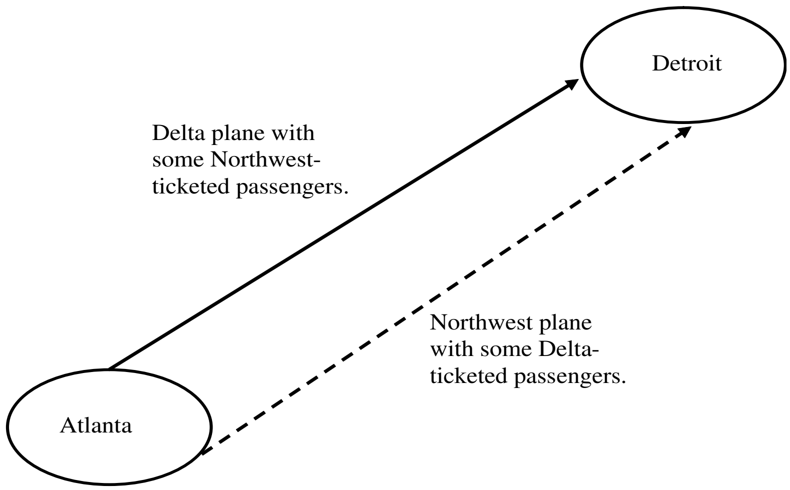

Figure 1: Route Network Diagram

Figure 2 shows an alternate situation in which the airlines' route networks may overlap. In Figure 2, Northwest operates a non-stop flight in the Atlanta to Detroit market, while Delta operates a one-stop itinerary in the Atlanta to Detroit market, but unlike Figure 1, Delta does not operate a non-stop flight in this market. Northwest and Delta's networks are still considered to be overlapping in Figure 2 even though Delta operates only a one-stop itinerary while Northwest operates a non-stop itinerary. Both carriers may virtual codeshare together in Figure 2.

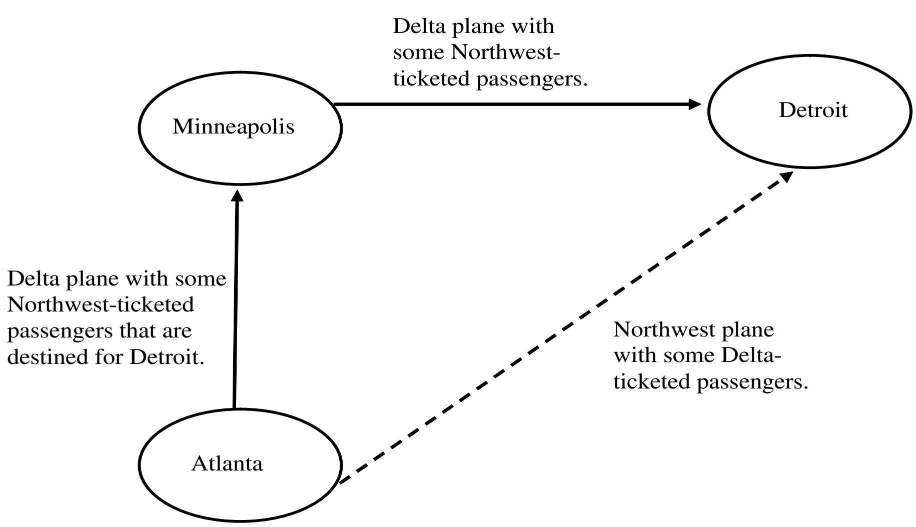

Figure 2: Modified Route Network Diagram

In Figure 2 it might seem counter-intuitive that a passenger would choose a one-stop itinerary even though a non-stop flight between the origin and destination is available. However, passengers often choose less convenient routes (flight itineraries that require intermediate stops) to get from their origin to destination when such alternate routing is competitively priced. In other words, within reasonable bounds, some passengers are willing to trade-off travel itinerary convenience for a lower price.

Figure 2 can also be used to illustrate a situation in which virtual codesharing is likely to have a demand-increasing effect associated with it. In the event that Northwest and Delta do not have a codeshare alliance, Northwest can only offer its Atlanta-based customers (some of whom may be members of Northwest's frequent-flyer program) a non-stop flight to Detroit. However, an alliance with Delta allows Northwest to offer its Atlanta-based customers both a non-stop flight on its own plane and a one-stop virtual codeshared itinerary operated solely by Delta. While passengers in Atlanta already had the option, prior to an alliance, to purchase either a pure online one-stop itinerary from Delta or a pure online non-stop flight from Northwest, Northwest's frequent-flyers could not accumulate frequent-flyer miles on the Delta operated flights. Thus, the alliance created a new opportunity for Northwest frequent-flyers to accumulate miles on a Delta operated one-stop itinerary. Similarly, Delta frequent-flyers that would like to travel on the non-stop Northwest flight also have a new opportunity to accumulate frequent-flyer miles on the Northwest operated flight. The new opportunity for passengers to accumulate frequent-flyer miles across partner carriers is one reason we expect a demand-increasing effect to be associated with a codeshare alliance. Our econometric model is designed to isolate and test for this potential demand-increasing effect.

Figure 2 is also useful to illustrate the main concern the DOT expressed in its review of the proposed alliance between Delta, Continental and Northwest. Since Delta and Northwest were competitors in the market shown in Figure 2, the DOT was concerned that forming an alliance would reduce the amount of competition between the two airlines. The econometric model we present below is designed to statistically test if collusive pricing behavior, rather than Bertrand Nash pricing behavior, between the three airlines better fit the data in markets that the three airlines virtual codeshare together during the post-alliance period.

3 Data

Data are drawn from the Origin and Destination Survey (DB1B), which is a 10% random sample of airline tickets from reporting carriers. DB1B is a database that is maintained and published by the U.S. Bureau of Transportation Statistics. Among other things, the database includes: (1) number of passengers that choose a given flight itinerary; (2) the fares of these itineraries; (3) the specific sequence of airport stops that each itinerary uses in getting passengers from the origin to destination city; (4) the carrier(s) that marketed and sold the travel ticket (ticketing carriers), and the carrier(s) that passengers actually fly on for their trip (operating carriers); and (5) the distance flown on each itinerary in a directional market. The distance associated with each itinerary in a market may differ since each itinerary may use different connecting airports in transporting passengers from the origin to destination city.

Unfortunately, the DB1B database does not include passenger-specific information. For example, relevant passenger-specific information that we do not have are: (1) whether or not a passenger has frequent-flyer membership with an airline; (2) the specific day of week of the travel; (3) the length of time in advance of travel that the passenger purchased the ticket; and (4) purpose of trip - leisure versus business. Therefore, we will have to rely on the econometric model's ability to tease out consumer choice behavior patterns from aggregated ticket purchase data. In addition, the database does not contain certain useful measures of travel itinerary convenience such as layover times or departure times. Notwithstanding these deficiencies in the data, we are able to construct useful measures of itinerary convenience from the available information in the data, which we discuss below.

The data we use link each product to a directional market rather than a mere non-stop route or segment of a market. For this research, we focus on U.S. domestic flights offered and operated by U.S. carriers in the fourth quarters of 2002 (pre-alliance) and 2003 (post-alliance).[5]

We arrive at the final sample used for estimation by applying a few filters to the original data set. First, itineraries with price less than $100 are excluded due to the high probability that these may be coding errors or passengers redeeming frequent-flyer miles to obtain a discounted fare. Second, itineraries with an inordinate number of intermediate stops (more than two) were dropped. Third, we focus on pure online and virtual codeshare products as defined previously. Fourth, following the standard practice for empirical analyses of airline codesharing, we recode

[5] Collecting data from the same quarter in both years will eliminate potential seasonal effects in demand.

regional feeder carriers to have their major carrier codes. In the absence of such recoding of feeder carriers, products that only include a major carrier and its associated regional feeder carrier(s) may mistakenly be counted as codeshare products since the operating and ticketing carrier codes would differ.[6]

Based on our previously stated research objectives, we focus on origin-destination markets in which at least two of the three airlines (Delta, Continental and Northwest) offered competing pure online products both in the pre and post-alliance periods. In other words, the three carriers' networks overlap in all of the markets that remain in our final sample. In addition, similar to Berry (1992) and Aguirregabiria and Ho (2012) among others, we focus on airports in the largest 50 U.S. cities as measured by city population estimates from the U.S. Census Bureau. Table 1 reports a list of the cities and airports included in our sample.

[6] We identify codeshare products as products where the ticketing and operating carriers differ.

Table 1
List of Cities and Airports

City, State	Airports		City, State	Airports
New York City, NY	LGA, JFK		Boston, MA	BOS
Newark, NJ	EWR		Louisville, KY	SDF
Los Angeles, CA	LAX		Washington, DC	DCA, IAD
Chicago, IL	ORD, MDW		Nashville, TN	BNA
Dallas, TX	DFW		Las Vegas, NV	LAS
Phoenix, AZ	PHX		Portland, OR	PDX
Houston, TX	IAH		Oklahoma City, OK	OKC
Philadelphia, PA	PHL		Tucson, AZ	TUS
San Diego, CA	SAN		Albuquerque, NM	ABQ
San Antonio, TX	SAT		New Orleans, LA	MSY
San Jose, CA	SJC		Cleveland, OH	CLE
Detroit, MI	DTW		Sacramento, CA	SMF
Denver, CO	DEN		Kansas City, MO	MCI
Indianapolis, IN	IND		Atlanta, GA	ATL
Jacksonville, FL	JAX		Omaha, NE	OMA
San Francisco, CA	SFO		Oakland, CA	OAK
Columbus, OH	CMH		Tulsa, OK	TUL
Austin, TX	AUS		Miami, FL	MIA
Memphis, TN	MEM		Colorado Springs, CO	COS
Minneapolis & St. Paul, MN	MSP		St. Louis, MO	STL
Baltimore, MD	BWI		Santa Ana, CA	SNA
Charlotte, NC	CLT		Raleigh & Durham, NC	RDU
El Paso, TX	ELP		Pittsburg, PA	PIT
Milwaukee, WI	MKE		Tampa, FL	TPA
Seattle, WA	SEA		Cincinnati, OH	CVG

After applying the above restrictions, we follow Gayle (2007a) and collapsed the data by averaging the price and aggregating the number of passengers purchasing products as defined by unique itinerary-airline(s) combination.[7] In other words, before the data are collapsed, there are several observations of a given itinerary-airline(s) combination that are distinguished by prices paid and number of passengers paying each of those prices. The final sample has 22,485 products contained in 1,170 origin-destination markets that span the pre and post-alliance periods.

Variables that we gathered and constructed from the database include: "Price", "Hub", "Stops", "Inconvenient", "Virtual", "Carrier Presence at Origin" and "Carrier Presence at Destination". These variables are the observable product characteristics. "Price" is the average price paid by

[7] A product remains in our sample only if at least 9 passengers purchase it throughout a quarter. Berry (1992) and Aguirregabiria and Ho (2012) among others use 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.

passengers who chose the specific itinerary-airline(s) combination. "Hub" is a zero-one dummy variable that takes the value one if the origin airport is a hub for the ticketing carrier. "Stops" is a variable that counts the number of intermediate stops associated with each product. For example, in the case of products that use non-stop flight itineraries, "Stops" takes the value zero. "Inconvenient" is the ratio of itinerary distance to the non-stop distance between origin and destination airports. The presumption is that an itinerary is less convenient the further its "Inconvenient" measure is from 1. "Virtual" is a zero-one dummy variable that takes the value one if the product is virtual codeshared. Both the "Carrier Presence at Origin" and "Carrier Presence at Destination" variables are airline-specific and vary across markets for each airline. "Carrier Presence at Origin" measures the number of different cities that an airline has non-stop flights from going into the origin city of the market, while "Carrier Presence at Destination" measures the number of different cities that the airline serves using non-stop flights from the destination city of the market. We leave discussing the rationale for using each of these variables until the results section since the main task now is to provide descriptive information on the data.

As in Berry and Jia (2010) and Berry, Carnal and Spiller (2006), we measure a market's size (subsequently denoted by M) by the geometric mean of population sizes across the origin and destination cities of the market. An air travel product's quantity sold (subsequently denoted by q_j) is the total number of passengers that purchase each specific itinerary-airline(s) combination. Therefore, a product's observed market share (subsequently denoted by upper case letter S_j) is computed as quantity of the product sold divided by our measure of market size, i.e. $S_j = \frac{q_j}{M}$.[8] How we use information on each product's observed market share will become clear after the econometric model and estimation procedure are discussed.

Table 2 provides a list of the airlines in the sample according to type of products the airlines are involved in. Table 3 reports sample summary statistics of the variables.

[8] We find that our measure of market size results in product shares that are extremely small. As such, we scaled up all product shares by a common factor. The common factor is the largest integer such that the share of the outside good ($S_0 = 1 - \sum_{j=1}^{J} S_j$) remains positive in all markets. In our data set the common factor is 42. We perform econometric estimations with and without scaling up product shares and find that econometric estimates are qualitatively similar.

Table 2
List of Airlines in the Data Set

Airlines Involved in Virtual Codeshare Products		Airlines Involved in Pure Online Products	
Airline Name	**Code**	**Airline Name**	**Code**
Alaska Airlines Inc.	AS	American Airlines Inc.	AA
Continental Air Lines Inc.	CO	Alaska Airlines Inc.	AS
Delta Air Lines Inc.	DL	JetBlue Airways	B6
Northwest Airlines Inc.	NW	Continental Air Lines Inc.	CO
United Air Lines Inc.	UA	Delta Air Lines Inc.	DL
US Airways Inc.	US	Frontier Airlines	F9
		AirTran Airways	FL
		America West Airlines	HP
		National Airlines	N7
		Spirit Air Lines	NK
		Northwest Airlines Inc.	NW
		Chautauqua Airlines	RP
		Sun Country Airlines	SY
		ATA Airlines	TZ
		United Air Lines Inc.	UA
		US Airways Inc.	US
		Midwest Airline	YX

Notes: Note that feeder carriers such as Chautauqua Airlines are not listed as involved in codeshare products. This is because we assign these carriers their major carrier codes (effectively not making a distinction between feeder and major carriers) for products where feeder carriers operate segment(s) of the trip but the ticketing carrier is the major carrier. However, the feeder carriers do offer pure online products, which is why they show up in the column labeled "Airlines involved in Pure Online Products". In the data section of the text we provide discussion on the rationale for assigning feeder carriers their major carrier code prior to identifying codeshare products.

Table 3
Summary statistics of variables

Variable	Mean	Std. Dev.	Min	Max
Price ($)	218.36	67.79	101.37	856.63
HUB	0.15	0.36	0	1
Stops	0.84	0.39	0	2
Inconvenient	1.12	0.18	1	2.65
Virtual	0.031	0.17	0	1
Carrier Presence at Origin	22.91	24.28	0	130
Carrier Presence at Destination	26.53	25.59	1	143
Market Size (mean population across the endpoint cities of a market)	901,784.90	687,820.50	264,747	5,439,591
Market nonstop flight distance (miles)	1,479.79	609.26	190	2,724

3.1 Preliminary Descriptive Analysis

Following many event studies [for example see Borenstein (1990) and Kim and Singal (1993)], we begin by using a difference-in-differences approach to get a sense of before and after relative changes in key variables of interest. At this point only descriptive evidence is being developed on the key variables. A more careful analysis of the relevant issues is laid out across subsequent sections of the paper.

In our study the relevant event is implementation of the codeshare alliance. Therefore, the difference-in-differences approach seeks to identify effects associated with implementation of the codeshare alliance based on the extent to which variables of interest change before and after implementation of the codeshare alliance across markets that should be impacted by the alliance ("treatment" markets) versus markets that should not be impacted by the alliance ("control" markets). Our treatment markets are origin-destination markets in which Delta, Continental and Northwest codeshare together during the post-alliance period, while our control markets are origin-destination markets that the three airlines compete in but do not codeshare together during the post-alliance period. Among the 1,170 origin-destination markets in the data set, the three airlines virtual codeshare together in 852 of the markets, and therefore compete but did not virtual codeshare together in 318 of the markets.

A variable of interest that we apply the difference-in-differences approach to is the three airlines' average price. Specifically, before and after relative change in Delta, Continental and Northwest average price is computed by:

$$\Delta DCN\_price = \log\left(\frac{DCN\_price^{Codeshare\_mkt}_{post-alli-period}}{DCN\_price^{Codeshare\_mkt}_{pre-alli-period}}\right) - \log\left(\frac{DCN\_price^{Non-Codeshare\_mkt}_{post-alli-period}}{DCN\_price^{Non-Codeshare\_mkt}_{pre-alli-period}}\right), \quad (1)$$

where subscripts $post-alli-period$ and $pre-alli-period$ refer to the time period used for computing the variable; the superscript $Codeshare\_mkt$ refers to origin-destination markets in which Delta, Continental and Northwest codeshare together during the post-alliance period; while superscript $Non-Codeshare\_mkt$ refers to origin-destination markets that the three airlines compete in but do not codeshare together during the post-alliance period. Therefore, $DCN\_price^{Codeshare\_mkt}_{pre-alli-period}$ represents Delta, Continental and Northwest average price during the pre-alliance period in origin-destination markets that they eventually codeshare together in during the post-alliance period; $DCN\_price^{Codeshare\_mkt}_{post-alli-period}$ represents the three airlines average price during the post-alliance period in origin-destination markets that they codeshare together during the post-alliance period;

$DCN\_price^{Non-Codeshare\_mkt}_{post-alli-period}$ represents the three airlines average price during the post-alliance period in origin-destination markets that they compete in but do not codeshare together during the post-alliance period; while $DCN\_price^{Non-Codeshare\_mkt}_{pre-alli-period}$ represents the three airlines average price during the pre-alliance period in origin-destination markets that they compete in but do not codeshare together during the post-alliance period.

Analogous to equation (1), we specify before and after relative changes in the three airlines' joint passenger traffic and joint passenger share as follows:

$$\begin{aligned}\Delta DCN\_total\_pass &= \log\left(\frac{DCN\_total\_pass^{Codeshare\_mkt}_{post-alli-period}}{DCN\_total\_pass^{Codeshare\_mkt}_{pre-alli-period}}\right) \\ &\quad - \log\left(\frac{DCN\_total\_pass^{Non-Codeshare\_mkt}_{post-alli-period}}{DCN\_total\_pass^{Non-Codeshare\_mkt}_{pre-alli-period}}\right),\end{aligned} \tag{2}$$

$$\begin{aligned}\Delta DCN\_pass\_share &= \log\left(\frac{DCN\_pass\_share^{Codeshare\_mkt}_{post-alli-period}}{DCN\_pass\_share^{Codeshare\_mkt}_{pre-alli-period}}\right) \\ &\quad - \log\left(\frac{DCN\_pass\_share^{Non-Codeshare\_mkt}_{post-alli-period}}{DCN\_pass\_share^{Non-Codeshare\_mkt}_{pre-alli-period}}\right).\end{aligned} \tag{3}$$

The before and after relative change in the three airlines' average price, ΔDCN_price, is 0.0179. One way to interpret this before and after relative price change is that changes in the three airlines' average price leave average price 1.79% higher in their codeshare markets relative to their non-codeshare markets. Before and after relative change in the three airlines total passenger traffic, ΔDCN_total_pass, is -0.018. Therefore, before and after changes in the three airlines' passenger traffic leave their passenger traffic 1.8% lower in their codeshare markets relative to their non-codeshare markets. The direction of the relative price and passenger traffic changes suggest that collusive effects could be associated with virtual codesharing between the three airlines in their overlapping markets.

Before and after relative change in the three airlines joint passenger share, ΔDCN_pass_share, is 0.019. Therefore, changes in the three airlines' joint passenger share leave their joint passenger share 1.9% higher in their codeshare markets relative to their non-codeshare markets. So even though the partner airlines' passenger traffic declined in their codeshare markets relative to their non-codeshare markets, the partners end up making relative gains in passenger share in their codeshare markets since other airlines' passengers traffic fell by more in these markets. This result suggest that there could be a demand-increasing effect associated with virtual codesharing, which in this case resulted in increase passenger share via slower decline in passenger traffic.

It must be noted that the difference-in-differences analysis captured by equations (1), (2) and (3), has caveats and provide only rough estimates of the effects associated with virtual codesharing between the three airlines. For example, these difference-in-differences computations do not control for persistent demand or cost conditions/shocks that may differ across codeshare versus non-codeshare markets. In evaluating the market effects associated with virtual codesharing between the three airlines, the formal econometric model presented below, while not perfect, will do a better job at controlling for potential differences in demand and cost conditions across codeshare versus non-codeshare markets.

Last, it is also useful to get a sense of exogenous characteristics of origin-destination markets that may influence the three airlines' choice of markets in which to virtual codeshare together during the post-alliance period. For this descriptive analysis we rely on a reduced-form logit regression model that uses exogenous market characteristics to explain the three alliance partners' codeshare versus non-codeshare markets. The variable being explained by the logit regression is denoted, $Codeshare\_mkt$, which is a zero-one indicator variable that only takes the value 1 if the three alliance partners virtual codeshare together in the origin-destination market during the post-alliance period. Results from this logit regression are reported in Table 4. The unit of observation for data used in the regression is origin-destination level.

Table 4
Reduced-form Codeshare Market Logit Regression

Dependent Variable: *Codeshare_mkt* = 1 if alliance partners virtual codeshare together in origin-destination market during the post-alliance period; otherwise *Codeshare_mkt* = 0.		
Variable	**Coefficient**	**Standard Error**
Constant	0.89	1.45
Market Size (measured in 10,000,000 people)	-52.41**	11.41
(Market Size)^2	64.97**	21.52
Market Nonstop Flight Distance (measured in 10,000 miles)	35.85**	9.75
(Market Nonstop Flight Distance)^2	-45.68	31.18
Market origin fixed effects	Yes	
Market destination fixed effects	Yes	
Pseudo R^2	0.4752	
Log likelihood	-359.20	
Number of Observations	1170	

Notes: ** indicates statistical significance at the 1% level. Model is estimated with market origin dummies and market destination dummies even though these dummy coefficients are not reported in the table.

In addition to observed market characteristics such as market size and nonstop flight distance, the regression in Table 4 also controls for unobserved (to the researchers) market endpoint characteristics using a set of dummy variables for origin fixed effects and destination fixed effects. Due to economy of presentation purposes, the coefficient estimates on these dummy variables are not reported in the table. The coefficient estimates on $Market\ Size$ and $(Market\ Size)^2$ suggest that markets with mean endpoint population greater than 4,033,400 people[9] are more likely to be codeshare markets. Also, the coefficient estimates on $Market\ Nonstop\ Flight\ Distance$ and $(Market\ Nonstop\ Flight\ Distance)^2$ suggest that the probability of a market being a codeshare market increases monotonically with nonstop flight distance between the origin and destination.

There is evidence that the regressors jointly do a good job in explaining the $Codeshare\_mkt$ variable. For example, the Pseudo R^2 of the logit regression is 0.4752, suggesting that almost 50% of the variation in $Codeshare\_mkt$ is jointly explained by the regressors. Second, the fitted values of the dependent variable from the logit regression, i.e. $Codeshare\_mkt\_hat = Prob(Codeshare\_mkt = 1)$, has a 0.717 correlation with $Codeshare\_mkt$.

[9] This population threshold is computed using the coefficient estimates on $Market\ Size$ and $(Market\ Size)^2$. Specifically, the population threshold is computed by: $10,000,000 \times \frac{52.41}{2\times 64.97}$.

4 The Model

We proceed by first describing the demand-side of the model. The supply-side is then laid out, which is where we model competitive interactions between airlines.[10]

4.1 Demand

In the spirit of Peters (2006), Berry, Carnall and Spiller (2006), Gayle (2007a, 2007b, 2013), Gayle and Wu (2012), Armantier and Richard (2008) and Berry and Jia (2010) among others, air travel demand is modeled using a discrete choice framework. Specifically, we use a nested logit model.[11] Potential passenger i in market l during time period τ faces a choice between $J_{\tau l}+1$ alternatives. There are $J_{\tau l}+1$ alternatives because we allow passengers the option ($j=0$, the outside good) not to choose either one of the $J_{\tau l}$ differentiated air travel products considered in the empirical model.

Products in a market are assumed to be organized into $G+1$ exhaustive mutually exclusive groups/nests, $g=0,1,...,G$, in which the outside good, $j=0$, is assumed to be the only member of group 0. A group or nest here refers to the set of products offered by an airline within a market. We explore alternate nesting structures in an appendix available upon request.

A passenger solves the following optimization problem:

$$\underset{j\in\{0,\ldots,J_{\tau l}\}}{Max}\left\{U_{ij\tau l}=\delta_{j\tau l}+\sigma\zeta_{i\tau lg}+(1-\sigma)\,\varepsilon_{ij\tau l}\right\}, \tag{4}$$

where $U_{ij\tau l}$ is the level of utility passenger i will obtain if product j is chosen, while $\delta_{j\tau l}$ is the mean level of utility across passengers that consume product j. $\delta_{j\tau l}$ is a function of the characteristics of product j, which we subsequently describe. $\zeta_{i\tau lg}$ is a random component of utility that is common to all products in group g, whereas the random term $\varepsilon_{ij\tau l}$ is specific to product j and is assumed to have an extreme value distribution. The parameter σ lies between 0 and 1, and measures the correlation of the consumers' utility across products belonging to the same group. Since products

[10] Armantier and Richard (2008) also use a structural econometric model to examine a codeshare alliance. However, a fundamental difference between our model and the model in Armantier and Richard (2008) is that we model both demand and supply aspects of codesharing, while Armantier and Richard (2008) only model the demand side. This crucial methodological difference affords us the advantage of being able to separately identify demand and supply effects of codesharing, which further allows us to more meticulously examine short-run market effects within a market equilibrium framework.

[11] We concede that a nested logit model is not as flexible and therefore less desirable compared to a random coefficients logit model. However, it is well-known that the random coefficients model is more computationally demanding to estimate relative to the nested logit model. As we discuss further in the results section, our nested logit demand model provides elasticity estimates that are comparable to much of the literature, including papers that use a random coefficients logit specification. As such, we decide to go with the less computationally intensive nested logit model. For checks of robustness of qualitative results we explore alternate nesting structures, as further discussed in an appendix available upon request.

are grouped by airlines, σ can also be thought of as measuring the correlation of the consumers' utility across products offered by a given airline. As σ approaches 1, the correlation of preferences among products offered by the same airline within a market increases. Conversely, as σ decreases, the correlation of preferences for products offered by the same airline within a market decreases.

The rationale for the product grouping structure above is to capture the possibility that passengers view an airline's products as closer substitutes for each other compared to the substitutability of these products across airlines [Gayle (2007b)]. One reason why this could be the case is that a passenger may be heavily invested (accumulated miles flown) in a given airline's frequent-flyer program and therefore, on the margin, would prefer to choose among alternate flights offered by this airline in order to build up accumulated miles towards the required threshold necessary for a discounted trip. Second, some consumers may just have a strong brand-loyalty to a given airline based on past experience. In any event, since σ is a parameter we estimate, the data will reveal whether or not a sufficient number of passengers are brand-loyal to render $\sigma > 0$.

The mean level of utility obtained across the population of consumers that consume product j is given by:

$$\begin{aligned}\delta_{j\tau l} \;=\; & x_{j\tau l}\beta - \alpha p_{j\tau l} + a_r + mkt_l^{origin} + mkt_l^{dest} + \lambda_0 Codeshare\_mkt \\ & +\lambda_1 DCN \times Codeshare\_mkt + \lambda_2 T + \lambda_3 T \times Codeshare\_mkt \\ & +\lambda_4 T \times DCN + \lambda_5 T \times DCN \times Codeshare\_mkt \\ & +\lambda_6 T \times DCN \times Codeshare\_mkt \times DCN\_pre - alli\_pass\_share + \xi_{j\tau l},\end{aligned} \tag{5}$$

where $x_{j\tau l}$ is a vector of observed product characteristics ["Stops" - the number of intermediate stops used by an itinerary; "Inconvenient" - the ratio of itinerary distance to the market non-stop distance; "Hub" - a zero-one dummy variable that takes the value one if the origin airport is a hub for the carrier offering the product for sale; "Virtual" - a zero-one dummy that takes the value one if the product is virtual codeshared], β is a vector of consumer taste parameters (marginal utilities) associated with the product characteristics in $x_{j\tau l}$, $p_{j\tau l}$ is the price of product j, α represents the marginal utility of price, a_r are airline fixed effects, where subscript r indexes ticketing carriers (ticketing carrier dummies), mkt_l^{origin} are market origin fixed effects, mkt_l^{dest} are market destination fixed effects, $Codeshare\_mkt$ is a zero-one dummy which is equal to 1 if a virtual codeshare product between Delta, Continental or Northwest was offered in the origin-destination market, T is a zero-one time dummy which is equal to 1 if the itinerary occurred in the post-alliance period, DCN is a zero-one dummy which is equal to 1 if product j is being offered for sale by either Delta, Continental or Northwest, $DCN\_pre-alli\_pass\_share$ is the pre-alliance joint passenger share of Delta, Continental and Northwest in the origin-destination market, and $\xi_{j\tau l}$ captures unobserved (by the econometricians but observed by passengers) product characteristics.

It is likely that there exists several non-price characteristics that are responsible for passengers' choice of one product over others, where these non-price characteristics are observed by passengers and airlines but not by us the researchers given limitations of the data available. This is the rationale for including $\xi_{j\tau l}$ in the demand model, i.e., the inclusion of $\xi_{j\tau l}$ effectively acknowledges that there will be passenger choice behavior outcomes observed in the data that cannot be fully explained by the measured product characteristics in the data.

λ_0, λ_1, λ_2, λ_3, λ_4, λ_5, and λ_6 are taste parameters to be estimated. λ_0 captures any persistent difference in mean utility for non-Delta/Continental/Northwest products across markets in which the three airlines eventually virtual codeshare together compared to markets in which they compete but do not codeshare together. Likewise, λ_1 captures any persistent difference in mean utility for the three airlines' products across markets in which the three carriers eventually virtual codeshare together compared to markets in which they compete but do not codeshare together. We therefore control for any persistent systematic difference across the three airlines' codeshare versus non-codeshare markets that may affect demand.

λ_2 captures the change in mean utility over the pre and post-alliance periods for products offered by airlines other than Delta, Continental or Northwest, while λ_3 captures whether this change in mean utility for other airlines' products differs across the three airlines codeshare versus non-codeshare markets. λ_4 captures the change in mean utility over the pre and post-alliance periods for products offered by Delta, Continental or Northwest, while λ_5 captures whether this change in mean utility for the three airlines' products differs across markets in which they virtual codeshare together versus markets in which they compete but do not virtual codeshare together. In other words, $\lambda_5 > 0$ implies that virtual codesharing has a demand-increasing effect associated with it, which is one of the main hypotheses we want to test. Last, λ_6 captures whether or not the demand effect of virtual codesharing depends on the size of the partner airlines' pre-alliance joint passenger share in a market that they eventually begin to codeshare in.

As we previously discussed, frequent-flyer membership with any one of the three carriers suddenly becomes more valuable with implementation of the codeshare alliance, since the alliance allows frequent-flyer members of any one of the three carriers to accumulate and redeem frequent-flyer points across any of the three partner carriers. The larger is the pre-alliance joint passenger share of Delta, Continental and Northwest in an origin-destination market, then we should expect a larger proportion of consumers in the market to have frequent-flyer membership with at least one of the three airlines. If this argument holds true, then we should expect $\lambda_6 > 0$.

The discussion above reveals that a key component of our demand specification that allows us to identify demand effects associated with the Delta/Continental/Northwest codeshare alliance (λ_5 and λ_6), is that equation (5) effectively compares consumers' choice behavior before and after implementation of the alliance in markets where the three airlines virtual codeshare together ("treatment" markets) versus markets in which they compete but do not virtual codeshare together ("control" markets). A reasonable criticism to raise at this point is that $Codeshare\_mkt$ in equation (5) is not strictly exogenous since airlines choose the markets in which to codeshare. The reader will subsequently observe that we do account for the possible endogeneity of $Codeshare\_mkt$ by replacing this variable with the estimated $\Pr ob(Codeshare\_mkt = 1)$ obtained from the previously discussed logit regression in Table 4. Therefore, the logit regression in Table 4 serves as one first-stage reduced-form regression that is used to account for possible endogeneity when estimating the structural demand model.

Finally, the demand for product j is given by,

$$d_j = M \times s_j(\mathbf{x}, \mathbf{p}, \boldsymbol{\xi}; \theta_d),$$

where M is a measure of market size, which we assume to be the geometric mean of population sizes across the origin and destination cities of the market, $s_j(\cdot)$ is the predicted product share function based on the nested logit model, [12] $\mathbf{x}$ and $\mathbf{p}$ are vectors of observed non-price product characteristics and price, respectively, $\boldsymbol{\xi}$ is a vector of unobserved (by the researchers) product characteristics, and $\theta_d = (\beta, \alpha, \lambda, \sigma)$ is the vector of demand parameters to be estimated. We dropped the market and time subscripts (l and τ) only to avoid a clutter of notation.

[12] The well-known formula for the predicted share function in the case of the nested logit model is: $s_j = \frac{\exp\left(\frac{\delta_j}{(1-\sigma)}\right)}{D_g^{\sigma}\left[1+\sum_{g=1}^{G} D_g^{(1-\sigma)}\right]}$, where δ_j is the previously discussed mean level of utility obtained from consuming product j, $D_g = \sum_{j \in \mathcal{G}_g} \exp\left(\frac{\delta_j}{(1-\sigma)}\right)$, and $\mathcal{G}_g$ is the set of products in group g.

4.2 Supply

What is commonly known about how a codeshare agreement works is that the ticketing carrier markets and sets the final price for the round-trip ticket and compensates the operating carrier for operating services provided. Details on compensation mechanisms actually used by partner airlines are not usually made known to the public and may even vary across partnerships. Therefore, we face the challenge of coming up with 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. We concede that the following is possibly a simplistic approximation of the actual contracting used by partners to compensate each other for services needed to provide a codeshare product.

One way to proceed, as pointed out in Chen and Gayle (2007) and Gayle (2013), is to think of a codeshare agreement as a privately negotiated pricing contract between partners (w, Γ), where w is a per-passenger price the ticketing carrier pays over to an operating carrier for transporting the passenger, while Γ represents a potential lump sum transfer between partners that determines how the joint surplus is distributed. As we develop the supply-side of the model further, it will become clear that only the level of w affects equilibrium final product prices. Since for the purposes of this paper we are not concerned how the surplus is distributed between partners through the lump sum transfer Γ, we do not attempt to derive an equilibrium value of Γ.[13]

Assume that the final price of a codeshare product is determined within a sequential price-setting game. In the first stage of the sequential process, the operating carrier sets the price for transporting a passenger, w, and privately makes it 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.

Let $r = 1, ..., R$ index competing ticketing carriers in a market and let $f = 1, ..., F$ index the corresponding operating carriers. Further, let $\mathcal{F}_r$ be a subset of the J products, both pure online and virtual codeshare, that are offered for sale by ticketing carrier r in the origin-destination

[13] See Chen and Gayle (2007) for a similar theoretical modeling approach of an airline codeshare agreement.

market.[14] Carrier r solves the following profit maximization problem for each $j \in \mathcal{F}_r$:

$$\underset{p_j}{Max} \sum_{j \in \mathcal{F}_r} (p_j - z_j) q_j, \tag{6}$$

where $q_j = d_j(\mathbf{p})$ in equilibrium, q_j is the quantity of product j offered for sale on the market, $d_j(\mathbf{p})$ is market demand for product j, $\mathbf{p}$ represents a $J \times 1$ vector of final prices, and z_j is the effective marginal cost that ticketing carrier r incurs by offering product j for sale. In the event that product j is a codeshare product, then $z_j = w_j^f$, where w_j^f is the price the ticketing carrier pays to operating carrier f for its transportation services.[15] On the other hand, if product j is a pure online product, then $z_j = c_j^r$, where c_j^r is the marginal cost that carrier r incurs by using its own plane(s) to provide product j. Note that in the pure online product case $f = r$ since carrier r is the sole ticketing and operating carrier of product j.

We posit that the marginal cost function is given by:

$$z_j = W_j \gamma + a_f + \eta_j, \tag{7}$$

where W_j is a vector of variables that shift marginal cost ("Itinerary Distance", "Carrier Presence at Origin", "Carrier Presence at Destination", market origin fixed effects, and market destination fixed effects) and γ is the associated vector of parameters, a_f captures operating carrier-specific portion of marginal cost, and η_j is a mean-zero, random error term that captures unobserved determinants of marginal cost. When product j is pure online, implying that $z_j = c_j^r$, then equation (7) simply relates a carrier's own marginal cost of providing a product to factors that influence this marginal cost. On the other hand, if product j is virtual codeshared, implying that $z_j = w_j^f$, then equation (7) is saying that w_j^f depends on factors that influence the marginal cost of the carrier that provides operating services for the codeshare product. This is an implication of the assumed sequential price-setting game that determines equilibrium prices of codeshare products. The reason is as follows. In the first stage of the sequential price-setting game, operating carriers each optimally choose w_j^f. Therefore, the equilibrium level of w_j^f in this first stage game depends on the marginal cost of the operating carrier that offers transportation services for codeshare product j. So, like c_j^r, w_j^f is a function of factors that shift the marginal cost of the operating carrier. As

[14] For most of the subsequent equations, we intentionally omit a market subscript for variables and equations only to avoid a notation clutter. Notwithstanding our omission of market subscripts, the reader should continue to interpret equations in a market-specific way.

[15] We implicitly assume here that the ticketing carrier of a virtual codeshare product only incurs fixed expenses in marketing the product to potential passengers.

such, the marginal cost function is effectively:

$$W_j\gamma + a_f + \eta_j = \begin{cases} w_j^f & \text{if } j \text{ is virtual codeshare} \\ \\ c_j^r & \text{if } j \text{ is pure online} \end{cases} \tag{8}$$

A pure strategy Nash equilibrium in final prices requires that p_j of any product j offered by carrier r must satisfy the first-order condition:

$$d_j(\mathbf{p}) + \sum_{k \in \mathcal{F}_r} (p_k - z_k) \frac{\partial d_k(\mathbf{p})}{\partial p_j} = 0.$$

The first-order conditions are a set of J equations, one for each product. A few additional definitions allow for a more convenient representation of the first-order conditions using matrix notation.[16]

First, let Ω be a $J \times J$ matrix which describes the ticketing carriers' ownership structure of the J products. Let $\Omega(j,k)$ denote an element in Ω, where

$$\Omega(j,k) = \begin{cases} 1 & \text{if products } k \text{ and } j \text{ are offered by the same ticketing carrier} \\ \\ 0 & \text{otherwise.} \end{cases} .$$

Second, let $\triangle$ be a $J \times J$ matrix of first-order derivatives of product market shares with respect to final prices, where element $\triangle(j,k) = \frac{\partial d_k}{\partial p_j}$. In vector notation, the system of J first-order conditions for the ticketing carriers can now conveniently be expressed as:

$$\mathbf{d}(\mathbf{p}) + (\Omega. * \triangle)(\mathbf{p} - \mathbf{z}) = 0, \tag{9}$$

where $\mathbf{d}(\cdot)$, $\mathbf{p}$, and $\mathbf{z}$ are $J \times 1$ vectors of product demands, final prices, and ticketing carriers' effective marginal costs, respectively, while $.*$ means element-by-element multiplication of two matrices. Equation (9) implies the following product markups:

$$\mathbf{mkup}(\beta, \alpha, \lambda, \sigma, \Omega) = \mathbf{p} - \mathbf{z} = -(\Omega. * \triangle)^{-1} \mathbf{d}(\mathbf{p}), \tag{10}$$

which reveals that product markups are a function of demand parameters and the product ownership structure matrix.

In the event that the codeshare alliance allows Delta, Continental and Northwest to practice collusive pricing in markets where they codeshare together during the post-alliance period, then we can account for such collusive pricing behavior by appropriately modifying the product ownership

[16] See Nevo (2000) for similar notation in a merger analysis setting.

structure matrix. In particular, let $\Omega^{Collude}$ be the modified $J \times J$ product ownership structure matrix in which the three alliance partners are treated as a single carrier rather than distinct carriers. Let $\Omega^{Collude}(j,k)$ denote an element in $\Omega^{Collude}$, where

$$\Omega^{Collude}(j,k) = \begin{cases} 1 & \begin{array}{c}\text{if distinct products } k \text{ and } j \text{ are offered by the same ticketing carrier,} \\ \text{where Delta, Continental and Northwest are treated as a single carrier}\end{array} \\ & \\ 0 & \text{otherwise.} \end{cases}.$$

Therefore, under collusive alliance pricing the appropriate first-order conditions in markets where the three airlines codeshare together during the post-alliance period are:

$$\mathbf{d}(\mathbf{p}) + \left(\Omega^{Collude}.*\triangle\right)(\mathbf{p}-\mathbf{z}) = 0, \tag{11}$$

where Ω in equation (9) is replaced with $\Omega^{Collude}$ to obtain equation (11). Product markups under collusive alliance pricing are:

$$\mathbf{mkup}^{Collude}\left(\beta,\alpha,\lambda,\sigma,\Omega^{Collude}\right) = -\left(\Omega^{Collude}.*\triangle\right)^{-1}\mathbf{d}(\mathbf{p}), \tag{12}$$

4.2.1 Alternate Supply Equation Specifications

At this point we do not know whether the three alliance partners practice collusive pricing, which further implies that we do not know which product markup specification, equation (10) versus equation (12), is most appropriate to characterize pricing behavior. If the codeshare alliance does not allow Delta, Continental and Northwest to practice collusive pricing in the markets where they codeshare together during the post-alliance period, then the appropriate parametric supply equation specification, which we define as *Model h*, is given by:

$$\textit{Model } h : p_j = W_j\gamma_h + a_f + \eta_j + mkup_j, \tag{13}$$

where η_j is the structural supply error term, and the product markup variable, $mkup_j$, is computed based on equation (10). On the other hand, if the codeshare alliance allows Delta, Continental and Northwest to practice collusive pricing in markets where they codeshare together during the post-alliance period, then the following parametric supply equation specification, which we define as *Model g*, should provide a better statistical fit of the data compared to *Model h*:

$$\textit{Model } g : p_j = W_j\gamma_g + a_f + \eta_j + mkup_j^{Collude}, \tag{14}$$

where the product markup variable, $mkup_j^{Collude}$, is computed based on equation (12).

We first estimate the demand parameters, use these demand parameter estimates to compute product markups under each alternate pricing behavior ($mkup_j$ versus $mkup_j^{Collude}$), then use these product markups as variables when estimating the alternate supply equations, *Model h* and *Model g*. Finally, in the spirit of Villas-Boas (2007), we use non-nested statistical tests based on Vuong (1989) to see which supply specification best fits the data. Note that the estimated markups ($mkup_j$ versus $mkup_j^{Collude}$) are different under each alternate pricing behavior, as such, the competing estimated supply equations are not nested, which is why a non-nested statistical test is needed to evaluate which supply model best fits the data.

5 Estimation

The parameters to be estimated are $\theta_d = (\beta, \alpha, \lambda, \sigma)$ for demand and γ for marginal cost. Following Berry (1994), the estimation strategy for demand parameters involves choosing parameter values such that *observed* product shares, S_j, are equal to *predicted* product shares, s_j, that is,

$$S_j = s_j\left(\delta, \sigma\right), \ \ \forall \ j. \tag{15}$$

As previously stated in the data section, observed product shares are computed by $S_j = \frac{q_j}{M}$. In the case where the predicted share function, $s_j\left(\cdot\right)$, is based on the nested logit model, the above estimation strategy yields the following well-known linear estimating equation:

$$\begin{aligned}
\ln\left(S_j\right) - \ln\left(S_0\right) \ = \ & x_j\beta - \alpha p_j + \sigma \ln\left(S_{j|g}\right) + a_r + mkt^{origin} + mkt^{dest} + \lambda_0 Codeshare_mkt \\
& + \lambda_1 DCN \times Codeshare_mkt + \lambda_2 T + \lambda_3 T \times Codeshare_mkt \\
& + \lambda_4 T \times DCN + \lambda_5 T \times DCN \times Codeshare_mkt \\
& + \lambda_6 T \times DCN \times Codeshare_mkt \times DCN_pre - alli_pass_share + \xi_j,
\end{aligned} \tag{16}$$

where S_0 is the observed share of the outside option, $S_{j|g}$ is the observed within group share of product j, and ξ_j is the structural demand error term.[17]

Provided we have valid instruments for p_j and $S_{j|g}$, equation (16) is straightforward to estimate using a linear instrumental variables technique such as two-stage least squares (2SLS), which is the estimator we use. As previously discussed, we also instrument for variables associated with $Codeshare\_mkt$ using $Prob(Codeshare\_mkt = 1)$ to replace $Codeshare\_mkt$, where

[17] The observed share of the outside option is computed by $S_0 = 1 - \sum_{g=1}^{G} S_g$, where S_g is computed by $\sum_{j \in \mathcal{G}_g} S_j$. The observed within group share of product j is computed by $S_{j/g} = \frac{\mathbf{S}_j}{\mathbf{S}_g}$.

$Prob(Codeshare\_mkt = 1)$ is computed from a previously estimated reduced-form logit model reported in Table 4.

Supply *Model h* and *Model g* can be re-arranged as $p_j - mkup_j = W_j\gamma_h + a_f + \eta_j$ and $p_j - mkup_j^{Collude} = W_j\gamma_g + a_f + \eta_j$, where $p_j - mkup_j$ and $p_j - mkup_j^{Collude}$ are effectively the dependent variables for the supply regressions respectively. Once we use the estimated demand parameters to compute alternate product markups, $mkup_j$ and $mkup_j^{Collude}$, the dependent variables for the re-arranged supply equations can be constructed, and then marginal cost parameters, γ_h and γ_g, can be estimated consistently using ordinary least squares.

An alternate estimation strategy would be to estimate the demand and marginal cost parameters jointly. However, a crucial objective of the analysis is to figure out what is the most appropriate specification for the supply equation - Bertrand Nash versus collusive pricing by the partner carriers. In other words, the correct specification of the supply equation is unclear *a priori*. An incorrectly specified supply equation could introduce bias in demand parameter estimates when demand and marginal cost parameters are jointly estimated. Therefore, in our case it is preferable to estimate the demand parameters separately from the marginal cost parameters. Villas-Boas (2007) also recommends separately estimating demand and marginal cost parameters when the correct specification of the supply equation is unclear.

5.1 Instruments

We recognize that a product's price and its within group share (p_j and $S_{j|g}$ respectively) are likely to be correlated with the residual portion of the product's quality captured in ξ_j (where ξ_j is unobserved to the researchers but observed to passengers and airlines). As such, we need to find instruments for p_j and $S_{j|g}$ in equation (16). We make the well-known identifying assumption found in the literature on discrete choice models of demand that observed non-price product characteristics are uncorrelated with the residual portion of product quality left in ξ_j.[18] In other words, given that airline fixed effect, market origin fixed effects, and market destination fixed effects are controlled for in the regression, then the residual shocks to product quality that are left in ξ_j are unlikely to be correlated with observed non-price product characteristics. This allows us to use various combinations of non-price product characteristics to form valid instruments for p_j and $S_{j|g}$.

The instruments we use include: (1) itinerary distance; (2) the number of competing products

[18] For example, see Berry and Jia (2010) and Peters (2006) for similar identifying assumptions.

offered by other airlines with equivalent number of intermediate stops; (3) the number of competitor products in the market; (4) the number of other products offered by an airline in a market; and (5) the sums and averages, by airlines in a market, of the "Inconvenient" and "Stops" variables.[19] As described in Gayle (2007a and 2013), instruments (1) to (4) are motivated by supply theory, which predicts that the equilibrium price is affected by changes in marginal cost and changes in product markup. For example, itinerary distance (instrument (1)) is a marginal cost-shifting variable, instruments (2) to (3) proxy for the degree of competition facing a product, which in turn affects the size of a product's markup, and instrument (4) recognizes the fact that the more substitute products an airline offers in a market, *ceteris paribus*, the airline is better able to charge a higher markup on each of these products. Last, instruments in (5) are likely to be correlated with reasons why passengers may prefer the set of products offered by one airline over the set of products offered by other airlines, and therefore serve as instruments for within group product shares.

6 Results

6.1 Demand Equation Estimates

Results from the demand estimation are reported in Table 5. Estimation A in Table 5 shows ordinary least squares (OLS) estimates. OLS estimation ignores that price, within group product share, and variables associated with the $Codeshare\_mkt$ variable are likely endogenous, and therefore coefficient estimates associated with these variables are most likely biased. In fact, an immediate red flag is that the OLS estimate of the coefficient on price is positive, which is contrary to standard demand theory.

Estimation B and Estimation C each uses two-stage least squares (2SLS) to account for suspected endogeneity. Estimation B only takes into account the suspected endogeneity of price and within group product share, while Estimation C takes into account all suspected endogenous variables. In Estimation C the predicted probability variable $Prob(Codeshare\_mkt = 1)$, which is obtained from the previously estimated reduced-form logit model in Table 4, is used to replace $Codeshare\_mkt$ in each demand equation regressor associated with this variable, while in Estimation B variable $Codeshare\_mkt$ is used directly in each demand equation regressor associated with it.

We first evaluate the endogeneity of price and within group product share (p_j and $S_{j|g}$ respec-

[19] See the data section for definition and explanation of the "Inconvenient" and "Stops" variables.

tively) by using a Hausman statistical test to compare estimates from Estimation A and Estimation B. The endogeneity of variables associated with the *Codeshare_mkt* variable are then evaluated by again using a Hausman test to compare estimates from Estimation B and Estimation C. The Hausman test in each case confirms, at conventional levels of statistical significance, that the variables suspected to be endogenous are indeed endogenous.[20] As such, the following discussion of results in Table 5 is based on Estimation C.

First, as expected, an air travel product's price has a negative effect on the utility obtained from choosing the product, *ceteris paribus*. Second, the more intermediate stops an air travel product has, the lower the utility obtained from choosing that product, *ceteris paribus*. The number of intermediate stops that an air travel product has is one measure of the inherent convenience of the travel itinerary - the negative coefficient for "Stops" is consistent with our expectation.

Gayle (2007a) points out that the number of intermediate stops may only capture a portion of the inherent convenience of an itinerary. For example, two itineraries may each have one intermediate stop, but depending on where the intermediate stop is located in relation to origin and destination cities, two one-stop itineraries in the same market may have very different travel distances and travel time associated with them. As such, passengers could view these two itineraries as having very different levels of convenience even though the itineraries have the same number of intermediate stops. Our "Inconvenient" variable, which measures the ratio of itinerary distance to non-stop distance between the origin and destination cities, is supposed to capture aspects of itinerary convenience that are not picked up by number of intermediate stops.[21] We therefore expect the coefficient on "Inconvenient" to be negative, which is indeed the estimated sign in Table 5.

[20] In a first stage OLS regression in which price is the dependent variable and the instruments are the regressors, R^2 is 0.115. When the dependent variable of such a regression is within group product share, R^2 is 0.444. Recall from the data section that the Pseudo R^2 from the reduced-form *Codeshare_mkt* logit regression is 0.475, while the correlation between *Codeshare_mkt* and $Prob(Codeshare\_mkt = 1)$ is 0.717. Therefore, the instruments do have explanatory power of variations in the endogenous variables.

[21] The minimum value that the "Inconvenient" variable can take on is 1. As such, the further an itinerary's "Inconvenient" measure is from 1, the less convenient is the itinerary.

Table 5
Demand Parameter Estimates

	Potential endogeneity of the *Codeshare_mkt* variable not taken into account.				Potential endogeneity of the *Codeshare_mkt* variable taken into account by using its associated fitted values from a first-stage logit regression.	
	Estimation A: Ordinary Least Squares (OLS)		**Estimation B:** Two-Stage Least Squares (2SLS)		**Estimation C:** Two-Stage Least Squares (2SLS)	
Variable	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
Constant	-3.62**	0.092	-1.41**	0.209	-1.60**	0.189
Price (in hundreds of $)	0.21**	0.011	-0.96**	0.096	-0.66**	0.097
$ln\ (S_{j\|g})\ (\sigma)$	0.46**	0.005	0.10**	0.013	0.12**	0.013
Stops	-0.82**	0.017	-1.20**	0.030	-1.13**	0.029
Inconvenient	-1.61**	0.038	-1.48**	0.051	-1.56**	0.049
Hub	0.95**	0.020	0.92**	0.034	0.85**	0.032
Virtual	-0.75**	0.036	-1.28**	0.057	-1.19**	0.054
Codeshare_mkt (λ_0)	-0.11**	0.029	-0.05	0.037	-0.42**	0.078
DCN × Codeshare_mkt (λ_1)	0.26**	0.039	0.15**	0.051	0.26**	0.063
T (λ_2)	0.11**	0.030	0.06	0.039	0.07	0.046
T × Codeshare_mkt (λ_3)	-0.02	0.035	0.02	0.046	0.01	0.06
T × DCN (λ_4)	-0.05	0.047	-0.09	0.062	-0.09	0.072
T × DCN × Codeshare_mkt (λ_5)	-0.52**	0.062	-0.48**	0.081	-0.50**	0.097
T × DCN × Codeshare_mkt × DCN_ Pre-alli_Pass_share (λ_6)	0.99**	0.059	0.98**	0.080	1.03**	0.081
Carrier fixed effects	Yes		Yes		Yes	
Market origin fixed effects	Yes		Yes		Yes	
Market destination fixed effects	Yes		Yes		Yes	
R^2	0.589		0.301		0.407	
Hausman exogeneity test	**Estimation A versus Estimation B**: Hausman statistic = 943.87 Critical χ^2 (0.95, 2)= 5.99				**Estimation B versus Estimation C**: Hausman statistic = 97.52 Critical χ^2 (0.95, 5)= 11.07	

Notes: ** indicates statistical significance at the 1% level. Models are estimated with ticketing carrier dummies, market origin dummies and market destination dummies even though these dummy coefficients are not reported in the table.

It has been argued that passengers are more likely to choose itineraries offered by hub airlines for the following reasons: (1) flight schedules offered by hub airlines may be more convenient; and (2) it is more likely that passengers have frequent-flyer membership with an airline that has a hub at the passenger's origin airport.[22] As described in the data section, a hub product means that the origin airport on the itinerary is a hub for the airline that offers the product for sale. Consistent

[22] See Proussaloglou and Koppelman (1995), Berry (1990), Schumann (1986).

with our expectation, the coefficient on "Hub" is positive, suggesting that passengers are more likely to choose hub products, *ceteris paribus*.

As previously discussed in the data section, data on layover times and departure times, which are also measures of itinerary convenience, are not available in the DB1B database. Therefore, we cannot explicitly control for these aspects of itinerary convenience. However, it is reasonable to assume that the "Hub" dummy variable picks up some of these itinerary conveniences, which also explains the positive coefficient on this variable.

Ito and Lee (2007) argue that passengers that are members of an airline's frequent-flyer program may view the airline's virtual codeshare product as an inferior substitute to its pure online product since virtual tickets often do not allow the frequent-flyer to upgrade to first class even though the flights on the two itineraries (pure online and virtual) are the same. This argument leads us to expect the negative sign of the coefficient on the "Virtual" dummy variable in Table 5. In other words, the negative sign suggests that passengers perceive virtual codeshare products as inferior substitutes to pure online products.

The estimate of σ is statistically greater than zero, but its value is closer to zero than one. As such, there is statistical (but weak economic) evidence that passengers perceive the set of products offered by an airline as closer substitutes for each other compared to the substitutability of these products with products offered by other airlines [Gayle (2007b)]. In other words, passengers' choice behavior does have some element of airline brand-loyalty associated with it, even though this brand-loyalty does not seem to be very strong.

The estimate of λ_0 is negative and statistically different from zero, suggesting that demand for non-Delta/Continental/Northwest products is persistently/systematically lower across markets in which the three airlines eventually virtual codeshare together versus markets in which they compete but do not virtual codeshare together. In contrast, the estimate of λ_1 is positive and statistically significant, suggesting persistently higher demand for the three airlines' products across markets in which the three carriers eventually virtual codeshare together compared to markets in which they compete but do not virtual codeshare together. λ_2, λ_3 and λ_4 are not statistically different from zero, suggesting that: (1) there is no change in demand over the pre and post-alliance periods for non-Delta/Continental/Northwest products; and (2) in the case of markets where Delta, Continental and Northwest compete but did not virtual codeshare together, demand did not change for products offered by the three carriers over the pre and post-alliance periods.

Interestingly, we find that $\lambda_5 < 0$ and $\lambda_6 > 0$ at conventional levels of statistical significance. In addition, $\left|\frac{\lambda_5}{\lambda_6}\right| = 0.49$. Therefore, the sign pattern and actual values taken by λ_5 and λ_6 suggest that markets in which Delta, Continental and Northwest have a joint pre-alliance passenger share greater than 0.49 and eventually virtual codeshare together during the post-alliance period, experience an increase in demand for the three carriers products over the pre and post-alliance periods. In other words, there is evidence of a demand-increasing effect of virtual codesharing, but this demand-increasing effect is only evident in markets that the partner carriers have a substantial joint pre-alliance passenger share. Interestingly, these are the type of markets that you would expect a relatively larger share of consumers to hold frequent-flyer membership with at least one of the carriers prior to implementation of the alliance. Therefore, this structural demand estimation result provides strong support for the argument that a key source of the demand-increasing effect of codesharing is via the new opportunities that consumers have to accumulate and redeem frequent-flyer points across the partner carriers.

In an appendix available upon request we explore alternate and more detailed nesting structures for the demand model. We find that all qualitative results discussed above are robust to these alternate nesting structures.

Last, the demand model yields a mean own-price elasticity estimate of -1.52. Oum, Gillen and Noble (1986), and Brander and Zhang (1990) argue that 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 produces own-price elasticity estimates ranging from -3.2 to -3.6, while Berry and Jia (2010) find own-price elasticity estimates ranging from -1.89 to -2.10 in their 2006 sample. Therefore, the elasticity estimates generated from our model are reasonable and consistent with evidence in the existing literature.

6.2 Computed Product Markups and Marginal Costs

Table 6 reports summary statistics on price, computed product markups, and recovered marginal cost.[23] First, we see that during the post-alliance period mean price is lower in markets where the three partner carriers virtual codeshare together relative to markets in which they compete but

[23] Each reported sample mean in Table 6 has an associated sample standard error, and these associated sample standard errors are reported in parentheses. Once the sample mean is more than 2.58 times as large as the associated sample standard error, then we can conclude that the sample mean is statistically different from zero at the 1% level of significance. In Table 6 we use ** to indicate that the sample mean is statistically different from zero at the 1% level of significance.

do not codeshare together. Second, DL, CO and NW products have lower mean price relative to the mean price of products offered by other airlines, and this relatively lower mean price is more pronounced in markets that the three partner carriers virtual codeshare together.

Interestingly, product markups generated from the structural model reveal a different pattern than we see for price. In particular, even when assuming that the three partner carriers compete with each other in markets where the three airlines virtual codeshare together, the mean markup on their products ($157.15) is slightly higher (approximately 1 percent) relative to mean markup on products offered by other airlines ($155.19) in these markets. Therefore, the higher mean price of products offered by other airlines in these markets is likely due to cost factors, as evidenced by recovered marginal cost in the last column of the table.

Table 6
Price, Product Markups and Recovered Product Marginal Cost (in Dollars $)

	Price			**Product Markups**		**Marginal Cost**
	Non-codeshare markets during the post-alliance period	Codeshare markets during the post-alliance period	Assumption on price-setting behavior of airlines	Non-codeshare markets during the post-alliance period	Codeshare markets during the post-alliance period	Codeshare markets during the post-alliance period
	Mean (Std. error)	Mean (Std. error)		Mean (Std. error)	Mean (Std. error)	Mean (Std. error)
All products	225.30** (1.342)	213.37** (0.686)	All compete	155.64** (0.155)	156.03** (0.158)	57.34** (0.681)
Products not offered by DL, CO or NW	226.09** (1.857)	215.87** (0.956)	All compete	155.84** (0.183)	155.19** (0.087)	60.68** (0.935)
DL, CO and NW products	224.15** (1.882)	210.06** (0.965)	All compete	155.34** (0.271)	157.15** (0.347)	52.91** (0.982)
	-	-	DL, CO and NW collude in their codeshare markets	-	165.91** (1.043)	44.15** (1.403)

Notes: ** indicates statistical significance at the 1% level. Marginal cost is recovered from each supply model as follows: *Marginal Cost* = $\mathbf{p} - mkup$ and *Marginal Cost*^{Collude} = $\mathbf{p} - mkup^{Collude}$, where $\mathbf{p}$ is the vector of observed product prices.

If we assume that DL, CO and NW collude in markets that they virtual codeshare together, the comparative patterns on mean markups and mean marginal cost described above are more pronounced in these markets. In particular, mean markup on products offered by the three partner carriers ($165.91) is substantially higher than mean markup on products offered by other carriers ($155.19), and mean marginal cost across products offered by the three carriers ($44.15) is substantially lower than mean marginal cost across products offered by other carriers ($60.68). The table also shows that assuming the three partner carriers collude, instead of compete, in markets that they virtual codeshare together will result in higher mean markups on their products in these markets ($157.15 versus $165.91). This difference in assumed price-setting behavior yields a substantial difference in markup on their products both in terms of dollars and percent increase ($8.76 and 5.10% respectively), and the differences are statistically significant at the 1% level.

Note we have not yet resolved which price-setting behavior between the partner carriers in their codeshare markets is better supported by the data. All we have done so far is to summarize what markups and marginal cost levels are under each assumed price-setting behavior. To investigate which assumed price-setting behavior is more appropriate, we subsequently turn to a formal non-nested statistical test for model selection. But first we show estimation results of supply equations under each assumed price-setting behavior.

6.3 Results from Supply Equation Estimation

Note that the markets in which price-setting behavior is in question are markets in which DL, CO and NW virtual codeshare together during the post-alliance period. This is because the policy-relevant issue is whether virtual codesharing together facilitates collusive price-setting behavior between the partner carriers. As such, the remainder of the analysis focuses on this subsample of markets. Therefore, the supply equations are estimated on this subsample of markets.

Table 7 reports parameter estimates for supply *Model h* and *Model g* respectively. Recall that *Model h* assumes Delta, Continental and Northwest do not practice collusive pricing in markets that they virtual codeshare together during the post-alliance period, while *Model g* assumes that they practice collusive pricing in these markets during the post-alliance period. The markup variables, $mkup$ and $mkup^{Collude}$, capture this assumed difference in pricing behavior, and are the only variables that differ across *Model h* and *Model g*. Note that coefficients on these markup variables are not estimated, but instead set equal to 1 to be consistent with theoretical derivations of the supply equations in the model section. This coefficient restriction on the markup variables effectively implies that $p_j - mkup_j$ and $p_j - mkup_j^{Collude}$, which are recovered product marginal costs, are the dependent variables in *Model h* and *Model g* respectively. As such, the coefficients that are estimated in the supply equations are marginal cost parameters.

The coefficient estimate on "Itinerary Distance" in each supply model is positive and statistically significant. This is evidence that marginal cost is increasing in itinerary distance, as we expect.

As previously described in the data section, both the "Carrier Presence at Origin" and "Carrier Presence at Destination" variables are airline-specific and vary across markets for each airline. "Carrier Presence at Origin" measures the number of different cities that an airline has non-stop flights from going into the origin city of the market, while "Carrier Presence at Destination" measures the number of different cities that the airline serves using non-stop flights from the destination city of the market. These variables should be correlated with the volume of passengers an airline channels through a market even though the endpoint cities of the market may not be the origin or final destination for many of the passengers. As such, we use these variables to indirectly capture the presence of economies of passenger-traffic density. Economies of passenger-traffic density means that an airline's marginal cost of transporting a passenger in a market falls as the volume of passengers that the airline transports in the market increases [Brueckner and Spiller (1994)].

An anonymous referee correctly points out that since the coefficients on the carrier-presence variables measure an airport-level effect on a carrier's prices in individual markets, perhaps these coefficients capture some blend of economies of density along with the cost of running a hub. As such, when drawing economic inferences from the sign of these coefficient estimates, it is advisable to remember that economies of density might not be the only factor that influences these coefficient estimates.

The coefficient estimate on "Carrier Presence at Origin" is positive and statistically significant for both supply models, but the coefficient estimate on "(Carrier Presence at Origin)^2" is not statistically significant in *Model h*, and is positive with weak statistical significance in *Model g*. Therefore, results for the "Carrier Presence at Origin" variable are not consistent with the presence of economies of passenger-traffic density. However, the sign pattern of coefficient estimates on "Carrier Presence at Destination" and "(Carrier Presence at Destination)^2" does suggest the presence of economies of passenger-traffic density once the airline's "presence" measure at the destination city is sufficiently large. Economies of passenger-traffic density has stronger statistical support in *Model h* compared to *Model g* since the negative coefficient estimate on "(Carrier

Presence at Destination)^2" is statistically significant in *Model h*, but not statistically significant in *Model g*. The coefficient estimates on "Carrier Presence at Destination" and "(Carrier Presence at Destination)^2" in *Model h* suggest that an airline has to provide nonstop flight to at least 133 different cities ($\frac{0.008}{2\times(0.00003)}$) from the destination city in order to achieve economies of passenger-traffic density in the relevant market. This "presence" threshold is not widely attained in the sample given that the maximum value for the "Carrier Presence at Destination" variable is 143, with a mean of 26. However, since economies of passenger-traffic density might not be the only factor driving these coefficient estimates, then these coefficient estimates might not yield a precise "presence" threshold estimate for achieving economies of passenger-traffic density.

Table 7
Supply Equation Parameter Estimates

	Model h: No collusion between DL, CO and NW in post-alliance period.	*Model g*: Collusion between DL, CO and NW in their codeshare markets during the post-alliance period.
Variable	**(1)**	**(2)**
Constant	-0.596** (0.122)	-0.559** (0.139)
Itinerary Distance (in 1,000 miles)	0.227** (0.024)	0.296** (0.035)
Carrier Presence at Origin	0.005** (0.001)	0.003** (0.001)
(Carrier Presence at Origin)^2	2.44E-07 (0.00001)	0.000018[a] (0.000011)
Carrier Presence at Destination	0.008** (0.001)	0.006** (0.002)
(Carrier Presence at Destination)^2	-0.00003** (8.93E-06)	-0.000019 (0.000016)
mkup	1	-
$mkup^{Collude}$	-	1
Carrier fixed effects	Yes	Yes
Market origin fixed effects	Yes	Yes
Market destination fixed effects	Yes	Yes
R^2	0.373	0.295
Number of observations: 8,165		

Notes: ** indicates statistical significance at the 1% level. [a] indicates statistical significance at the 10% level. Equations are estimated using ordinary least squares. Standard errors are in parentheses. The coefficients on the markup variables, *mkup* and $mkup^{Collude}$, are not estimated but set equal to 1 based on theoretical derivations of the supply equations in the model section. Models are estimated with operating carrier dummies, market origin dummies and market destination dummies even though these dummy coefficients are not reported in the table

6.4 Statistical Non-nested Tests for Model Selection

To determine which of the two alternate supply model specifications provides the best statistical fit of the data, we rely on a likelihood-based non-nested statistical test in Vuong (1989).[24] The

[24] Similar to how we use a likelihood-based non-nested test for supply model selection, Villas-Boas (2007) and Gayle (2013) use a generalized methods of moments-based non-nested statistical test for supply model selection.

non-nested statistical test is a modification of the well-known likelihood ratio test. The likelihood ratio statistic for comparing *Model h* and *Model g* is given by:

$$LR = \sum_{j=1}^{n} \left(LL_j^h - LL_j^g \right), \tag{17}$$

where j index observations in the data, and n is the sample size. LL_j^h is the optimal value of the log likelihood function for *Model h* evaluated at observation j. Specifically, assuming that the error term of the supply equation is normally distributed, $LL_j^h = \log \left[\phi \left(\frac{p_j - mkup_j - W_j \widehat{\gamma}_h}{\widehat{\sigma}_h} \right) \right]$, where $\phi(\cdot)$ is the standard normal probability density function, $\widehat{\gamma}_h$ is the vector of marginal cost parameter estimates for *Model h* that we report in Table 7, and $\widehat{\sigma}_h$ is an estimate of the standard deviation of the residuals from *Model h*.[25] LL_j^g is computed analogously, i.e. $LL_j^g = \log \left[\phi \left(\frac{p_j - mkup_j^{Collude} - W_j \widehat{\gamma}_g}{\widehat{\sigma}_g} \right) \right]$, where $\widehat{\gamma}_g$ is the vector of marginal cost parameter estimates for *Model g* that we report in Table 7, and $\widehat{\sigma}_g$ is an estimate of the standard deviation of the residuals from *Model g*.

Vuong (1989) shows that the likelihood ratio statistic in (17) can be normalized by its variance:

$$v^2 = \frac{1}{n} \sum_{j=1}^{n} \left(LL_j^h - LL_j^g \right)^2 - \left[\frac{1}{n} \sum_{j=1}^{n} \left(LL_j^h - LL_j^g \right) \right]^2. \tag{18}$$

Furthermore, the resulting non-nested test statistic:

$$Q = n^{-0.5} \frac{LR}{v}, \tag{19}$$

is asymptotically distributed standard normal under the null hypothesis that the two models being compared by the test are asymptotically equivalent.[26] As such, for this one-tale test at a 5% level of significance, $Q > 1.64$ implies that supply model g is statistically rejected in favor of supply model h, $Q < -1.64$ implies that supply model h is statistically rejected in favor of supply model g, while $-1.64 < Q < 1.64$ implies that we cannot statistically distinguish between the two models being compared.

For the estimated supply models in Table 7, we find that $Q = 5.27$, suggesting that model g is statistically rejected in favor of supply model h. In other words, the supply model that assumes the three carriers do not collude (*Model h*) in their codeshare markets during the post-alliance period is statistically superior to the supply model that assumes the three airlines collude (*Model*

[25] Note that supply *Model h* and *Model g* are linear regression models. In the case of a linear regression model, least squares parameter estimates and maximum likelihood parameter estimates are equivalent.

[26] Equations (17), (18) and (19) above correspond to equations (3.1), (4.2) and (5.6) on pages 312, 314 and 318 respectively in Vuong (1989).

g) in these markets. To the best of our knowledge, this is the first paper to explicitly test and statistically reject that collusive pricing behavior is associated with a codeshare alliance.

Unlike international alliance partners that often receive antitrust immunity, i.e. antitrust authorities have granted some international partners the right to explicitly collude, domestic alliance partnerships have not been granted such rights [see Brueckner and Proost (2010); Brueckner, Lee and Singer (2011); and Brueckner (2003)]. However, even though domestic alliance partners are forbidden to explicitly collude, it is reasonable to suspect, as policymakers did in the case of the DL/CO/NW alliance, that the cooperation between domestic partners required to make their interline service seamless, could facilitate illegal tacit collusion. So prior to the formal analysis in this paper, tacit collusion between domestic partners could not be ruled out.

7 Conclusion

The main contribution of our present paper is to specify and estimate a structural econometric model that allows us to disentangle demand changes from possible changes in airline pricing behavior that are associated with a codeshare alliance. We focus on the Delta/Continental/Northwest codeshare alliance, which was formed in June 2003. This alliance is particularly interesting to study because, before ultimately allowing the alliance to go forward, the U.S. Department of Transportation expressed concern that the alliance could facilitate collusion on prices and/or service levels in the partners' overlapping markets. In addition, previous reduced-form econometric analysis of this alliance found evidence that virtual codesharing between Delta, Continental and Northwest is associated with higher price [see Gayle (2008)]. Therefore, our analysis focuses on better understanding the market effects, both from the demand and supply sides of the market, of virtual codesharing between the three airlines in their overlapping markets.

Our key findings are as follows: First, the econometric estimates for the air travel demand equation suggest that the Delta/Continental/Northwest alliance has a demand-increasing effect associated with it. Importantly, the demand-increasing effect is only evident in markets that the partners have a substantial joint passenger share (greater than 49%) prior to implementation of the alliance. Since a relatively larger proportion of passengers in a market are more likely to have frequent-flyer membership with at least one of the three carriers in markets that the carriers jointly dominate prior to the alliance, this finding is consistent with the argument that these frequent-flyer passengers will increase their demand for the alliance partners' products given that the alliance creates new opportunities for passengers to accumulate and redeem frequent-flyer points across partner carriers.

Second, a statistical non-nested test applied to air travel supply model selection suggests that Bertrand Nash pricing behavior, rather than collusive pricing behavior, between the three airlines better fit the data in markets where the three airlines codeshare together. To the best of our knowledge, this is the first paper to explicitly test and statistically reject that collusive pricing behavior is associated with a codeshare alliance.

In summary, if increased collusive pricing behavior of the partner carriers is the primary concern of policymakers with allowing the Delta/Continental/Northwest alliance to go forward, then the evidence does not suggest implementation of the alliance facilitated collusive pricing.

References

[1] Aguirregabiria, V., 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.

[2] Armantier, O., and O. Richard. 2008. "Domestic Airline Alliances and Consumer Welfare," *RAND Journal of Economics,* Vol. 39, No. 3, 875-904.

[3] 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, 195-222.

[4] Berry, S. 1990. "Airport Presence as Product Differentiation." *The American Economic Review*, Vol. 80, 394-399.

[5] Berry, S. 1992. "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, Vol. 60 (4), 889-918.

[6] Berry, S. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics*, Vol. 25, 242-262.

[7] Berry, S., M. Carnall, and P. Spiller. 2006. "Airline Hubbing: Costs and Demand." In Darin Lee, ed., *Advances in Airline Economics,* Vol. 1, Elsevier Press.

[8] 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.

[9] Bilotkach, V. 2007. "Price Effects of Airline Consolidation: Evidence from a Sample of Transatlantic Markets," *Empirical Economics*, Vol. 33, 427 – 448.

[10] Borenstein, S. 1990. "Airline Mergers, Airport Dominance, and Market Power," *The American Economic Review Papers and Proceedings*, Vol. 80, No. 2, pp. 400-404.

[11] Brander, James A., and Anming Zhang. 1990. "Market Conduct in the Airline Industry: An Empirical Investigation," *RAND Journal of Economics*, Vol. 21, 567-583.

[12] Bresnahan, T., S. Stern, and M. Trajtenberg. 1997. "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *RAND Journal of Economics,* Vol. 28, 17-44.

[13] Brueckner, J., D. Lee, and E. Singer. 2011. "Alliances, Codesharing, Antitrust Immunity and International Airfares: Do Previous Patterns Persist?", *Journal of Competition Law & Economics*, May 2011.

[14] Brueckner, J., and S. Proost. 2010. "Carve-Outs Under Airline Antitrust Immunity", *International Journal of Industrial Organization*, 28(6), 657-668.

[15] Brueckner, J. 2003. "International Airfares in the Age of Alliances," *Review of Economics and Statistics,* Vol. 85, 105-118.

[16] Brueckner, J., and W.T. Whalen. 2000. "The Price Effects of International Airline Alliances," *Journal of Law and Economics,* Vol. XLIII, 503-545.

[17] Brueckner, J., and P. Spiller. 1994. "Economies of Traffic Density in the Deregulated Airline Industry," *Journal of Law and Economics,* Vol. 37, 379-415.

[18] Chen, Y., and P. Gayle. 2007. "Vertical Contracting Between Airlines: An Equilibrium Analysis of Codeshare Alliances," *International Journal of Industrial Organization*, Vol. 25, Issue 5, pp. 1046-1060.

[19] Gayle, P. and C-Y. Wu. 2012. "Are Air Travel Markets Segmented Along the Lines of Nonstop versus Intermediate-stop(s) Products?" *Manuscript, Kansas State University.*

[20] Gayle, P. 2013. "On the Efficiency of Codeshare Contracts Between Airlines: Is Double Marginalization Eliminated?" *American Economic Journal: Microeconomics*, Vol. 5, Issue 4, pp. 244-273.

[21] Gayle, P. 2008. "An Empirical Analysis of the Competitive Effects of the Delta/Continental/Northwest Codeshare Alliance," *Journal of Law and Economics*, Vol. 51, November, pp. 743-766.

[22] Gayle, P. 2007a. "Airline Code-share Alliances and their Competitive Effects," *Journal of Law and Economics,* Vol 50, November, pp. 781-819.

[23] Gayle, P. 2007b. "Is Virtual Codesharing A Market Segmenting Mechanism Employed by Airlines?" *Economics Letters,* Vol. 95, No. 1, pp. 17-24

[24] Goldberg, P., and F. Verboven. 2001. "The Evolution of Price Dispersion in the European Car Market," *The Review of Economic Studies,* Vol. 68, No. 4, 811-848.

[25] Ito, H. and D. Lee. 2007. "Domestic Codesharing, Alliances and Airfares in the U.S. Airline Industry," *Journal of Law and Economics*, Vol. 50, No. 2, pp. 355-380.

[26] Kim, E. H. and V. Singal. 1993. "Mergers and Market Power: Evidence from the Airline Industry," *The American Economic Review*, Vol. 83, No. 3, pp. 549-569.

[27] Lederman, M. 2007. "Do enhancements to loyalty programs affect demand? The impact of international frequent flyer partnerships on domestic airline demand," *RAND Journal Economics*, Vol 38, No. 4, pp. 1134-1158.

[28] McFadden, D. 1978. "Modelling the Choice of Residential Location," in A. Karlqvist et al. (eds.), *Spatial Interaction Theory and Planning Models* (Amsterdam: North-Holland).

[29] Nevo, A. 2000. "Mergers with Differentiated Products: The case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics*, Vol. 31, pp. 395-421.

[30] 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.

[31] 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.

[32] Proussaloglou, K., and F. Koppelman. 1995. "Air Carrier Demand: An Analysis of market Share Determinants." *Transportation*, Vol. 22, 371-388.

[33] Schumann, H. 1986. "Oligopolistic Nonlinear Pricing - A Study of Trading Stamps and Airline Frequent Flyer Programs." *Unpublished Ph.D. Dissertation, Northwestern University.*

[34] U.S. Department of Transportation, Office of the Secretary. 2003. "Termination of review under 49U.S.C. § 41720 of Delta/Northwest/Continental Agreements."

[35] Villas-Boas, Sofia Berto. 2007. "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data," *Review of Economic Studies*, Vol. 74, No. 2, pp. 625 - 652.

[36] Vuong, Quang. 1989. "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses," *Econometrica*, Vol. 57, No. 2, pp. 307 - 333.