GLOBALIZATION, LONG-HAUL FLIGHTS AND INTER-CITY CONNECTIONS

Valeria Bernardo

GiM-IREA, University of Barcelona. Address: Diagonal Av. 690, 08034, Barcelona, Spain. Mail to: valeriabernardo@ub.edu.

Xavier Fageda*

GiM-IREA, University of Barcelona. Address: Diagonal Av. 690, 08034, Barcelona, Spain. Mail to: xfageda@ub.edu

Abstract: We estimate a model that identifies the impact of long-haul non-stop flights on traffic flows in inter-city markets, using data that account for the true origin and destination of passengers. In particular, we analyze whether long-haul non-stop flights may have a positive and significant effect on traffic flows in a context where indirect connectivity is high. Furthermore, we examine whether the impact of non-stop flights on traffic is associated to the income levels of the destination. We find that even modest increases of non-stop flights lead to significant traffic gains but such traffic gains are only clearly relevant when the destination city is located in a high-income country.

Keywords: Air transportation; inter-city connections; long-haul non-stop flights; matching analysis; instrumental variables

JEL codes: R10, R11, R40, R41

*Corresponding author

0. Introduction

The availability of long-haul non-stop flights may promote the increase of investments, trade, tourism, immigration and knowledge flows between the cities that benefit from those flights. In this regard, total air passenger flows between two distant cities provide a precise measure of the economic and social links between them. Thus, we can examine the expected positive economic effects of long-haul non-stop flights through their impact on those air passengers flows.

The number of connecting passengers may be high in long-haul inter-city markets even if these city-pair markets are not served by non-stop flights.[1] Thus, the new availability of non-stop flights may simply lead to a switch from connecting to non-stop traffic that does not necessarily imply the generation of additional traffic. Disentangling the impact of direct connectivity may be particularly relevant for cities that have airports that are not major hubs but that are well connected with major hubs.

Furthermore, the main advantage of non-stop flights over connecting flights should be the reduction in time costs. Thus, the traffic gains of non-stop flights may be concentrated in rich city-pairs with a high proportion of passengers sensitive to time costs.

In this paper, we estimate a model that quantifies the impact of long-haul non-stop flights on traffic flows in inter-city markets, using data that account for the true origin and destination of passengers. Hence, our bilateral air traffic data include non-stop and connecting traffic. In particular, we analyze whether long-haul non-stop flights may have a positive and significant effect on traffic flows in a context where indirect connectivity is high.

A second goal in our analysis is to examine whether the impact of non-stop flights on traffic is associated to the income levels of the destination. If the impact of non-stop flights on traffic can be mainly explained by the reduction in time costs, then such impact could be only relevant for high-

---

[1] Looking at the data from the Official Airlines Guide (OAG), we can observe that the amount of traffic between distant cities may be high even without non-stop flights.

income destinations. This would suggest a certain limitation to the positive economic impact of long-haul non-stop flights as they would be only relevant where income levels are already high.

Long-haul non-stop flights between city pairs have traditionally been operated by network airlines through their major hub airports. [2] However, several factors have led to an increase in the provision of long-haul flights from airports other than the main hubs. Congestion at the main hubs is associated with lower quality levels in the form of delays and cancellations (Brueckner, 2002; Mayer and Sinai, 2003). Passengers may have a higher preference for direct flights than they had in the past (Berry and Pia, 2010). New aircraft models like the Airbus A350 or the Boeing 787 Dreamliner may allow for the provision of long-haul flights in thin routes. Finally, in Europe changes in the airline market, such as the arrival of low-cost carriers (e.g. Norwegian) in the long-haul market, or the increasing presence of airlines from North America and Asia benefit non-hub airports.

These trends may imply that cities with mid-sized airports benefit from globalization through their increased connectivity with the rest of the world. Having said this, such cities are generally well connected by means of frequent flights (and usually low fares) to the major hub airports. Thus, they may benefit from global connections through major hub airports as well.

We take advantage of the recent shift in the supply of non-stop flights from European cities with mid-sized airports to analyze the causal relationship between long-haul non-stop flights and total passenger flows. To this end, we use quarterly data on air traffic from European cities to large urban areas across the rest of the world for the period 2010-2016.

A major econometric challenge in this analysis is the potential endogeneity bias of the variable for the number of non-stop flights. To deal with this issue we apply a matching procedure to compare comparable routes so that we built a sample with city-pairs with and without non-stop flights similar in terms of the covariates and traffic levels in the baseline period. Furthermore, we include city-pairs

---

[2] By adopting this strategy, airlines are able to reduce their costs through the exploitation of density economies and by covering their network of routes with fewer flights (Brueckner and Spiller, 1994).

fixed effects in the estimation to control for differences between city-pairs due to unobservable factors that are time invariant and time fixed effects to control for unobserved shocks that are common to all city-pairs.

Finally, we use an instrumental variables approach to overcome the possible bias due to heterogeneity that is unobservable and changing over time. In this regard, the instruments that seem to correct the expected upward bias are based on lags of the potentially endogenous explanatory variable.

Various studies have undertaken empirical analyses concerning the economic impact of air traffic. The traditional approach is to examine the causal relationship between an aggregated indicator of airport size and measures of urban or regional economic performance, such as, employment, income or number of firms. Here, we can mention some studies using samples of urban areas in the United States (Bilotkach, 2015; Bloningen and Cristea, 2015; Brueckner, 2003; Cristea and Danila, 2017; Green, 2007; Lakew and Bilotkach, 2018; Sheard, 2014, forthcoming), European regions (Albalate and Fageda, 2016; Bel and Fageda, 2008), or individual European countries, like Italy (Percoco, 2010; Cattane et al., 2016). A major shortcoming of these studies is that the economic effect of air services is related primarily to enhanced connections to specific destinations. Indeed, air transportation may have different economic effects on the host urban area or region. An expanding airport might increase numbers of employees, directly or indirectly. Additionally, air traffic can facilitate face-to-face contact, thus enhancing the delivery of tradable services, increasing tourism and knowledge flows and improving attractiveness for business investment. These latter effects can be captured most appropriately by referring to bilateral air traffic data.

Accordingly, we reviewed a number of recent studies using bilateral air traffic data that are more pertinent to this paper. Giroud (2012) uses data from the United States' manufacturing sector to show that new airline routes bring about an increase in investments within firms and an increase in a plant's total factor productivity, while Hovhannisyan and Keller (2015) report a significant positive impact of international business travel from the United States on patenting rates. Fageda (2017),

4

using data on traffic from Barcelona to the rest of the world, shows that the availability of non-stop flights offered at a sufficient frequency increases the amount of foreign direct investment. Finally, Alderighi and Gaggero (2017) show that air transport services positively affect the export business of Italian manufacturers.

Previous studies using bilateral air traffic data analyze the causal relationship between a given measure of the economic links between cities (investments, trade, patenting rates) and the supply of non-stop flights. Furthermore, previous studies focus on the United States or on an individual European country or urban area. Note also that studies using international air travel data are based on country level information.

We use the available data on the total number of passengers for inter-city links, including those passengers that made one stop to reach their final destination. In this regard, city-pair data of investments, trade, employment, patents and so on is not available for a sample that includes cities of different countries. This explains that all studies about the impact of air transportation on some of these specific measures of economic links between cities are made for one single country. For a sample that include cities from several countries, bilateral air traffic –including connecting traffic- is the best available proxy of such economic links.

Indeed, we focus on intercontinental inter-city markets where aviation is the only feasible means to connect these cities, and connecting flights are viable substitutes for non-stop flights. Hence, we can use these data to arrive at a global measure of the economic and social links between the considered cities, including: i) business trips, ii) trips for leisure or visiting friends and relatives, which account for tourism and immigration flows, and iii) the trade of high value products, in relation to their weight, as they are usually transported on commercial passenger flights.[3]

We find that even modest increases of the number of long-haul non-stop flights may lead to significant traffic gains. Thus, the availability of long-haul non-stop flights imply not only a

---

[3] Of course, some flights are just for cargo but they represent a low proportion of the total activity of European airports and are usually dedicated to the heaviest products that can be moved by plane. The amount of cargo on commercial passenger flights is particularly relevant in long-haul flights as they are typically operated with big aircrafts.

redistribution of traffic from connecting to non-stop traffic, but it also generate new traffic. This result holds for a sample of city-pairs with high levels of indirect connectivity.

However, we also find that the impact of more long-haul non-stop flights on total traffic flows is only clearly relevant when the destination city is located in a high-income country where the proportion of passengers sensitive to time costs should be higher. This is consistent with the hypothesis that the traffic gains of long-haul non-stop flights are due to a reduction in time costs. This is a novel result in relation to previous literature about the economic impact on non-stop flights as we investigate the differential effect that they may have in destinations with different levels of income.

Results of our analysis complement those obtained by Bilotkach (2015) and Cristea and Danila (2017). These two studies, using data at the airport/city level in the United States obtain evidence of a strong economic impact derived from a greater availability of non-stop flights. Our analysis suggests a certain limitation to the positive economic impact of non-stop flights as they are only relevant where income levels are already relatively high.

The rest of this paper is organized as follows. In the next section, we explain the empirical equation that we estimate and discuss some econometric issues that we must address. Then, we detail the sample used and provide relevant information about the data used in the empirical analysis. In the following two sections, we show the results of the econometric estimates. First, we examine the impact of long-haul non-stop flights on traffic flows. Second, we examine the differential impact of long-haul non-stop flights according to the income levels of the destination. The last section is devoted to our concluding remarks.

1. Empirical Strategy

Our empirical model is based on a gravity model, where air passenger flows between two cities depend on the economic and demographic size of each city and on the distance between them (Zhang, Lin and Zhang, 2017). However, the effect of distance (along with other time-invariant unobserved variables) is captured by city-pair fixed effects. Taking this into account, the essential

variable in our analysis – an indicator for the provision of long-haul flights – may capture a reduction in the time spent on the trip and a more convenient service. So, we estimate the following model for the city pair k in year y and quarter q:

$Log(Total\_traffic)_{kyq} = \alpha + \beta_1 Non\text{-}stop\_flights_{kyq} + \beta_2 Flights\_to\_European\text{-}hubs_{kyq} + \beta_3 Log(Pop\_origin)_{ky} + \beta_4 Log(Pop\_destination)_{ky} + \beta_5 Log(GNIpc\_origin)_{ky} + \beta_6 Log(GNIpc\_destination)_{ky} + \gamma'_k + \eta'_y + \nu'_q + \varepsilon_{kt},$ (1)

where the dependent variable is the total number of passengers, including non-stop and connecting traffic. As the main explanatory variable, we include the total number of non-stop flights. As controls, we include the population and gross national income of both origin and destination cities. Data for population refer to the urban area while data for income is at the country level. Data are for travel in only one direction within a city-pair market being a European city as the origin and a non-European city as the destination. While origin income affects demand for travel out of the origin city, roughly half the measured one-way traffic in a city-pair market is for passengers returning home to the destination city (which is their origin). Hence destination income should matter as well for measured traffic. The same argument applies to population, so both origin and destination populations will affect the demand for travel out of the origin city. Furthermore, we include a variable of flights of the origin city to the major European hubs as a partial indicator of indirect connectivity.

The data used present a panel structure so we employ the techniques typically applied within the framework of panel data models. A clear advantage of the city-pair fixed effects model is that it allows us to control for omitted variables that are correlated with the variables of interest and which do not change over time (Verbeek, 2000). The city-pair fixed effects model focuses on the within variation in the data so that it controls for the effect of time-invariant variables, such as, the distance covered in the non-stop link or historical links between countries. Furthermore, we add year dummies to control for yearly effects that are common to all city-pairs and quarter dummies to control for seasonal effects.

All variables are expressed in logarithms except the variables for non-stop flights that have many observations with 0 values. For the total number of passengers, 0.1% of observations have a 0 value so that the use of logarithms entails losing very few observations (93). In gravity models, it is usual to express the continuous variables in logs, so we prefer to use logs, as far as possible.

The estimates may present heteroscedasticity and temporal and cross-sectional autocorrelation problems. We apply the Breusch-Pagan/Cook-Weisberg test for heteroscedasticity and the Wooldridge test for autocorrelation in panel data. Both tests show that we may have problems of heteroscedasticity and of autocorrelation, which must be addressed. Hence, the standard errors are robust to heteroscedasticity. Following Bertrand et al. (2004), we allow for an arbitrary variance-covariance structure by computing the standard errors in clusters by city pair to correct for autocorrelation in the error term both at the cross-sectional and temporal levels.

As we mention above, the main variable of the analysis is that for the non-stop flights. We may expect that the total number of passengers will be higher in city pairs with non-stop air services. The variable of non-stop flights should capture the lower time costs and the more convenient service in travelling between the origin and destination cities.

Hence, the availability of non-stop flights should imply a redistribution of traffic from connecting to non-stop traffic, but it may also generate new traffic. That is, some potential travelers that did not fly before may now use the direct flight option. Thus, we may have both a qualitative and a quantitative effect. Recall that travelers in long-haul city-pairs do not have as option alternative transportation modes.

An important econometric challenge is the potential simultaneous determination of demand and supply. To deal with this problem, we use different techniques to identify properly the effect of nonstop flights on traffic.

We apply the logic of differences in differences with matching that is a common methodology employed within the treatment evaluation framework (see Angrist and Picke, 2009; Gertlet et al, 2016 for details). In this regard, one of the main goals of this study is to examine whether long-haul

8

non-stop flights may have a positive effect on traffic flows. Hence, we have an outcome variable (traffic) and a treatment variable (having non-stop flights). Furthermore, we have a panel dataset that includes city-pairs without non-stop flights (control city-pairs) and city-pairs with non-stop flights (treated city-pairs). In this regard, we take advantage of a large sample that may include similar treated and control city-pairs. Note that we focus on treated city-pairs with a change in their treatment status in the considered period because we drop those city-pairs with air services in all periods. These latter routes are clearly denser than the rest of routes in our sample. Finally, we estimate a city-pair fixed effects model that focuses on the within variation of the data so that we identify changes on traffic flows due to new non-stop services in treated city-pairs in comparison to traffic flows in control city-pairs.

To identify the causal effect of the treatment status on outcomes we need to compare comparable city-pairs. We may expect that the treatment is strongly correlated with the explanatory variables and with the previous levels of the outcome variable. Indeed, those city-pairs with higher levels of traffic and higher values of the covariates (population, income, flights to European hubs) will have a higher probability of having non-stop flights. This may produce an upward (selection) bias in the estimated impact of non-stop flights given that the treatment status is far from being random.

Thus, we apply a propensity score matching procedure and re-estimate Equation 1 with the observations that have common support. Matching procedures eliminate possible bias by pairing observations in the treated city-pairs (with non-stop flights) with control city-pairs (without non-stop flights flights) having similar characteristics. In the matching analysis therefore, we pair observations that have similar levels of traffic and similar values of all the covariates in the initial year of the considered period. Following Rosenbaum and Rubin (1983), we first estimate the probability of being treated, conditional on the pre-existing characteristics that differ between groups using a logit model, to obtain a propensity score for each observation. In a second step, we match the observations in the treated and control groups with respect to the propensity score, using the first nearest neighbor algorithm. This algorithm matches treated observations with the control

observations that have the closest propensity score. Then, we drop all the observations without common support and re-estimate Equation 1.

The implementation of the propensity score matching to produce the matching sample allow us to have a sample with treated and control observations similar in terms of the covariates and traffic levels in the baseline period. As we mention above, in the matching analysis we exclude those city-pairs with non-stop flights in all periods. These city-pairs may distort the identification of the effect of non-stop flights as they may be city-pairs with particularly high levels of demand and profitability. In this regard, city-pairs with particularly high levels of profitability may impose an upward bias in the identification of the effect of non-stop flights.

With the matching sample, the treated status should be random in terms of the observable factors. Hence, we are correcting the bias related with the fact that treated city-pairs will be more likely to be those city-pairs that connect richer and more populated cities, city-pairs with higher frequencies to European hubs and city-pairs with higher levels of traffic. The expected bias should be positive as observable factors may influence positively on traffic and non-stop flights. Hence, the matching procedure should reduce the estimated impact of non-stop flights on traffic.

However, a limitation of the propensity score matching is that it does not control for unobservable factors that may be correlated both with outcome and treatment status. To this point, the use of city-pair fixed effects allow us to control for time-invariant unobservable factors. Furthermore, we add time fixed effects to control for unobserved shocks that are common to all city-pairs.

With the matching sample that compares comparable city-pairs and the inclusion of city-pair and time fixed effects, the only source of bias are unobserved shocks that take place in the period considered and that may influence simultaneously traffic and non-stop flights. This could be an unobserved shock that increases route traffic like an increased tourist attractiveness of one of the endpoints or a stronger degree of economic cooperation between countries. Or it could be an unexpected shock that increase the number of non-stop flights like the expansion of capacity in a congested airport that provide services in one of the endpoints of the city-pair. This potential bias

should be modest as the considered period is relative short and the share of flights that an intercontinental route may have in an any airport is really very small. In any case, we also apply an instrumental variables procedure using lags of the endogenous explanatory variable to mitigate this potential (positive) bias.

2. Data

We restrict our sample of cities to allow us to build a feasible dataset with sufficient variation in the supply of non-stop flights. We focus on cities that are able to generate a high amount of traffic to many intercontinental destinations, excluding those cities with the biggest airports that have always had high levels of intercontinental connectivity.

Hence, our sample consists of European cities with mid-sized airports. We include cities with airports having more than 10 million passengers, excluding the main hubs. By main hubs, we mean airports with more than 50 million passengers (Amsterdam, Frankfurt, Istanbul, London, Madrid, Moscow, Paris, Rome). We also exclude tourist destinations like islands in the south of Europe, as our focus is on inter-city links. Overall, we have 32 cities in the sample, which generally conform to one of two descriptions: large cities within big countries (e.g. Ankara, Barcelona, Lyon, Manchester, Milan, Munich, Saint Petersburg), or the main cities of small countries (e.g. Athens, Brussels, Budapest, Dublin, Lisbon, Stockholm). The European cities in our sample have more than one million inhabitants, with the exception of Dusseldorf, Edinburgh, Geneva and Stuttgart, which have more than five hundred thousand inhabitants. Note that differences in the quality of passenger transportation networks across major cities in Europe are mainly related to the availability of direct intercontinental flights, as those cities are generally well serviced through a dense network of highways, trains and short-haul flights (Bel and Fageda, 2008).

We include non-European cities hosting the largest airports in different geographical areas; Africa, East Asia, South Asia, Latin America and the Caribbean, the Middle East and North America. We do not consider cities in nearby areas like North Africa or Central Asia because our focus is on long-haul flights. From the Middle East, we only include Abu Dhabi, Doha and Dubai,

11

given their leading position in air traffic flows between Europe and Asia. However, we exclude those links to Middle-East cities with less than 4,000 kilometers. Furthermore, we do not consider cities from Oceania because non-stop flights to Europe were not technically viable in the considered period. Overall, the 74 non-European cities each have more than one million inhabitants (except Doha) and most of them have more than two million inhabitants.

Table 1 indicates the cities included in our sample. Some of these cities may have more than one airport, so in those cases we aggregate the information of all their airports to arrive at the usable data for the inter-city link. Note that our sample represents a very high proportion of total traffic in long-haul city-pairs departing from European cities with mid-size airports.

Table 1. Sample of cities

Origin (Europe)	Destination (rest of the world)
Ankara, Athens, Barcelona, Berlin, Birmingham, Brussels, Bucharest, Budapest, Cologne, Copenhagen, Dublin, Dusseldorf, Edinburgh, Izmir, Geneva, Glasgow, Hamburg, Helsinki, Lisbon, Lyon, Manchester, Milan, Munich, Oslo, Prague, Saint Petersburg, Sofia, Stockholm, Stuttgart, Vienna, Warsaw, Zurich	Atlanta, Abu Dhabi, Bangkok, Bangalore, Baltimore, Belo Horizonte, Beijing, Bogotá, Bombay, Boston, Buenos Aires, Caracas, Chennai, Chicago, Dallas, Dhaka, Denver, Detroit, Doha, Dubai, Guangzhou, Guayaquil, Hanoi, Ho Chi Minh City, Havana, Hong Kong, Houston, Islamabad, Jakarta, Johannesburg, Karachi, Kinshasa, Kolkata, Kuala Lumpur, Lagos, Las Vegas, Los Angeles, Lima, Luanda, Manila, Memphis, Mexico City (DF), Miami, Minneapolis, Montreal, Nairobi, Nagoya, Nanjing, New Delhi, New York, New Orleans, Orlando, Osaka, Philadelphia, Phoenix, Rio de Janeiro, Salvador de Bahia, San Diego, San José (CR), San Francisco, Santa Cruz (BO), Santiago, Santo Domingo, Sao Paulo, Seattle, Shenzhen, Seoul, Shanghai, Singapore, Taipei, Vancouver, Washington, Tokyo, Toronto

We use the information at carrier-route level provided by the Official Airlines Guide (OAG) and RDC Aviation, on a quarterly basis, from 2010 to 2016. The OAG provides Marketing Information Data Tapes (MIDT) through a Traffic Analyzer. The information contained in the OAG dataset includes the total number of passengers for city-pair markets all over the world. RDC aviation provides information about the number of non-stop flights for city-pair markets all over the world. Thus, we use OAG data for our dependent variables and RDC Aviation data for the variables of non-stop flights.

The traffic data include non-stop passengers who fly directly between the two cities and connecting passengers who stop at one airport (or more) to reach their final destination. Connecting passengers include transfers through a European or a non-European airport. For example, on the

city-pair Barcelona-Atlanta, we may have passengers flying from Barcelona to Atlanta via Amsterdam (or other European cities) or passengers flying Barcelona to Atlanta via Miami (or other US cities). The number of connecting passengers in long-haul links may be high. Even with the availability of non-stop services, connecting passengers may still represent an important proportion of the total air market.

All the data used refer to a one-way journey. In this regard, the city-pair markets in our sample only consider one direction. The origin is a European airport and the destination is a Non-European airport.

Obviously, there is a strong correlation between the total number of passengers and the total number of flights across city-pair markets. In fact, this is the main challenge of the econometric analysis. However, looking at 2016, for illustrative purposes, we may find thin city-pairs with non-stop flights and thick city-pairs without non-stop flights. Indeed, some city-pairs with fewer than 2,000 passengers (including non-stop and connecting passengers) have non-stop flights, for example, Brussels-Luanda, Budapest-Doha, Edinburgh-Doha, Lisbon-Caracas or Zurich-Nairobi. In contrast, if we look at the city-pair markets within the densest 10% of our sample (having more than 5,800 passengers), we find that 53 out of 219 city-pairs do not have direct flights. This is the case with, for example, Dusseldorf-Shanghai, Milan-Manila, Geneva-Bangkok, Vienna-Tokyo, Budapest-New York, Barcelona-Mexico City, Berlin-Bangkok, Barcelona-Tokyo, Manchester-Bangkok and Barcelona-Seoul.

The data on population reflect the urban area only and the information comes from the United Nations (World Urbanization Prospects), while for income, the data represent the national situation and are compiled by the World Bank (World Development Indicators).

We also consider a variable of flights of the origin city to the major European hubs. We have traffic data for non-stop and connecting passengers but our supply data is just for non-stop flights. This makes it difficult to consider a variable that completely measures the supply of connecting flights. Having said this, our variable of indirect connectivity refers to the sum of the non-stop flights

from the cities of origin to the major European airports that are hubs of a large network airline (London-LHR, Paris-CDG, Frankfurt-FRA, Madrid-MAD, Amsterdam-AMS, Rome-FCO, Istanbul-IST and Moscow-SVO). This variable is built as the sum of the flights from the origin city (which is a European city) to each European hub airport weighted by the share of flights of such hub to the destination city (which is a non-European city). The share is based on the proportion of total flights of each hub over the total flights of all major European hubs to the destination cities. Although this is a partial indicator of the levels of indirect connectivity, the vast majority of destinations in our sample have non-stop flights to several European hub airports. Note that this variable may also suffer from an endogeneity bias as it is the case of the main variable of our analysis. However, we may expect that such bias is modest as it could only come from unobserved shocks that affect the origin city and it is not a goal of this study to examine the impact of the variable of non-stop flights to major European hub airports on traffic flows.

We have 2,250 routes (city pairs) that represent about 60,000 observations as we have data for 28 time periods (quarterly data from 2010 to 2016). 191 of these routes operate non-stop services in all periods, and 37 in discontinued periods. [4] Furthermore, we have 116 new routes with non-stop air services. By new routes we mean that scheduled non-stop air services have been initiated after the first quarter of 2010. The rest of the potential city-pair markets did not have non-stop air services throughout the observed period but they may still have a lot of (connecting) traffic. By scheduled non-stop air services, we mean provision of at least one flight per week. Overall, 13% of all observations involves city-pairs with scheduled non-stop air services.

Note that many airlines provide non-stop flights in several routes in our sample. Among the European airlines, we have Air Lingus, Air Berlin, Alitalia, Austrian, SN Brussels, Czech airlines, Finnair, Jet2, LOT, Lufthansa, Meridiana, Norwegian, SAS, SATA, Swiss, TAP, Thomas Cook and Transaero. Among the non-European airlines, we have Air Canada, Air China, Air India, Air

---

[4] Discontinued routes have non-stop air services in the initial period but later the services are cancelled for some or all periods. Some of these routes have seasonal services; routes that have non-stop flights in just some seasons of the year (usually the spring and summer).

Transat, American Airlines, ANA, Cathay, Continental, Delta, Emirates, Etihad, FlyDubai, Hainan airlines, JAL, Jet Airways, Korean air, Pakistan International Airlines, Qatar Airways, Singapore Airlines, Thai Airways, United, US Airways and Virgin Atlantic Airways

Table 2 shows the descriptive statistics of the variables used in the empirical analysis. One remarkable fact reported in this table is that more than half of total traffic is, on average, connecting traffic. Note also that the mean value for the variable of non-stop flights imply that, on average, a little bit less of one flight per week is offered (recall that data is quarterly). However, this value is so low because a high number of observations have zero values when looking at the number of non-stop flights. If we focus just on routes with non-stop flights in at least some quarter of the considered period, such mean value increases up to almost one daily flight. Thus, it is also remarkable that, on average, new non-stop flights are provided with high frequency. So the change may be big; from no services at all to a daily service.

Table 2. Descriptive statistics of the variables used in the empirical analysis

Variable	Mean	St. Deviation	Minimum value	Maximum value
Total traffic (passengers per city-pair per quarter)	2026.92	4154.34	0	86593
Non-stop traffic (passengers per city-pair per quarter)	728.72	3135.17	0	72516
Connecting traffic (passengers per city-pair per quarter)	1298.19	1806.90	0	25077
Non-stop flights (number of flights per city-pair per quarter)	10.67	38.37	0	639
Flights to European hubs (number of flights per city-pair per quarter)	511.18	344.66	0	4735
Pop_origin (thousands of inhabitants per year)	1953.83	1207.26	479.85	5308.6
Pop_destination (thousands of inhabitants per year)	8313.02	6987.74	528.50	38139.6
GDPpc_origin (US dollars per year)	40179.13	19140.96	8297.48	89818.3
GDPpc_destination (US dollars per year)	25934.07	22176.14	318.08	72671

3. The impact on non-stop flights on the connectivity of cities

3.1. Marginal change in the number of non-stop flights

As we have mentioned above, the first step in our empirical strategy is to implement the propensity score matching to produce the matching sample that allow us to have a sample with treated and control observations similar in terms of the covariates and traffic levels in the baseline period. Table 3 reports the mean t-test differences between treated city-pairs (city-pairs having non-stop flights) and control city-pairs (city-pairs not having non-stop flights) when we consider the full sample and when we consider the matching sample in the baseline period. In the matching sample, the treated and control city-pairs have similar values for all the variables at the initial year of the considered period. Moreover, we cannot reject that they are equal in most of the variables. In this regard, the matching sample only uses about 10 percent of the total number of observations that we have available. Thus, data in table 3 confirm that we have comparable treated and control city-pairs in terms of observable characteristics in the matching sample. [5]

Table 3. Mean t-test for variables used in the matching procedure in 2010

Sample	All sample			Matching sample		
	Mean values			Mean values		
	Control city-pairs (mean values)	Treated city-pairs mean values)	T-test (differences in mean values)	Control city-pairs (mean values)	Treated city-pairs (mean values)	T-test (differences in mean values)
Total traffic (passengers per city-pair per quarter)	1597.30	2904	-7.67***	2791.92	2821.09	-0.17
Flights_to_European_hubs (number of flights per city-pair per quarter)	527.11	663.20	-7.82***	686.83	669.22	0.78
Pop_origin (thousands of inhabitants per year)	1911.67	1795.12	2.06**	1895.52	1780.94	1.50
Pop_destination (thousands of inhabitants per year)	7770.08	7289.32	1.49	5842.46	7341.37	-3.34***
GNIpc_origin (US dollars per year)	38930.37	4292.59	-4.36***	42748.77	43003.73	-0.22
GNIpc_destination (US dollars per year)	24790.75	37667.89	-12.37***	34668	37656.36	-2.24*
Number observations	8513			888		

---

[5] The propensity score matching method identifies treatment and control routes with similar probabilities being selected in a treatment (conditioned on the observable characteristics). Data in table 3 reveals that this lead to similar treated and control routes.

However, as we mention above, matching techniques do not overcome the potential bias related to differences in participation due to the heterogeneity in unobservable factors. To overcome this problem we use two different techniques. First, we include city-pair fixed effects in the estimation to control for differences between city-pairs due to unobservable factors that are time invariant and time fixed effects to control for unobserved shocks that are common to all city-pairs. Second, we apply an instrumental variables procedure to the matching sample to overcome the possible bias in participation due to heterogeneity that is unobservable and changing over time. In this regard, we may expect that the simultaneous bias is upward because there is a positive relationship between non-stop flights and total traffic.

It is difficult to find good instruments as potential factors affecting non-stop flights may have also an influence on total traffic. We use a standard approach to deal with the endogeneity bias that is using lags of the endogenous explanatory variable.[6] We have supply data up to 2005 so that the use of lags does not imply to reduce the sample size. Given that we use quarterly data, one-year lag implies using four lags, two-year lags implies using eight lags and so on. As we might expect, the impact of non-stop flights on traffic is lower when we use one-year or two-year lags with respect to the estimate that uses contemporary data of the supply variable. Therefore, the lags approach reduces the upward bias derived from the simultaneous determination of traffic and non-stop flights. We prefer the use of two-year lags rather than one-year lag as instrument. In this regard, the errors are serially correlated so that a potential drawback of using lagged values of the endogenous explanatory variable is the correlation with the lagged errors. Hence, the use of two-year lags instead of one-year

---

[6] We have also run regressions using two related subsets of instruments; i) the total number of intercontinental flights from the origin city and the total number of flights from the destination city to European airports - excluding the flights of the corresponding city-pair- and ii) all flights from the origin and destination cities - excluding the flights of the corresponding city-pair. However, results with these two sets of instruments are not satisfactory because they do not correct the upward bias. Indeed, the use of these instruments lead to a stronger impact of non-stop flights than that obtained with the variable not instrumented.

lag of the endogenous explanatory variable as instrument may help to mitigate the bias associated with this correlation.

Table 4 shows the results of two regressions. In the first regression, we consider the full sample and the contemporaneous values of the variable for non-stop flights. The second regression consider the matching sample and two-year lags as instrument of the variable for non-stop flights. In all regressions, we include all the covariates, city-pair and time fixed effects.

In order to examine the impact of the number of non-stop flights on total traffic, we look at the elasticities (evaluated at sample means) that show the increase of total traffic given a one percent increase of non-stop flights. However, changes in frequencies (if any) are always greater than one percent. Therefore, the use of elasticities that show the marginal effect of the change in frequencies on traffic does not provide the most accurate measure of what the actual effect of these changes may be. Thus, we also look at the increase of traffic with a one standard deviation increase of non-stop flights in relation to the mean. Note that this implies a modest increase of non-stop flights. For both samples (the full and matching samples), a one standard deviation increase of non-stop flights implies the shift from one to almost three weekly flights. In practice, the change once an airline chooses to offer non-stop services may be much higher especially when we depart from a scenario with no direct flights.

Results in table 4 show a significant increase of traffic when the number of non-stop flights is higher. When we consider the full sample and the contemporaneous values of non-stop flights, a one percent increase in the number of non-stop flights lead on average to about a 0.064 per cent increase of traffic. We also find that a one standard deviation increase in non-stop flights lead to about 23 percent increase of traffic. When we consider the matching sample and two-year lags of the non-stop flights variable as instrument, which is our preferred regression, a one percent increase in the number of non-stop flights lead on average to about a 0.036 per cent increase of traffic. Furthermore, a one standard deviation increase in non-stop flights lead to about 8 percent increase of traffic.

Thus, we find that even modest increases of non-stop flights may lead to significant traffic gains. This result holds for a sample of city-pairs with high levels of indirect connectivity. This is the first main result of this paper. Furthermore, our identification strategy corrects the expected upward bias in the estimated impact on non-stop flights on traffic. Indeed, the use of the propensity score matching and instrumental variables lead to a much lower estimated impact of non-stop flights on traffic flows. The magnitude of the estimated impact of non-stop flights on traffic flows when we use the matching sample and two year lags as instrument is two-three times lower than the estimated impact when we use the full sample and contemporaneous values of the non-stop flights variables.

Table 4. Estimation results (demand equation)

Sample	All	Matching
Instrument	**None**	**Two-year lags**
		Results first-stage
Instrument	-	0.12 (0.004)***
R²	-	0.29
		Results second-stage
Non-stop_flights	0.0061	0.0025
	(0.001)***	(0.0006)***
R²	0.22	0.15
No. observations	59968	6104

Notes: We include all controls, city-pair, year and quarter fixed effects. Standard errors in parentheses (robust to heteroscedasticity and clustered by city-pair). Statistical significance at 1% (***), 5% (**), 10% (*).

3.2. Openings and closures of non-stop flights

The interpretation of the results in table 4 focuses on a marginal change in the number of non-stop flights. While that the identification of the impact of a marginal change in the number of non-stop flights on total passenger flows is the main focus of our analysis, we also examine the impact based on the variation in openings and closures of non-stop flights. Here we consider that a city-pair have scheduled non-stop flights if at least one non-stop flight per week is offered in such city-pair.

In this regard, we report the results of two additional regressions in table 5. In the first regression, we consider a dummy variable for openings of non-stop flights as the main variable of interest and the matching sample is used. The second regression considers a dummy for closures of non-stop flights as the main variable of interest. We cannot apply a matching procedure in this latter regression

as the control city-pairs would be city-pairs with non-stop flights in all periods. Recall that we exclude these city-pairs in the analysis because they may impose an upward bias in the identification of the effect of non-stop flights. Furthermore, the matching sample would be very small. Hence, we follow a before-and-after approach and the sample is restricted to city-pairs with closures during the considered period.

In all these regressions, we drop those city-pairs with seasonal services. By seasonal services, we mean city-pairs that have non-stop flights in just some seasons of the year (usually the spring and summer). Note that a potential drawback in these regressions is that we cannot use lags as instrument given that the variable of interest is a dummy.

In any case, results of these additional regressions shows that the impact of non-stop flights on total traffic is much higher than in the approach based on the marginal change. The impact of openings lead to about a 60 per cent increase in traffic, while the closures lead to about a 40 per cent decrease in traffic. While our data reveal that the impact of non-stop flights on traffic flows may be substantial, we prefer to be cautious in the formulation of the first of the main results of this paper; even modest increases of non-stop flights may lead to significant traffic gains. Indeed, our focus on the marginal change is to derive a conservative measure of the impact of non-stop flights on traffic flows.

Table 5. Estimation results (demand equation – openings and closures)

Sample	Matching	City-pairs with closures in the considered period
$D^{openings}$	0.60 (0.01)***	-
$D^{closures}$	-	-0.39 (0.04)***
R^2	0.44	0.34
No. observations	6104	504

Notes: We include all controls, city-pair, year and quarter fixed effects. Standard errors in parentheses (robust to heteroscedasticity and clustered by city-pair). Statistical significance at 1% (***), 5% (**), 10% (*).

4. The differential impact of non-stop flights in destinations with different levels of income

As we mention above, the second goal of the study is to examine whether the impact of non-stop flights on traffic flows is associated to the income levels of the destination. Recall here that our one-

way data only consider one direction. The origin is a European airport and the destination is a Non-European airport. Hence, the analysis related to the differential effects of non-stop flights according to income levels is focused on the destination. Our sample of destination cities include rich, middle-income and developing countries. In contrast, differences in income between the origin cities are less marked.

We use two different strategies to investigate the role of the destination income on the impact of non-stop flights on traffic. First, we add an interaction between the variable of non-stop flights and the GNI per capita of the country of destination. Second, we run regressions of the traffic equation for subsamples with richer and poorer destination countries. We make the distinction between "richer" and "poorer" by considering destination countries with income levels higher and lower than the mean sample, respectively. To correct the potential bias in the analysis of the relationship between non-stop flights and traffic flows, we use the same methodologies as in the previous section: a propensity score matching that produces the matching sample, city-pair and time fixed effects, and two year lags as instrument for the non-stop flights variable.

The first column in table 6 shows the results of the traffic equation including the interaction variables. We find clear evidence of a stronger effect of non-stop flights on total traffic flows for richer destination countries. Indeed, the interaction variable of non-stop flights and GNI per capita of the country is positive and statistically significant and, in fact, the interaction variable eliminates the effect of the non-stop flights variable without the interaction. This provides some evidence that the effect of non-stop flights is concentrated in city-pairs of high-income levels. However, it should be noted that the correlation between the three interaction variables could distort the individual identification of each variable, so we also analyze the role of the income of the destination country with the additional regressions based on subsamples.

Table 6. Estimation results (demand equation –differences by income of the destination)

Sample	Matching (all)	Matching (high-income)	Matching (low-income)
	Results first-stage		
Instrument	0.0097 (0.001)***	0.12 (0.0005)***	0.06 (0.007)***

R²	0.89	0.34	0.20
	Results second-stage		
Non-stop_flights	-0.06 (0.001)***	0.003 (0.0006)***	-0.003 (0.003)
Non-stop fligths X GNIpc_destination	1.42e-06 (3.10e-07)***	-	-
R²	0.91	0.44	0.28
No. observations	6104	4508	1596

Notes: We include all controls, city-pair, year and quarter fixed effects. Standard errors in parentheses (robust to heteroscedasticity and clustered by city-pair). Statistical significance at 1% (***), 5% (**), 10% (*).We use the matching sample and two-year lags as instrument.

The second and third column in table 6 show the results of these additional regressions. We find clear evidence that the impact of non-stop flights on total traffic is only relevant when we consider city-pairs that have destination cities located in relatively rich countries. Indeed, the coefficient of the non-stop flights variable is only positive and statistically significant in the sub-sample that is based on city-pairs with relatively rich destination countries. This is consistent with the idea that the main mechanism that explain the traffic increase associated with non-stop flights is the reduction in time costs. Indeed, high-income travelers may have a higher willingness to pay for non-stop flights to save time and have a more convenient service. In contrast, the incentives to fly of low-income travelers may be more related with the fares that they have to pay. This suggests a limitation to the positive economic impact of non-stop long-haul flights as they are only relevant where income levels are already relatively high.

In the previous section, we obtained the first main result of this paper. Even modest increases of non-stop flights may lead to significant traffic gains. In this section, we have obtained the second main result. Such traffic gains seem to be concentrated in city-pairs where the destination is located in a high-income country.

5. Concluding remarks

Our results suggest that even modest increases of the number of non-stop flights lead to significant traffic gains. However, the impact of more non-stop flights on total traffic is only clearly relevant when the destination city is located in a high-income country. Hence, such traffic gains could be explained by a quality effect related with a reduction in time costs.

Our results have been obtained in a context where connecting flights could be potentially competitive in relation to non-stop flights, given that the sample considered includes as origin European cities that are well connected with the main hubs and as destination distant cities.

In spite of this substantial impact on traffic, our data show that many inter-city markets are not connected by non-stop flights, although they have close relationships, as suggested by their dense traffic levels. This suggests that the interests of airlines may not coincide with those of cities. Airlines may prefer to operate in fewer airports to save on costs, although it may also be the case that the airline market is imperfect in the sense that some profitable flight options are not being offered.

Given the positive impact of long-haul non-stop flights on total passenger flows, it may be advisable to promote airport policies to attract airlines that are willing to offer them. The tools available for airport managers include lowering charges, increasing capacity, offering preferential allocation of the best gates and check-in desks, or engaging in marketing initiatives with the cities' promotional agencies. Furthermore, at an international level, the promotion of open skies policies may help in procuring non-stop long-haul flights beyond the main hubs. Europe has had just such an open skies agreement with the United States and Canada since 2008, and also with some nearby countries like Morocco, Israel and Georgia. Results of our analysis suggest that it could be advisable to expand liberalization policies to other countries.

Our analysis also suggests potential negative externalities associated with the concentration of traffic in main hubs, aside from the well-known externality associated with congestion. Such negative externality manifests in the loss of direct connectivity between mid-sized cities. However, this does not mean that our paper justifies public intervention to promote de-hubbing strategies by airlines. A project for future research could be to investigate this issue.

Finally, the results of our analysis suggest that the positive impact of non-stop flights on total passenger flows between distant cities is concentrated in countries with high income levels. This

suggests a certain limitation to the role of air transport in the economic growth of cities insofar as their effects are only relevant where income levels are already relatively high.

References

Albalate, D., Fageda, X. (2016). High-tech employment and transportation: Evidence from the European regions, *Regional Studies*, 50(9), 1564-1578.

Alderighi, M., Gaggero, A. (2017). Fly and trade: Evidence from the Italian manufacturing industry. *Economics of Transportation*, 9, 51-60.

Angrist, J. D. , Pischke, J.S (2009). *Mostly Harmless Econometrics*. New York; Princeton University Press

Bel, G., Fageda, X. (2008). Getting there fast, Globalization, intercontinental flights and location of headquarters, *Journal of Economic Geography* 8, 471-495.

Bertrand, M., E. Duflo, Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119, 249-275.

Berry, S., Jia, P. (2010). Tracing the Woes, an Empirical Analysis of the Airline Industry, *American Economic Journal: Microeconomics*, v. 2(3), pp. 1–43

Bilotkach, V. (2015). Are Airports Engines of Economic Development? A Dynamic Panel Data Approach", *Urban Studies* 52, 1577-1593.

Blonigen, B., Cristea, A. (2015). Air Service and Urban Growth: Evidence from a Quasi-Natural Policy Experiment. *Journal of Urban Economics* 86, 128-146.

Brueckner, J.K. (2003). Airline traffic and urban economic development, *Urban Studies* 40, 1455-1469.

Brueckner, J.K. (2002). Airport Congestion When Carriers Have Market Power," *American Economic Review* 92, 1357-1375.

Brueckner, J.K., Spiller, P.T. (1994). Economies of traffic density in the deregulated airline industry. *Journal of Law and Economics* 37, 379-415.

Button, K., Taylor, S. (2010). International air transportation and economic development. *Journal of Air Transport Management* 6, 209-222.

Cattaneo, M., Malighetti, P., Paleari, S., Redondi, R. (2016). The role of the air transport service in interregional long-distance students' mobility in Italy. *Transportation Research Part A: Policy and Practice* 93, 66-82.

Cristea, A., Danila, L. (2017), *The role of aviation networks for urban developments*, Working Paper.

Fageda, X. (2017). International air travel and foreign direct investments: Evidence from Barcelona. *Journal of Regional Science*, 57, 858-883.

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B., Vermeersch, C.M.J. (2016). *Impact Evaluation in Practice*. Washington, DC: Inter-American Development Bank and World Bank.

Giroud, X. (2013). Proximity and investment: Evidence from plant-level data, *Quarterly Journal of Economics* 128, 861-915.

Green, R.K. (2007). Airports and Economic Development, *Real Estate Economics* 35, 91–112.

Hovhannisyan, N., Keller, W. (2015). International business travel: an engine of innovation?, *Journal of Economic Growth* 20, 75-104.

Lakew. P.A, Bilotkach V. (2018). Airport delays and metropolitan employment. *Journal of Regional Science* 58, 424–450.

Mayer, C., Sinai, T. (2003). Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays Are Evil, *American Economic Review*, 93 (4), 1194-1215.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70, 41-55.

Sheard, N. (2014). Airports and urban sectoral employment, *Journal of Urban Economics* 80, 133-152.

Sheard, N. (in press). Airport Size and Urban Growth, *Economica*, in press.

Percoco, M. (2010). Airport Activity and Local Development, Evidence from Italy, *Urban Studies* 47, 2427, 2443.

Verbeek, M. (2000). *A guide to modern econometrics*. West Sussex; John Wiley and Sons.

Zhang, Y., F. Lin and A. Zhang (2017), "Gravity models in air transport research: A survey and an application," in: Bruce A. Blonigen and Wesley W. Wilson (eds.), Handbook of International Trade and Transportation, Edward Elgar.