IMPACTS OF COMPETITION ON CONNECTING TRAVELERS: EVIDENCE FROM THE TRANSATLANTIC AVIATION MARKET

Valeria Bernardo
Escola Superior de Ciències Socials i de l'Empresa, TecnoCampus, Universitat Pompeu Fabra, C/ Ernest Lluch 32, 08302, Mataró, Barcelona. Mail to: vbernardo@tecnocampus.cat

Xavier Fageda*
Department of Applied Economics & GiM-IREA, University of Barcelona. Address: Diagonal Av. 690, 08034, Barcelona, Spain. Mail to: xfageda@ub.edu

Abstract: This paper examines the effects of competition driven by low-cost airlines and high-speed rail on long-haul connecting routings. We use passenger, supply and fare data for 2010-2017 on transatlantic routings with a stopover in large European hub airports. Competition from low-cost airlines on the short segments of transatlantic routings reduces the number of connecting passengers channeled by hubbing airlines, who have to pay higher fares. In contrast, airline competition in the long segments of the transatlantic routings does not seem to have a detrimental effect on connecting travelers. Finally, high-speed rail services are not found to lead to a reduction in the number of connecting passengers or to an increase in the price they have to pay.

Keywords: network airlines, low-cost airlines, high-speed rail, competition, connecting routings, passengers, fares.

*Corresponding author.

1. Introduction

Network airlines operate through hub-and-spoke systems to take advantage of the economies of traffic density that characterize the airline industry. To this end, they concentrate flights at their (few) hub airports where passengers on short-haul flights feed their long-haul services. This means a sizable proportion of passengers channeled by network airlines are connecting passengers that use the hub airport as a stopover to reach their final long-haul destination. In contrast, low-cost carriers (LCCs) operate point-to-point routes with the consequence that most of their passengers are point-to-point passengers who do not need a stopover to reach their final destination. LCCs have been able to exploit several cost advantages that make them highly competitive on short-haul and medium-haul trips.

The hubbing operations of network airlines generate a number of positive externalities including the high direct connectivity afforded by hub airports and a high indirect connectivity afforded by smaller airports with good connections to these hub airports. Elsewhere, travelers on short-haul flights typically benefit from the low prices offered by LCCs.

In Europe, the competition between hubbing airlines and LCCs is intensifying as the presence of the latter increases at large hubs (Dobruszkes et al., 2017; Wong et al., 2019). Here, competition on the short segments of intercontinental routings may have detrimental effects both for hubbing operations and for connecting travelers due to a poorer exploitation of density economies. However, competition on these short segments might have its benefits for connecting travelers due to the weaker market power exercised by the hubbing airline. Determining whether the presence of LCCs is beneficial or not for connecting passengers is relevant to the extent that LCCs utilize the capacity of what are highly congested airports.

Additionally, some hubbing airlines in Europe also face competition from high-speed rail (HSR) lines. This is especially true of countries with a large domestic air market such as France, Spain and Italy. On short- and medium-haul routes, HSR usually offers a highly competitive service in terms of fares, times, frequency and comfort. However, an increase in the number of passengers travelling on the long segments (yet arriving at the hub airport by train) could offset this negative competition effect for hubbing airlines on the short segments. Shifting short-haul services from planes to trains could reduce the external costs associated with aviation in terms of pollution and congestion.

In this paper, we examine the effects of competition driven by LCCs and HSR on the hubbing operations of European network airlines at their main hubs. Specifically, we analyze the effects of competition on passengers channeled and on fares charged by hubbing airlines on transatlantic routings with a stopover in their hubs.

We focus our analysis on six large European hubs that concentrate a high proportion of connecting flights from Europe to North America: namely, Amsterdam, Paris-Charles de Gaulle, Rome-Fiumicino, Frankfurt, London-Heathrow and Madrid. We use passenger, fare and supply data on transatlantic routings with a stopover in one of these hubs for 2010-2017.

All previous studies about the impact of LCCs on fares and traffic focus on short-haul non-stop routes (Windle and Dresner, 1995, 1999; Dresner et al., 1996; Morrison, 2001; Hofer et al., 2008, Goolsbee and Syverson; 2008; Oliveira and Huse, 2009; Murakami; 2011; Huse and Oliveira, 2012, among others). Here, we add to this literature by providing a direct test of the impact of LCCs on traffic and on fares in connecting long-haul flights.

Additionally, several studies have examined the impact of HSR on air traffic and fares (Clewlow, 2012; Albalate et al., 2015; Bergantino et al., 2015; Wei et al., 2017; Zhang et al., 2017; Wang et al., 2018; Ma et al., 2019; Wang et al., 2018; Zhang et al., 2018; Ma et al., 2019, among others). All these studies focus on the impact of HSR on non-stop routes, as does the literature examining the impact of LCCs. Thus, here, we add to this literature by providing a direct test of the impact of HSR on connecting travelers on long-haul flights.

The rest of this paper is organized as follows. The next section briefly reviews the related empirical literature relevant to our study and develop the main hypothesis of the empirical analysis. In the third section, we detail the sample used and provide relevant information about the data used in the empirical analysis. In the following section, we explain the empirical equation that we estimate. In the fifth section, we discuss the main results of the empirical analysis. After that, we run several robustness checks to examine different issues that may have an influence on the estimated results. The last section is devoted to our concluding remarks.

2. Literature review and hypothesis

Figure 1 illustrates a simplified version of the network that we consider in the empirical analysis. We focus on routings with one stopover. These routings have a short leg (ie; O1H) and a long leg (HD). The network in figure 1 is a hub-and-spoke network where the origin points (O1….O4) are connected to the destination (D) with a stopover in the hub airport (H). Hence, (network) airlines may exploit density economies by using bigger planes at higher load factors because their flights are filled with point-to-point-passengers (passengers that fly from several O points to H, or that fly from H to D) and connecting passengers (passengers that fly from several O points to D with a stopover in H). In this regard, there is consensus in the literature about the relevance of density economies in air transportation (Caves et al., 1984; Brueckner and Spiller, 1994; Berry et al., 2006).

Insert Figure 1

Previous empirical studies provide evidence about the impact of LCCs and HSR on air traffic and fares in short-haul non-stop routes. In our network, these routes would be those that link any of the O points with H. To this point, our contribution is to examine how competition in OH points may have an influence on the connecting routings (OHD).

The downward pricing pressure that LCCs exert on the routes they operate is well documented (eg; Dresner et al., 1996; Windle and Dresner, 1995, 1999). In particular, those studies that analyze the reaction of incumbent (network) airlines to competition from LCCs generally find that they reduce prices when LCCs enter in the route (Hofer, et al. 2018; Morrison, 2001; Murakami, 2011; Goolsbee and Syverson, 2008; Oliveira and Huse, 2009; Huse and Oliveira, 2012). Note also that some recent studies have examined the viability of long-haul low-cost operations in terms of both costs and revenues (De Poret et al., 2015; Soyk et al., 2017; Soyk et al., 2018).

Other studies have focused on the capacity reactions of incumbent airlines to LCC competition. In particular, Fageda (2014) estimates a frequency equation to explain the determinants of network airline services in large European hub airports. He finds that network airlines reduce their flight frequencies when the share of low-cost airlines increases both on the route and at the hub airport. In contrast, some studies for Brazil find that incumbent airlines increase capacity to compete with LCCs as a pre-emption strategy (Bettini et al., 2018; Bettini and Oliveira, 2008).

Furthermore, several studies have examined the effect of HSR services on air traffic and airfares (see Givoni and Dobruszkes, 2013; and Zhang et al., 2019, for a detailed literature review). As with analyses of LCC impact, the focus is on non-stop short-haul routes and usually the studies concentrate on domestic routes. However, various studies provide some insights about the expected impact for connecting routes.

It has been found that competition from HSR affects negatively airline fares although the impact is greater in the case of LCCs (Bergantino et al. 2015; Wei et al., 2017; Wang et al., 2018; Ma et al., 2019). Furthermore, it has been found that HSR services impose a substitution effect on short-haul (less than 500 km) and medium-haul routes (between 500 and 800-1000 km), causing a decrease in air traffic after HSR entry (Albalate et al., 2015; Wan et al., 2016; Zhang et al., 2018). However, for long-haul trips (more than 800-1000 kilometers), the opposite outcome is reported for analyses of the Chinese experience (Zhang et al., 2018; Wan et al., 2016; Liu et al., 2019).

Additionally, at the airport level, it is found a lower air connectivity in the domestic market after HSR entry (Zhang et al., 2017; Li et al., 2019), but this effect is reported as not being significant at the international level (Zhang et al., 2017). Furthermore, the effect seems to differ between hubs and non-hubs (Albalate et al., 2015; Clewlow et al., 2012; Dobruszkes et al., 2014) and, in fact, some studies find an increase in air traffic when the airport has its own HSR station (Clewlow et al., 2012, Zhang et al. 2018).

Taken as a whole, this body of literature provides clear evidence that LCCs and HSR can have substantial negative impacts on airline traffic and fares on short-haul non-stop routes. However, there is some indirect evidence that HSR may play a complementary role in longer international services.

In the context of our network, we may expect that competition in the short leg of the routing have an influence on connecting travelers that make a stopover in the hub airport to reach their long-haul final destination.

Services provided in short-haul routes by LCCs and HSR may be very competitive in terms of fares and quality. Thus, both LCCs and HSR may capture much of the point-to-point traffic that could otherwise be channeled by the hubbing airline. Such reduction in the number of point-to-point passengers may result in a poorer exploitation of density economies in the short-haul flight. An expected reaction of the hubbing airline to less point-to-point traffic would be to reduce its supply either by cutting flight frequencies and/or operating smaller planes. Lower frequencies may be detrimental to the connections in terms of layover time and smaller planes may lead to higher operating costs. Finally, previous studies clearly show that network airlines reduce prices when they have to compete with LCCs or HSR. Network airlines could compensate the lower profitability in the short-segment by increasing the fares charged to travelers in the long-segment.

All these factors may imply that competition on short segments leads to fewer connecting passengers being channeled by the hubbing airline and that the surviving connecting passengers have to pay higher fares. This is what we call the "density economies hypothesis".

However, the intensity of competition may have an effect on the ability of the hubbing airline to exploit its market power. Indeed, competition could force the hubbing airline to contain its prices, which in turn could boost demand. This is what we call the "market power hypothesis".

Finally, the negative effects linked to the "density economies hypothesis" could be countered if HSR or LCCs services may feed the long flight with connecting passengers that take the train or the LCC flight in the short segment of the routing. Hence, the hubbing airline might reduce supply on the short segments without sacrificing passengers on the long segments. In this regard, higher frequencies and/or lower fares on the long segments could lead to more connecting passengers by the hubbing airline. This is what we call the "complementary effect" hypothesis.

We argue that this latter hypothesis should be more relevant for HSR than for LCCs, particularly for those airports that have their own high-speed train station. HSR generally provide high-frequency services that depart from the city-center. Furthermore, HSR may alleviate the severe congestion distortions that suffer most of hub airports. Finally, LCCs may affect to a much higher number of feeding short-haul routes to the hub than HSR so that LCC may have a higher negative influence on the overall profitability of the feeder services provided by hubbing airlines.

3. Data

Passenger and fare data are drawn from the Official Airlines Guide (OAG) and include information for one-way routings with one-stop. Specifically, the routings we consider link a European (i.e. European Union, Norway and Switzerland) airport with a North American (i.e. United States and Canada) airport via one large hub in Europe.[1] Hence, the origin is from a European airport and the destination is to a US/Canada airport. All routings comprise a long-haul transatlantic segment and a short-haul segment within Europe.

The hub airports included in our analysis are Amsterdam, Paris-Charles de Gaulle, Rome-Fiumicino, Frankfurt, London-Heathrow and Madrid. These airports are characterized by their big size and the fact that one network airline operates a high percentage of all its flights out of them. Additionally, these six airports concentrate a high proportion of connecting flights from Europe to North America.

Data are at the airline/routing level and our focus is on the traffic channeled and on the fares charged by the dominant hubbing airline on those connecting routings where the stopover is made at their main hub (i.e. IAG at London Heathrow and Madrid, Air France-

[1]Note that all the countries in our sample form part of an open skies agreement in all the years of the period considered (Bernardo and Fageda, 2017).

KLM at Amsterdam and Paris, Lufthansa at Frankfurt and Alitalia at Rome-Fiumicino). Examples of the routings in our sample, include, that flown by KLM between Athens and Atlanta via Amsterdam and that flown by British Airways between Barcelona and Boston via London Heathrow. The long segments in these examples correspond to Amsterdam-Atlanta and London Heathrow-Boston and the short segments to Athens-Amsterdam and Barcelona-London Heathrow, respectively.

We use quarterly data for the period 2010 (first quarter)–2017 (third quarter) comprising 131,474 observations in the case of passenger data and 113,829 observations in that of fare data. We restrict our attention to: *i)* routings with more than 20 passengers per quarter to guarantee a feasible dataset, and *ii)* urban areas with a population exceeding 300,000 inhabitants to ensure consistency of information for both North America and Europe. In this regard, our sample is composed by 8088 routing for several time periods.

Data for the number of passengers and average fares comes from OAG Traffic analyzer. Passenger data are based on marketing information data tapes (MIDT) that include passengers' true origin and destination. This dataset is constructed from bookings made through global distribution systems. Hence, direct airline bookings are excluded. Unfortunately, we do not have available information on the share of direct bookings although it can be expected that is lower in long-haul transatlantic routings served by network airlines than in non-stop routings served by low-cost airlines.

The OAG tool offers several reports with different information for fares. Estimated data by fare class is provided in some reports, while the one used here – mix report – is based on average fare values that are derived from Travelport issued tickets. Travelport is one of the top three global distribution systems along with Amadeus and Sabre. Given that fare data provided by the OAG Traffic Analyzer mix report tool is expressed in mean values, we cannot identify discrimination across fare classes. Some recent studies that use OAG data or similar data from other global distribution systems include Abate and Christidis (2020), Boonekamp et al. (2018), Button et al. (2019), Fageda et al. (2019b), Scotti and Volta (2018) or Nyoyaa et al. (2018).

The fare data used have some limitations because it provides a partial picture of the actual average fares paid by all travelers in each routing. However, it may have some advantages in relation to other sources used in previous papers that examine airline fares at the route level. Screen scraping techniques using electronic spiders linked to the websites of specific airlines are based on posted rather than booked data. Furthermore, they could only cover a very short period of time. The use of specific surveys would require to limit the analysis to a small number of airports and routes. For the Transatlantic aviation market, an alternative would be to use the airline origin and destination survey (DB1B) provided by the US Department of Transportation but international data are not available for non-US citizens and information only covers 10% of all tickets sold.

While the different ways to collect fare data may be helpful for addressing some specific questions, the OAG fare data is the best available source for the purposes of this paper given that it allows considering a large number of routings for several years. However, it must be recognized that the number of tickets used to calculate the average fares may be low in some routings so that the variable of fares may be affected by outliers or extreme values. In section 6, we analyze in detail the potential distortion that this limitation in the fare data may impose on results of the baseline regressions.

To show the relevance of connecting routings, figure 2 shows the proportion of connecting passengers over total passengers in the transatlantic market for the last quarter of the period considered (third quarter of 2017). In all hub airports in our sample, connecting passengers are more than half of the total market with percentages that range from 54% in London Heathrow to 77% in Frankfurt.

Insert Figure 2

We also use supply data at the airline/route level (flight frequency, distance) which are drawn from RDC aviation (Innovata data). Note that these data are for non-stop routes. We use these data to build variables that capture the effect of the intensity of airline competition on the short- and long-haul segments of the transatlantic routings.

As control variables, we consider the population of the urban areas at the origin and destination of the routes, and gross domestic product (GDP) per capita at both endpoints of the routes at the country level. Urban population data have been obtained from the United Nations (World Urbanization Prospects) and GDP per capita from the World Bank (World Development Indicators).

As main variables of the analysis, we consider a dummy variable that takes a value of one for those routes on which HSR services compete with airline flights (referring, logically, to the short segments of the transatlantic routings). The HSR variable applies to routes on which rail services operate at, at least, 250 km/h. We only take into account direct services with no connections/transfers and for which the entire line is a high-speed service (thus, routes with sections of conventional rail are excluded). We obtained information about each line from the International Union of Railways' (UIC) HSR maps, while information about direct services across Europe was collected from the search engines of *Voyages-sncf*, the commercial online ticket distributor of SNCF.

In our sample, the following routes offer HSR services (in parentheses, we report the year the service came into operation when later than the first quarter of 2010):

1) Paris to Amsterdam, Barcelona (first quarter of 2013), Bordeaux (third quarter of 2017), Cologne, Basel-Mulhouse (first quarter of 2012), Frankfurt, London, Luxembourg, Lyon, Marseille, Montpellier, Nantes, Rennes (third quarter of 2017) and Toulon.

2) Madrid to Alicante (fourth quarter of 2010), Barcelona, Málaga, Seville, Valencia (fourth quarter of 2010) and Zaragoza.

3) Rome to Bologna, Florence, Milan, Naples and Turin

4) Amsterdam to Brussels, Cologne, Frankfurt and Paris.

5) Frankfurt to Amsterdam, Brussels and Paris

6) London to Paris

Chapter 5.1 of the Manual on the Regulation of International Air Transport published by the International Civil Aviation Organization (ICAO) defines an LCC as "an air carrier that has a relatively low-cost structure in comparison with other comparable carriers and offers low fares and rates. Such an airline may be independent, the division or subsidiary of a major network airline or, in some instances, the ex-charter arm of an airline group". Based on these criteria, the ICAO provides the list of LCCs that we use here. The LCCs offering services on the short-haul segments of our sample are: Air Southwest, Belle Air, Blue Air, Blue1, bmibaby, Condor, Corendon, Easyjet, Eurowings, Flybe, Germania, Germanwings, Jet 2, Meridiana, Monarch, Niki, Norwegian, Pegasus, Ryanair, Smarwings, SunExpress, Thomas Cook, Thomson, Transavia, Tuifly, Volotea, Vueling, Wind Jet, Wizz air, Wow and XL Airways.

The presence of LCCs is modest in Frankfurt and London Heathrow, but their share has increased exponentially in the considered period to about 23% in Paris and more than 30% in Amsterdam, Rome and Madrid. Note also that LCCs may, in fact, offer short-haul flights to secondary airports that serve the same urban area as that in which the hub airport is located. Specifically, several LCCs offer flights to secondary airports in the urban area of London (Gatwick, Stansted, Luton, City and Southend), Paris (Beauvais-Tille, Orly) and Rome (Ciampino).

On the long-haul segments, the presence of LCCs is much lower in the period considered. The proportion of observations in our sample for which LCCs offer high frequency services on short segments (at least one daily flight) is about 14% while the proportion of observations for which LCCs offer high frequency services on long segments is about 0.005%. Thus, it is clear that here competition from LCCs concentrates above all on short segments.

Competition on long segments (ie; Amsterdam-Atlanta, London Heathrow-Boston) comes primarily from US and Canadian network carriers. Note here that British Airways, Iberia and American airlines signed a Joint Venture agreement in October 2010 (later joined by US Airways in 2014), Lufthansa signed a Joint Venture agreement with Air Canada, United and Continental in October 2009, and Air France-KLM signed a Joint Venture agreement with Delta/Northwest in June 2009 (later joined by Alitalia in July 2010). As part of these Joint Venture agreements, airlines cooperate as regards both costs and revenues and so they can be considered as virtual mergers on specific routes (Fageda et al., 2019a). This may weaken competition on long-haul segments.

4. Empirical model

We estimate the following equations where the dependent variable refers to demand or fares of the hubbing airline *a* on routing *k* in period *t*:

$$Passengers\_hubbing_{akt} = \beta_0 + \beta_1 Flights\_LCC\_short\_segment_{kt} + \beta_2 Flights\_LCC\_short\_segment\_secondary_{kt} + \beta_3 HHI\_short\_segment_{kt} + \beta_4 high\_speed\_rail_{kt} + \beta_5 HHI\_long\_segment_{kt} + \beta_6 Flights\_LCC\_long\_segment_{kt} + \beta_7 Non\_stop\_flights_{kt} + \beta_8 Pop\_origin_{kt} + \beta_9 Pop\_destin_{kt} + \beta_{10} Income\_origin_{kt} + \beta_{11} Income\_destin_{kt} + \gamma' routing + \lambda' year + \mu' quarter + \varepsilon_{akt} \tag{1}$$

$$Fare\_hubbing_{akt} = \beta_0 + \beta_1 Flights\_LCC\_short\_segment_{kt} + \beta_2 Flights\_LCC\_short\_segment\_secondary_{kt} + \beta_3 HHI\_short\_segment_{kt} + \beta_4 high\_speed\_rail_{kt} + \beta_5 HHI\_long\_segment_{kt} + \beta_6 Flights\_LCC\_long\_segment_{kt} + \beta_7 Non\_stop\_flights_{kt} + \beta_8 Pop\_origin_{kt} + \beta_9 Pop\_destin_{kt} + \beta_{10} Income\_origin_{kt} + \beta_{11} Income\_destin_{kt} + \gamma' routing + \lambda' year + \mu' quarter + \varepsilon_{akt} \tag{2}$$

The dependent variables in equations (1) and (2) are the number of passengers channeled and the fares charged by the hubbing airline, respectively. Demand in inter-city aviation markets is commonly modelled through a gravity equation that considers as explanatory variables income and population of the endpoints and route distance although some studies also include airline-specific factors like fares per km as explanatory factor (see Grosche et al, 2017; Zhang et al., 2018 for detailed surveys). Competition variables may have an influence on the number of passengers through their effects on fares and quality.

Prices are commonly modeled as a mark-up over total costs where demand shifters and distance control for the potential exploitation of density and distance economies and the mark-up is a function of competition variables. To this point, note that the effect of distance is captured by routing fixed effects given that it is a time-invariant variable.

We include three variables that seek to capture airline competition on the short segments. First, we consider the route concentration, measured using the Herfindahl–Hirschman index (HHI), in terms of flight frequencies. In addition to route concentration, we consider a variable that identifies LCC flights on each route. Given that both variables are correlated due to the notable presence of LCCs on short-haul routes, we estimate two different specifications to disentangle the effect of the general level of competition from that driven by the LCCs. In the first specification, we consider the number of flights offered by LCCs. In the second, we consider the HHI variable and a dummy for high-frequency services of LCCs (and the interaction between the two variables). This latter dummy variable takes a value of one when the LCCs offer at least one daily flight on the short segment of the routing. To this point, these two latter variables are based on the flights offered by LCCs to short segments of the routings (ie; flights from a European airport to one of the six large European hubs). For example, on the routing Lisbon-Amsterdam-Washington Dulles, for this variable we compute Vueling and Easyjet flights from Lisbon to Amsterdam.

Furthermore, we include a variable that measures the flights of LCCs from secondary airports on the short segment. This variable considers LCC flights from the same airport of origin of the routing to a secondary airport that serves the same urban area as that in which the given hub airport is located. For example, on the routing Malaga-Paris Charles de Gaulle-Atlanta, for this variable we compute Vueling and Ryanair flights from Malaga to Paris Orly.

We also consider the HHI variable measured in terms of flight frequencies and the number of flights provided by LCCs on the long segments of a routing. The modest presence of LCCs on the long segments makes it unnecessary to use two different specifications of the competition variables as we do when identifying competition on the short segments of a routing.

We also consider a variable to identify the influence of high-speed rail services on hubbing operations. This variable is constructed as a binary variable taking a value of 1 for short-segments of the routing affected for HSR connections.

Unfortunately, we do not have available data for the frequencies offered by HSR that would be the ideal way to measure its impact as we do when considering the variables for LCCs. The dummy variable for HRS is aimed to approximate the increase in frequencies and the reduction in travel time that HSR is always able to achieve in comparison to conventional rail services. The lack of rail frequency information for several years is common in previous studies. In fact, all previous studies that examine the impact of HSR on air traffic and fares use a dummy variable as we do here (Albalate et al., 2015; Bergantino et al. 2015; Wan et al., 2016; Wei et al., 2017; Wang et al., 2018; Zhang et al., 2018; Liu et al., 2019). Only the study of Ma et al. (2019) for the route Beijing-Shanghai is able to add a variable for HSR frequency while that Bergantino et al. (2015) provide some information on frequencies in their analysis for intra-modal competition in Italy.

We consider an additional variable that may capture the intensity of competition when we use the sample of transatlantic routings; the number of non-stop flights (if any) from origin to destination. For example, the routing Brussels-Frankfurt-Washington offered by Lufthansa has to compete with the non-stop flights from Brussels to Washington provided by United and Brussels Airlines. The non-stop service is clearly superior in quality to the connecting service and so we should expect fewer passengers and lower fares on the connecting routes that have to compete with non-stop services.

In all equations, population and income at both endpoints of the routing are included as control variables. Recall that the origin is a European airport and the destination a North American airport. Demand should be higher when the endpoints are richer and more highly populated. The population variables could have indirectly an effect on fares due to their effect on the demand generated, while the income variable might capture the willingness of travelers to pay.

All continuous variables without zero values are expressed in logs. Unreported year, quarter, and routing fixed effects are also added in the regressions. Note that the routing fixed effects imply that the regressions focus on the within variation of the data so that the effect of variables such as distance, joint venture agreements or the major tourist attractiveness of the origin and/or destination of the routings are captured by these fixed effects. In this regard, the use of routing fixed effects allows us to control for unobserved factors that do not vary over time. For example, the routing fixed effects may allow controlling for the fact that some routings may be systematically suffering more delays than others due to more congestion in the airports of the routing. However, we must recognize here that a limitation of our analysis is that we cannot control for time varying shocks that may have an influence on the on-time performance of the service provided by the hubbing airline in the routing.

The year fixed effects are dummies for each year of the considered period, being 2010 the omitted year. The inclusion of year fixed effects allows us to control for unobserved shocks that are common to all routings like the price of fuel. The quarter fixed effects, which control for seasonal variations that are usual in aviation, are dummies for each quarter of the year, being the first quarter the omitted one. Standard errors are robust to heteroscedasticity and clustered by routing.

Table 1 shows the descriptive statistics of the variables used in the empirical analysis. In this regard, note that the frequency of most variables is quarterly with the exception of the variables of population and income at the origin and destination for which the frequency of data is annual. In this table, we show the within and between variation of the data. In general, the within variation seems to be sufficiently high for all variables except the variables for the population of the origin and destination of the routing and to lower extent the dummy variable for HSR and income of the origin.

Insert Table 1

In table 2, we show the correlation matrix between the variables used in the empirical analysis. Data in this table provide evidence that multicollinearity does not seem to be a major concern given that the correlation between the explanatory variables is generally low. Only moderate levels of correlation can be identified between HHI and LCC flights in the short segment, population at the origin and HHI in the short segment, and population at destination and HHI in the long segment.

Insert Table 2

Finally, Table 3 provides information about the routes in our sample affected by HRS services and about secondary airports that are close to the considered hub airports. We compare the current travel times of HSR in relation to air services. We add 60 minutes to the time by air to take into account that the access time to the infrastructure is higher for air services. The travel time ratio is more favorable for HSR in the shorter routes except for Frankfurt. Furthermore, data in table 3 show that Paris, Rome and London have one or more nearby airports that are less than one hour away by car.

Insert Table 3

5. Results

Table 4 shows the results of the estimates. Recall that we estimate two different specifications to disentangle the effect of the general level of competition from that driven by LCCs. First, we consider the number of LCC flights on the short segments as our explanatory variable without including the HHI variable. Second, we consider a dummy variable for high-frequency services provided by LCCs on the short segments, the HHI variable and the interaction between the two variables.

The control variables work as expected with the exception of the variable of population at the origin that is negative and statistically significant. This surprising result may be explained by the low within variation of this variable.

Insert Table 4

We find that the greater the number of LCC flights on the short segment, the fewer the passengers and the higher the fares on the transatlantic connecting routings. This result holds both when we consider the number of LCC flights as our explanatory variable and when we consider a dummy variable for high-frequency services of LCCs as our explanatory variable.

Furthermore, a more intense general level of competition in the short segment reduces the number of passengers, given that the coefficient of the HHI variable is positive and statistically significant. However, the effect of the HHI variable is not statistically significant in the fare regression. Note also that the interaction variable almost completely counteracts the effect of the HHI variable. This means that the impact of competition on the short segment of the transatlantic routing on the passengers of the hubbing airline is driven almost entirely by LCCs when these carriers are present. This can be attributed to the much lower fares that LCCs are able to charge in comparison to those charged by network airlines.

Competition on the short segments from LCCs at secondary airports also results in fewer passengers and higher fares; although, this outcome as it affects passengers is not statistically significant.

We also find that weaker competition on long segments leads to fewer passengers and higher fares. Indeed, the coefficient of the HHI variable on the long segments is negative and statistically significant in the passenger equation and is positive (although not statistically significant) in the fare equation. However, we do not find a significant effect of the variable for LCC flights. The very modest presence of LCCs on the long segments in the period considered presumably accounts for this result.

Overall, we find evidence that the "density economies hypothesis" is more relevant in explaining the results of the competition variables on the short segments and the "market power hypothesis" is more relevant in explaining the results of the competition variables on the long segments. The signs of the competition variables for both the passengers and fare regressions point in the expected direction. However, we do not find a statistically significant effect of the HHI variable on the long segments in the fare regression, which means we are unable to provide strong evidence in favor of the "market power hypothesis" in the case of the competition variables on the long segments. In any case, what seems clear from our results is that connecting travelers benefit from competition on the long segments, while they are harmed by competition on the short segments.

Furthermore, we fail to find a negative impact of HSR services on connecting travelers. The HSR variable is positive and statistically significant in the regression for passengers, but it is not statistically significant in the fare regression. To this point, it should be borne in mind that the routing fixed effects regressions only capture the within variation of the data, which is modest in the case of the HSR variable. Only a few routes to Paris and Madrid present any changes in the period considered.

Table 5 shows the results of the HSR variable for subsamples that consider each hub airport separately. We report the results for these subsamples using routing fixed effects and airport fixed effects. Our preferred regressions are those using routing fixed effects as they allow us to control for unobserved time-constant factors at the routing level. However, we report the additional regressions using airport fixed effects instead of routing fixed effects to provide further insights into the effect of HSR services given the modest within variation of this variable. To provide a more accurate analysis of the HSR effects, we differentiate in these regressions between routings affected by HSR with short segments having more and less than 400 kms (when the subsample allows us to make such differentiation). In this regard, the competitiveness of HSR vs flights in terms of travel time and frequency is generally higher in shorter trips.

In the routing fixed effects regressions, we find a positive effect of HSR on traffic in the connecting routes via Paris but not those via Madrid. HSR does not seem to influence the price setting of hubbing airlines on connecting routes via Madrid but air fares are lower in the routings via Paris when the short segment is less than 400 kms. Recall that

Paris has its own HSR station while Madrid does not. Hence, connecting passengers using HSR services to reach Paris Charles de Gaulle airport may help to feed the long flight and more than offset the lower number of passengers on the short segment.

Insert Table 5

When we use airport fixed effects, we find that HSR services impact positively on passengers in Amsterdam and Frankfurt (both with their own HSR station), in routings that have short segments with less than 400 kms in Paris and Rome and in routings that have short segments with more than 400 kms in Madrid. We only find a negative impact of HSR on passengers in the case of London where just the route to Paris is affected by HSR services. This is a particularly dense route with a high proportion of point-to-point passengers. In terms of fares, only in the case of Frankfurt does the hubbing airline charge higher fares on routings affected by HSR. Importantly, HSR services do not seem to harm the hubbing airline or the connecting passengers at Paris, Rome and Madrid airports where several domestic routes are affected by HSR competition.

6. Robustness checks

In this section, we run several robustness checks to examine different issues that may have an influence on the estimated results for the competition variables.

Regarding the fare equation, the fare data available provide a partial picture of the actual average fares paid by all travelers in each routing. In some routings, the number of tickets used to calculate the average fares may be low. Hence, the variable of fares may be affected by outliers or extreme values (see figure A1 in the appendix). To examine the potential distortion that this limitation in the fare data may impose on our results, table 6 shows the results of additional regressions of the fare equation in which we exclude observations with extreme values for the variable of fares per kilometer. In particular, we report results of regressions using subsamples that exclude observations with fares per km below and above 1%, 5% and 10%, respectively. Figure A1 in the appendix shows a smoother distribution of the variable of fares per km when we consider these subsamples.

Results for competition variables related with low-cost carriers remain essentially identical to those obtained in the baseline regression. However, the variable for HSR turns out positive and statistically significant when we exclude observations with fares per km below and above 5% and 10%. Results for the subsamples based on Paris Charles de Gaulle and Madrid airports show that this latter result only holds for the routes with short segments having more than 400 kms in the subsample based on Paris Charles de Gaulle airport. Recall here that HSR does not lead to less connecting passengers in the routing fixed effects regression that use that subsample.

Insert Table 6

Regarding the passengers equation, air fares are not included as explanatory factor in the baseline regression because the goal of this paper is to examine the impact of competition variables on passengers and fares. Given that competition variables affect simultaneously both passengers and fares, the specific impact of such competition variables is cleaner if we estimate separately the passenger and fares equations. This is the approach that have been used in previous studies about the impact of competition either spurred by low-cost airlines or high-speed rails on air traffic (Goolsbee & Syverson, 2008; Bettini and Oliveira, 2008; Fageda, 2014; Albalate et al., 2015; Bergantino et al., 2015; Wan et al, 2016; Wang et al., 2018; ; Ma et al., 2019; Zhang et al., 2018). Only the studies of Bettini et al. (2018) and Liu et al. (2019) include fuel costs as explanatory factor. To this point, the use of year fixed effects allows us to control for the impact of fuel price changes given that such changes are a common shock for all airlines.

Bearing this in mind, we report the results of a regression in column 2 of table 7 in which the passenger equation includes fares per kms as explanatory variable. To deal with the potential endogeneity bias, we use lagged values instead of contemporaneous values of the fares per kms variable. As expected, the variable of fares per km is negative and statistically significant. Results for competition variables related with low-cost carriers and high-speed rail remain essentially identical to those obtained in the baseline regression. Only the variable that accounts for the intensity of competition in the long segment loses its statistical significance.

In column 3 of table 7, we also show the results of a regression of the passengers equation in which we include a variable that measures the flight frequency provided by the hubbing airline in the routing as explanatory variable. While this variable may account for an important airline service attribute as schedule convenience, we do not include it in the baseline regression because it may be correlated with competition variables. To deal with the potential endogeneity bias of this variable, we use lagged values instead of contemporaneous values. Note that in column 4 of table 7 we also report the results of a regression of the passengers equation including both the fares per km and the flight frequency provided by the hubbing airline as explanatory variables.

As expected, the variable of flight frequency provided by the hubbing airline is positive and statistically significant. Results for competition variables related with low-cost carriers and high-speed rail remain essentially identical to those obtained in the baseline regression. Again, only the variable that accounts for the intensity of competition in the long segment loses its statistical significance.

The sample considered in this study has the structure of panel data given that we have available information on 8088 routings for several time periods. The two main panel data models are random effects and fixed effects. The fixed effects model is usually the preferred model because it allows controlling for unobserved factors that do not vary over time and random effects may impose a bias in the estimation given that they may be correlated with the rest of explanatory variables. Hence, the baseline regression is based on a routing fixed effects model that is able to capture the effect of time invariant factors like distance, joint venture agreements, a more systematic congestion in any of the airports of the routings or the major tourist attractiveness of the origin and/or destination of the routings. However, a limitation of the fixed effects model is that time invariant variables like a dummy for major tourist destinations cannot be considered because their effect is already captured by the fixed effects. Furthermore, the routing fixed effects imply that the regressions focus on the within variation so that the estimation of the impact of variables with low within variation may be imprecise. In contrast, the random effect model exploits the between and within variation of the data.

Insert Table 7

To this point, in column 5 of table 7 we report the results of an additional regression where the estimation is made using random effects and adding as explanatory variables the distance of the routing and dummies for major tourist destinations for the origin (EU cities) and destination (US cities) of the routing. In the EU, all airports on the following islands are considered tourist destinations: the Balearic and Canary Islands (Spain), Sardinia and Sicily (Italy), Corsica (France), and many Greek islands, together with the airports of Alicante (ALC), Bari (BRI), Faro (FAO), Malaga (AGP) and Nice (NCE). In the US, Las Vegas (LAS), Orlando (MCO) and Fort Lauderdale-Hollywood (FLL) airports are considered major tourist destinations. Note that tourist variables seek to identity the impact on air traffic of origin and/or destination points where tourism intensity (number of tourists in relation to the inhabitants) is particularly high. Many big cities in our sample like London, Paris, Rome or New York also receive many international tourists each year but tourism intensity is not so high as for the places identified through the tourism dummy variables.

Results for competition variables related with low-cost carriers and high-speed rail remain essentially identical to those obtained in the baseline regression that uses routing fixed effects. Only the sign of the variable that measures the flights of LCCs from secondary airports on the short segment turns out positive and statistically significant. The distance variable is not statistically significant which may be because the sample is based on long-haul routings. The dummy for major tourist destinations at the origin is negative and statistically significant. This may be explained by the fact that few international tourists in these destinations come from US. In contrast, the dummy for major tourist destinations at the destination is positive and statistically significant. Note also that the variable that measures the population at the origin is positive and statistically significant so that the negative effect found for this variable in the baseline regression may be explained by its low within variation.

7. Conclusion

In this paper, we have examined the effects of competition driven by low-cost airlines and high-speed rail on the hubbing operations of European network airlines at their main hubs. More specifically, we have analyzed the effects of competition on the number of passengers channeled and on fares charged by these hubbing airlines on transatlantic routings with a stopover at their hubs.

Overall, our results suggest that competition from LCCs on short-haul segments of transatlantic routings reduces the number of connecting passengers channeled by hubbing airlines on transatlantic routings and this smaller number of connecting travelers have to pay higher fares.

Our results also provide some evidence that competition on the long-haul segments of transatlantic routings, which is usually driven by network airlines, may result in an increase in the number of connecting passengers on hubbing airlines. In these more competitive environments connecting passengers do not pay higher fares.

Hence, we conclude that connecting passengers may benefit from airline competition on long-haul segments of transatlantic routings, while they clearly suffer detrimental effects from such competition on short-haul segments.

Furthermore, we find that the presence of HSR does not generally lead to a reduction in the number of connecting passengers or an increase in prices. In fact, the number of connecting passengers channeled by hubbing airlines may increase on those routings where the short-haul segment is affected by HSR services, especially when the airport has its own HSR station. A potential explanation for this complementary role of HSR is that many travelers seem likely to use HSR services to arrive at the hub airport to take their long-haul flight.

In short, LCCs may provide benefits in terms of lower fares and more traffic on the short-haul routes for which they provide services. When these LCCs operate out of large hub airports, however, this positive effect may not offset the detrimental impact on both hubbing airlines and connecting passengers using their hub airports as a stopover. Bearing in mind that large hub airports are usually highly congested, measures to deter the expansion of LCCs at these sites could be beneficial from a general welfare perspective. Given that we may expect that LCCs react more to an increase of airport charges than the hubbing airline, such increase in airport charges may be positive for connecting travelers.

In contrast, HSR services do not seem to have a detrimental impact on hubbing operations. Travelers on short-haul routes may benefit from the convenience of the service provided by HSR and this does not seem to impact the number of connecting travelers negatively. Furthermore, the increase in HSR services at the expenses of airline services may be beneficial for the environment. These potential benefits, however, need to be balanced with the high costs associated with the building and maintenance of HSR lines.

Acknowledgements

We acknowledge financial support from the Spanish Ministry of Economy and Competitiveness and AEI/FEDER-EU (RTI2018-096155-B-I00) and Generalitat de Catalunya (2017SGR644). We are grateful to Xiaowen Fun and two anonymous reviewers for their useful comments

References

1. Abate, M., Christidis, P. (2020) ‘The impact of air transport market liberalization: Evidence from EU’s external aviation policy’, *Economics of Transportation*, 22, 100164.
2. Albalate, D., Bel, G., Fageda, X. (2015) ‘Competition and cooperation between high-speed rail and air transportation services in Europe’, *Journal of Transport Geography*, 42, pp. 166–174.

3. Bergantino, A. S., Capozza, C., Capurso, M. (2015) ‘The impact of open access on intra and inter-modal rail competition. A national level analysis in Italy’, *Transport Policy*, 39, pp. 77–86.
4. Bernardo, V., Fageda, X. (2017) ‘The effects of the open skies agreement between Morocco and the European Union: A differences-in-differences analysis’, *Transportation Research Part E: Logistics and Transportation Review*, 98, pp. 24–41.
5. Berry, S., Sarnall, M., Spiller, P. (2006) ‘Airline hubs: costs, markups and the implications of customer heterogeneity’, In: Lee, D. (Ed.), *Advances in Airline Economics*, vol. 1. Elsevier, Amsterdam, pp. 183–214.
6. Bettini, H. F., Oliveira, A. V. (2008) ‘Airline capacity setting after re-regulation: The Brazilian case in the early 2000s’, *Journal of air Transport management*, 14, pp. 289–292.
7. Bettini, H. F., Silveira, J. M. F., Oliveira, A. V. (2018) ‘Estimating strategic responses to the march of a low cost carrier to primary airports’, *Transportation Research Part E: Logistics and Transportation Review*, 109, pp. 190–204.
8. Boonekamp, T., Zuidberg,J., Burghouwt, G. (2018) ‘Determinants of air travel demand: The role of low-cost carriers, ethnic links and aviation-dependent employment’, *Transportation Research Part A: Policy and Practice*, 112, pp. 18-28.
9. Brueckner, J.K., Spiller, P.T. (1994) ‘Economies of traffic density in the deregulated airline industry’, *Journal of Law and Economics*, 37, pp. 379–415.
10. Button, K., Martini, G., Scotti, D., Volta, N. (2019) ‘Airline regulation and common markets in Sub-Saharan Africa’, *Transportation Research Part E: Logistics and Transportation Review*, 129, pp. 81-91.
11. Caves, D.W., Christensen, L.R., Tretheway, M.W. (1984) ‘Economies of density versus economies of scale: why trunk and local service airline costs differ’, *The RAND Journal of Economics*, 15, pp. 471–489.
12. Clewlow, R.R., Sussman, J.M., Balakrishnan, H. (2012) ‘Interaction of high-speed rail and aviation’, *Transport. Res. Rec.: Journal of the Transportation Research Board*, 2266, pp. 1–10.
13. De Poret, M., O'Connell, J.F., Warnock-Smith, D. (2015) ‘The economic viability of long-haul low cost operations: Evidence from the transatlantic market’, *Journal of Air Transport Management*, 42, pp. 272–281
14. Dresner, M., Lin, J.S.C., Windle, R. (1996) ‘The impact of low-cost carriers on airport and route competition’, *Journal of Transport Economics and Policy*, 30, pp. 309–329
15. Dobruszkes F., Givoni M., Vowles T. (2017), ‘Hello major airports, goodbye regional airports? Recent changes in European and US low-cost airline airport choice’, *Journal of Air Transport Management*, 59, pp. 50–62.
16. Dobruszkes, F., Dehon, C., Givoni, M. (2014) ‘Does European high-speed rail affect the current level of air services? An EU-wide analysis’, *Transportation Research Part A: Policy and Practice*, 69, pp. 461–475.

17. Fageda, X. (2014) ‘What hurts the dominant airlines at hub airports?’, *Transportation Research Part E: Logistics and Transportation Review*, *70*, pp. 177–189.
18. Fageda, X., Flores-Fillol, R., Theilen, B. (2019a) ‘Hybrid cooperation agreements in networks: The case of the airline industry’, *International Journal of Industrial Organization*, 62, pp. 194–227.
19. Fageda, X., Suárez, A., Serebrisky, T., Fioravanti, R. (2019b) ‘Estimating the Impact of Public Policies Worldwide on Air Connectivity in Remote Regions’, *Regional Studies*, 53, pp. 1161-1169.
20. Givoni M., Dobruszkes F. (2013) ‘A review of ex-post evidence for mode substitution and induced demand following the introduction of High-Speed Rail’, *Transport Reviews*, 33, pp. 720–742.
21. Goolsbee, A., Syverson, C. (2008) ‘How do incumbents respond to the threat of entry? Evidence from the major airlines’, *The Quarterly journal of economics*, *123*, pp. 1611-1633.
22. Grosche, T., Rothlauf, F., Heinzl, A. (2007) ‘Gravity models for airline passenger volume estimation’, *Journal of Air Transport Management*, 13, pp. 175–183.
23. Hofer, C., Windle, R.J. & Dresner, M.E. (2008) ‘Price premiums and low cost carrier competition’, *Transportation Research Part E: Logistics and Transportation Review*, 44, pp. 864–882.
24. Huse, C., & Oliveira, A. V. (2012) ‘Does product differentiation soften price reactions to entry? Evidence from the airline industry’, *Journal of Transport Economics and Policy*, 46, pp. 189–204.
25. Liu, S., Wan, Y., Ha, H.K., Yoshida, Y., Zhang, A. (2019) ‘Impact of high-speed rail network development on airport traffic and traffic distribution: Evidence from China and Japan’, *Transportation Research Part A*, 127, pp. 115–135.
26. Ma, W., Wang, Q., Yang, H., Zhang, A., & Zhang, Y. (2019) ‘Effects of Beijing-Shanghai high-speed rail on air travel and fare: Passenger types, airline groups and tacit collusion’ *Research in Transportation Economics*, *74*, pp. 64–76.
27. Mantin, B., Koo, B. (2009) ‘Dynamic price dispersion in airline markets’ *Transportation Research Part E: Logistics and Transportation Review,* 45, pp. 1020–1029.
28. Morrison, S.A. (2001) ‘Actual, adjacent and potential competition: estimating the full effect of Southwest airlines’, *Journal of Transport Economics and Policy*, 35, pp. 239–256.
29. Murakami, H. (2011) ‘Time effect of low-cost carrier entry and social welfare in US large air markets’ *Transportation Research Part E: Logistics and Transportation Review*, *47*, pp. 306–314.
30. Njoyaa, E..T, Christidis, P., Nikitasa (2018) ‘Understanding the impact of liberalisation in the EU-Africa aviation market’, *Journal of Transport Geography*, 71, 161-171.
31. Official Airlines Guide (2019) ‘Data Traffic Analyzer’.

32. Oliveira, A.V.M., Huse, C. (2009) ‘Localized competitive advantage and price reactions to entry: full-service vs. low-cost airlines in recently liberalized emerging markets’, *Transportation Research Part E: Logistics and Transportation Review*, 45, pp. 307–320.
33. Qin, M.S., John, G., Vitorino, M.A. (2018) ‘Planes, trains, co-opetition: evidence from China’, Working Paper (University of Minnesota).
34. RDC aviation (2019) ‘Capstats statistics’ (Innovata data).
35. Scotti, D., Volta, N. (2018) ‘Price asymmetries in European airfares’, *Economics of Transportation*, 14, pp. 42-52.
36. Soyk, C., Ringbeck, J., Spinler, S. (2018) ‘ Revenue characteristics of long-haul low cost carriers (LCCs) and differences to full-service network carriers (FSNCs)’, *Transportation Research Part E: Logistics and Transportation Review* 112, pp. 47–65.
37. Soyk, C., Ringbeck, J., Spinler, S. (2017) ‘Long-haul low cost airlines: Characteristics of the business model, and sustainability of its cost advantages’, *Transportation Research Part Part A: Policy and Practice*, 106, pp. 215–234.
38. United Nations (2019) ‘World Urbanization Prospects’, data available at: https://population.un.org/wup/
39. Varella, R., Frazão, J., Oliveira, A., (2017) ‘Dynamic Pricing and Market Segmentation Responses to Low-Cost Carrier Entry’, *Transportation Research Part E: Logistics and Transportation Review*, 98, pp. 151–170.
40. Wan, Y., Ha, H.-K., Yoshida, Y., Zhang, A. (2016) ‘Airlines’ reaction to high-speed rail entries: empirical study of the Northeast Asian market’, *Transportation Research Part A: Policy and Practice*, 94, pp. 532–557.
41. Wang, K., Xia, W., Zhang, A., Zhang, Q (2018) , ‘Effects of train speed on airline demand and price: Theory and empirical evidence from a natural experiment’, *Transportation Research Part B: Methodological*, 114, pp. 99-130
42. Wei, F., Chen, J., Zhang, L. (2017) ‘Demand shocks, airline pricing, and high-speed rail substitution: Evidence from the Chinese Market’, *Journal of Transport Economics and Policy*, *51*, pp. 266–289.
43. Windle, R., Dresner, M., (1999) ‘Competitive responses to low cost carrier entry’, *Transportation Research Part E: Logistics and Transportation Review*, 35, pp. 59–75.
44. Windle, R., Dresner, M. (1995) ‘The short and long run effects of entry on US domestic air routes’, *Transportation Journal*, 12, pp. 14–25.
45. Wong, W.H., Zhang, A, Cheung, T.K.Y, Chu, J. (2019) ‘Examination of low-cost carriers' development at secondary airports using a comprehensive world airport classification’, *Journal of Air Transport Management*,78, pp. 96–105.
46. World Bank (2019) ‘Word Bank Development indicators’, data available at: https://databank.worldbank.org/source/world-development-indicators
47. Zhang, F., Graham, D.J., Wong, M.S.C. (2018) ‘Quantifying the substitutability and complementarity between high-speed rail and air transport’, *Transportation Research Part A: Policy and Practice*, 118, pp. 191–215.

48. Zhang, Y., Lin, D. & Zhang, A (2018) 'Gravity models in air transport research: A survey and application' in Blonigen, B.A & Wilson, W. (eds), *Handbook of International trade and transportation*, 141-158, Edward Elgar, Cheltenham (UK).
49. Zhang, Y., Zhang, A., Zhu, Z., Wang, K. (2017) 'Connectivity at Chinese airports: the evolution and drivers', *Transportation Research Part A: Policy and Practice*, 103, pp. 490–508.
50. Zhang, A., Wan, Y., Yang, H. (2019) 'Impacts of high-speed rail on airlines, airports and regional economies: A survey of recent research', *Transport Policy*, 81, pp. A1-A19.

TABLES

Table 1. Descriptive statistics of the variables used in the empirical analysis

Variable	Frequency	Mean	Standard deviation	Standard deviation (between)	Standard deviation (within)	Ratio (Sd. between /Sd.within)	Number observations
Total passengers_hubbing	Quarterly	191.22	229.58	165.25	116.48	1.42	138,740
Fare_hubbing (USD)	Quarterly	623.32	622.85	465.56	504.83	0.92	119,859
Flights_hubbing	Quarterly	533.78	335.93	304.55	118.11	2.58	138,320
Flights_LCC_short_segment	Quarterly	32.06	90.47	85.53	41.55	2.06	136,292
Flights_LCC_secondary_airports	Quarterly	120.29	270.52	246.02	78.32	3.14	138,740
$D^{high\_frequency\_LCC\_short\_sgment}$	Quarterly	0.14	0.35	0.33	0.19	1.67	136,292
HHI_short_segment	Quarterly	0.75	0.24	0.23	0.08	2.77	136,292
$D^{High\_speed\_rail}$	Quarterly	0.07	0.27	0.25	0.03	8.09	138,740
HHI_long_segment	Quarterly	0.65	0.26	0.25	0.09	2.78	138,320
Flight_LCC_long_segment	Quarterly	3.02	13.04	11.01	9.87	1.12	138,740
Non-stop flights	Quarterly	22.72	87.28	75.09	16.74	4.49	138,740
Population_origin (000 inhabitants)	Annual	1667.57	2011.65	1989.25	51.21	38.84	134,031
Population_destination (000 inhabitants)	Annual	6578.08	5060.87	5094.46	127.45	39.97	138,627
Income_origin (USD)	Annual	40252	16517.18	16519.76	1588.97	10.40	137,072
Income_destination (USD)	Annual	50223.73	1478.50	1104.52	1306.31	0.85	138,740
Distance (kms)	Time invariant	7892.26	1303.29	1303.29	0	-	138,740

Table 2. Correlation matrix of the variables used in the empirical analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Passengers_hubbing (1)	1													
Fare_hubbing (2)	0.009	1												
Flights_LCC_short_segment (3)	0.03	-0.07	1											
Flights_LCC_secondary_airports (4)	0.26	0.09	-0.04	1										
HHI_short_segment (5)	-0.13	0.04	-0.40	-0.06	1									
$D^{High\_speed\_rail}$ (6)	0.05	-0.005	-0.02	-0.08	-0.08	1								
HHI_long_segment (7)	-0.26	0.07	0.02	-0.14	0.02	0.01	1							
Flight_LCC_long_segment (8)	0.15	-0.01	-0.01	0.08	0.01	-0.001	-0.22	1						
Non-stop flights (9)	0.06	-0.02	-0.02	0.01	-0.19	0.21	-0.12	0.06	1					
Population_origin (10)	0.20	-0.01	0.12	0.09	-0.40	0.20	0.001	-0.01	0.30	1				
Population_destination (11)	0.21	-0.12	0.05	-0.07	-0.001	0.02	-0.43	0.12	0.19	-0.07	1			
Income_origin (12)	0.03	0.13	0.01	0.05	0.07	0.07	0.001	0.09	0.09	-0.13	-0.02	1		
Income_destination (13)	0.04	-0.01	0.04	0.10	0.05	-0.02	-0.06	0.01	0.02	0.01	0.11	0.02	1	
Flights_hubbing_airline (14)	0.39	0.11	0.04	0.27	-0.20	0.27	-0.13	0.10	0.43	0.34	0.08	0.23	-0.05	1

Table 3. Data on HSR and secondary airports

Airport	Routes with less than 400 kms		Routes with more than 400 kms		Secondary airports	
	Number routes	Ratio travel time HSR/air	Number routes	Ratio travel time HSR/air	Distance to hub airport (kms)	Minimum driving time (minutes) to hub airport
Amsterdam	4	1.8	0	-	-	-
Paris	6	1.2	8	1.5	Orly (43), Beauvais-Tille (82)	Orly (35), Beauvais-Tille (51)
Rome	3	0.9	2	1.6	Ciampino (41)	Ciampino (35)
Frankfurt	2	1.8	1	1.7	-	-
London	1	1.1	0	-	Gatwick (61), Stansted (101), Luton (57), Southend (121), City (39)	Gatwick (42), Stansted (62). Luton (41), Southend (83), City (68)
Madrid	3	1.1	2	1.2	-	-

Source: Voyages-SNCF and Google maps

Table 4. Results of estimates (baseline regressions)

Dependent variable	Passsengers		Fares	
Flights_LCC_short_segment	-0.0001 (0.000005)***	-	0.0002 (0.00006)***	-
Flights_LCC_secondary_airports	-0.00003 (0.00002)	-0.00003 (0.00002)	0.0001 (0.00002)***	0.0001 (0.00002)***
HHI_short_segment	-	0.09 (0.02)***	-	0.02 (0.02)
$D^{high\_frequency\_LCC}$	-	-0.06 (0.02)***	-	0.09 (0.02)***
$D^{high\_frequency\_LCC}$ X HHI_short_segment	-	-0.07 (0.03)**	-	0.03 (0.03)
$D^{High\_speed\_rail}$	0.22 (0.06)***	0.22 (0.06)***	0.06 (0.04)	0.03 (0.04)
HHI_long_segment	-0.08 (0.01)***	-0.08 (0.01)***	0.02 (0.01)	0.02 (0.01)
Flight_LCC_long_segment	0.00005 (0.0002)	0.00005 (0.0002)	0.0001 (0.0002)	0.0001 (0.0002)
Non-stop flights	-0.002 (0.0001)***	-0.002 (0.0001)***	-0.0008 (0.0001)***	-0.0008 (0.0001)***
Population_origin	-1.39 (0.26)***	-1.32 (0.26)***	-0.15 (0.23)	-0.13 (0.24)
Population_destination	0.65 (0.15)***	0.64 (0.15)***	-0.12 (0.14)	-0.12 (0.14)
Income_origin	0.35 (0.09)***	0.35 (0.09)***	0.28 (0.07)***	0.29 (0.07)***
Income_destination	5.78 (0.57)***	5.75 (0.57)***	3.74 (0.51)***	3.68 (0.51)***
Year fixed effects	YES	YES	YES	YES
Quarter fixed effects	YES	YES	YES	YES
Routing fixed effects	YES	YES	YES	YES
R2	0.13	0.13	0.03	0.03
Observations	131,474	131,474	113,829	113,829

Notes: Standard errors in parentheses (robust to heteroscedasticity and clustered by routing). Statistical significance at 1% (***), 5% (**), 10% (*).

Table 5. Results for the HSR variable (subsamples by hub airports)

	Routing fixed effects				Airport fixed effects			
	Passengers		Fares		Passengers		Fares	
	$D^{High\_speed\_rail}$ (< 400 kms)	$D^{High\_speed\_rail}$ (> 400 kms)	$D^{High\_speed\_rail}$ (< 400 kms)	$D^{High\_speed\_rail}$ (> 400 kms)	$D^{High\_speed\_rail}$ (< 400 kms)	$D^{High\_speed\_rail}$ (> 400 kms)	$D^{High\_speed\_rail}$ (< 400 kms)	$D^{High\_speed\_rail}$ (> 400 kms)
Amsterdam	-	-	-	-	0.74 (0.08)***	-	-0.20 (0.13)	-
Paris	0.47 (0.12)***	0.16 (0.03)**	-0.80 (0.12)***	0.02 (0.05)	0.28 (0.16)*	-0.02 (0.10)	-0.72 (0.12)***	0.06 (0.04)
Rome	-	-	-	-	4.10 (1.98)**	2.41 (1.65)	-0.47 (1.84)	-0.16 (1.53)
Frankfurt	-	-	-	-	2.26 (0.57)***	3.87 (1.31)***	1.87 (0.53)***	2.12 (1.11)**
London	-	-	-	-	-3.38 (1.53)**	-	-0.36 (1.56)	-
Madrid	0.02 (0.08)	-	-0.13 (0.12)	-	0.01 (0.08)	5.05 (1.45)***	-0.07 (0.12)	0.81 (2.28)

Notes: Standard errors in parentheses (robust to heteroscedasticity and clustered by route/routing). Statistical significance at 1% (***), 5% (**), 10% (*).

Table 6. Robustness checks (fare regressions)

Sample	Excluding observations with fares per km below and above 1% (1)	Excluding observations with fares per km below and above 5% (2)	Excluding observations with fares per km below and above 10% (3)
Flights_LCC_short_segment	0.0002 (0.00005)***	0.0002 (0.00005)***	0.0002 (0.00004)***
Flights_LCC_secondary_airports	0.0001 (0.00003)***	0.0001 (0.00003)***	0.0001 (0.00002)***
$D^{High\_speed\_rail}$	0.07 (0.04)	0.10 (0.04)***	0.13 (0.03)***
HHI_long_segment	0.01 (0.01)	-0.007 (0.02)	-0.01 (0.01)
Flight_LCC_long_segment	0.0001 (0.0002)	-0.00007 (0.0001)	-0.0001 (0.0001)
Non-stop flights	-0.0008 (0.0001)***	-0.0007 (0.0001)***	-0.0005 (0.0001)***
Population_origin	0.25 (0.22)	0.53 (0.19)***	0.59 (0.17)***
Population_destination	-0.19 (0.13)	-0.15 (0.12)	-0.20 (0.11)*
Income_origin	0.29 (0.07)***	0.33 (0.06)***	0.30 (0.06)***
Income_destination	4.35 (0.49)***	4.91 (0.43)***	4.50 (0.40)***
R2	0.03	0.02	0.02
Observations	111,911	103,113	91,811

Notes: All regressions include year, quarter and routing fixed effects. Standard errors in parentheses (robust to heteroscedasticity and clustered by routing). Statistical significance at 1% (***), 5% (**), 10% (*). Results of the HSR variable for the subsample based on Paris CDG. (1) & less 400 km: -0.76 (0.12)***, (1) & more 400 km: 0.05 (0.04); (2) & less 400 km: -0.56 (0.12)***, (2) & more 400 km: 0.09 (0.04)**; (3) & less 400 km: -0.37 (0.12)***, (3) & more 400 km: 0.10 (0.04)***. Results of the HSR variable for the subsample based on Madrid. (1): -0.15 (0.02); (2) -0.10 (0.10); (3): -0.04 (0.11).

Table 7. Robustness checks (passengers equation)

	fares per km as covariate (1)	Flights_hubbing_airline as covariate (2)	(1) & (2) as covariates	Random effects model
Fares per km	-0.03 (0.002)***	-	-0.03 (0.003)***	-
Flights_hubbing_airline	-	0.0002 (0.00002)***	0.0002 (0.00002)***	-
Flights_LCC_short_segment	-0.0002 (0.000005)***	-0.0002 (0.000005)***	-0.0002 (0.000003)***	-0.0001 (0.00005)***
Flights_LCC_secondary_airports	-0.00006 (0.00003)**	-0.00007 (0.00003)**	-0.00007 (0.00003)**	0.0001 (0.00002)***
$D^{High\_speed\_rail}$	0.28 (0.05)***	0.27 (0.05)***	0.28 (0.05)***	0.13 (0.03)***
HHI_long_segment	-0.008 (0.02)	-0.02 (0.02)	-0.02 (0.02)	-0.14 (0.01)***
Flight_LCC_long_segment	-0.0002 (0.0002)	-0.0002 (0.0002)	-0.0003 (0.0002)	0.00002 (0.00002)
Non-stop flights	-0.002 (0.0002)***	-0.002 (0.0002)***	-0.002 (0.0002)***	-0.002 (0.0001)***
Distance	-	-	-	-0.07 (0.06)
Population_origin	-1.55 (0.30)***	-1.38 (0.28)***	-1.48 (0.30)***	0.32 (0.01)***
Population_destination	0.37 (0.17)**	0.41 (0.16)**	0.34 (0.17)**	0.37 (0.01)***
Income_origin	0.34 (0.10)***	0.22 (0.09)**	0.28 (0.10)***	0.21 (0.02)***
Income_destination	6.28 (0.63)***	5.98 (0.62)***	6.30 (0.63)***	4.81 (0.52)***
Tourism_origin	-	-	-	-0.13 (0.03)***
Tourism_destination	-	-	-	0.34 (0.05)***
Year fixed effects	YES	YES	YES	YES
Quarter fixed effects	YES	YES	YES	YES
Routing fixed effects	YES	YES	YES	NO
R2	0.10	0.10	0.10	0.13
Observations	81,663	90,872	81,511	131,474

Notes: Standard errors in parentheses (robust to heteroscedasticity and clustered by routing). Statistical significance at 1% (***), 5% (**), 10% (*). We use lagged values of Fares per km and Flights_hubbing_airline.

FIGURES

Figure 1. Network

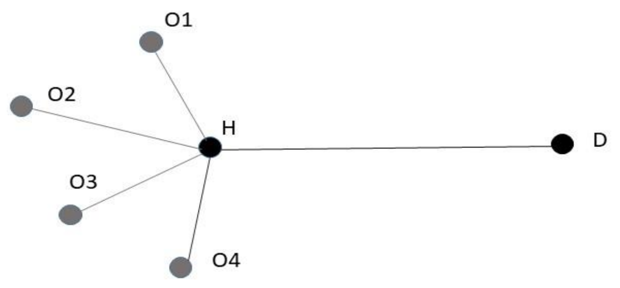
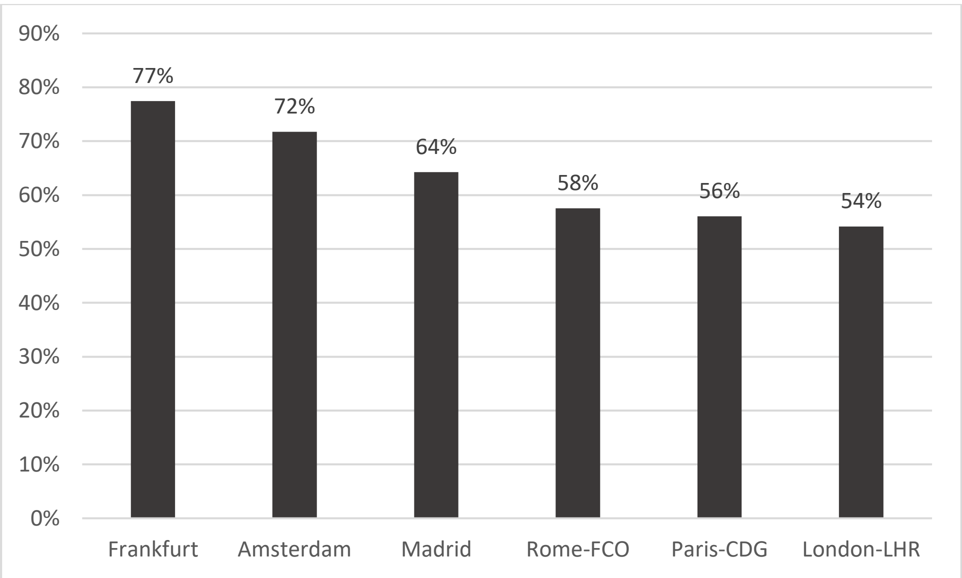


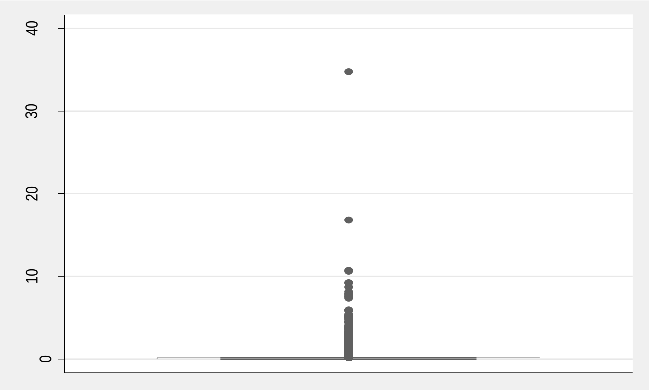
Figure 2. Proportion of connecting passengers over total passengers in the transatlantic market



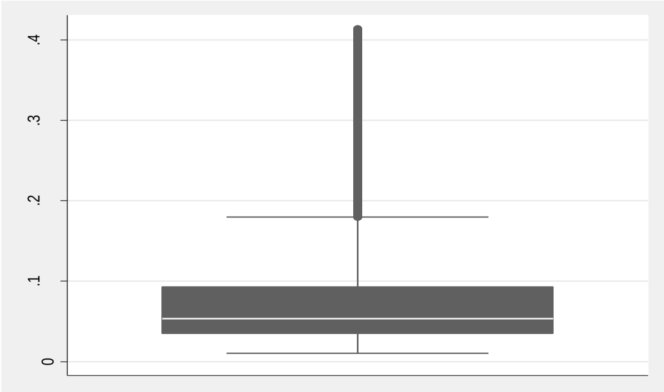
Note: Data comes from OAG (third quarter of 2017)

APPENDIX

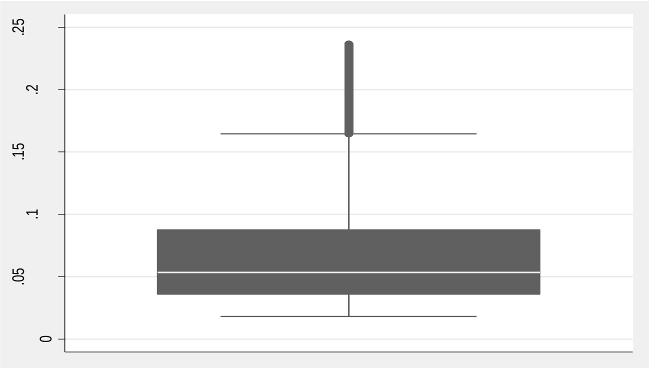
Figure A1. Box plot of fares per kilometer



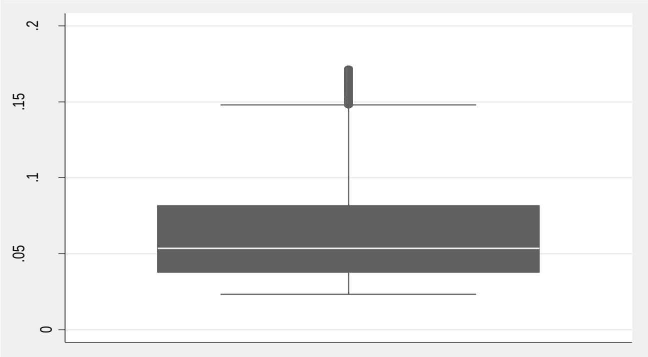
Note: All sample



Note: Excluding observations with fare per kilometer below and above 1%



Note: Excluding observations with fare per kilometer below and above 5%



Note: Excluding observations with fare per kilometer below and above 10%