Journal of Air Transport Management 72 (2018) 32–38

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

When the mall is in the airport: Measuring the effect of the airport mall on passengers' consumer behavior

José I. Castillo-Manzano, Lourdes López-Valpuesta*, Antonio Sánchez-Braza

Universidad de Sevilla, Departamento de Análisis Económico y Economía Política, Facultad de Ciencias de Económicas y Empresariales, Avda. Ramón y Cajal, 1, 41018 Sevilla, Spain

ARTICLE INFO

Keywords:
Airport management
Non-aeronautical revenue
Airport shopping mall
Statistical causal inference
Propensity score matching

ABSTRACT

This paper provides a new approach to evaluating the influence on passenger F&B consumption and expenditure of terminals that approximate to the concept of an airport shopping mall. Using a broad database of 37,226 passengers interviewed at eight different Spanish airports, including two Spanish hub airports, Madrid-Barajas and Barcelona, with a methodology framed within statistical causal inference with Kernel and Radial matching, the results robustly demonstrate that passengers alter their consumption behavior in hub airport malls compared to how they behave at regional airports with a smaller commercial and F&B offer. Specifically, there is an increase of between 3.7 and 4.1% in the likelihood that hub passengers will make a consumption and between 1.2 and 1.3% in the likelihood that they will make a purchase, while mean per-passenger spending increases by 3.53€.

1. Introduction

A great deal of literature currently exists on the growing importance of non-aeronautical revenue in airport management (Del Chiappa et al., 2016; Fasone et al., 2016; Yokomi et al., 2017). A large number of papers address the possible determinants that would explain passengers' consumer behavior at airports. According to these papers, their behavior could be explained using variables that range from waiting time (see the debate surrounding the importance or lack of importance of this variable in Chung et al., 2013); passenger characteristics, from income level (Castillo-Manzano, 2010) to age (Graham, 2008) and gender (Geuens et al., 2004); trip characteristics, including whether the passenger's motive for flying is leisure or business (Lu, 2014) or, for example, whether s/he is flying on a domestic or international flight (Fuerst et al., 2011) or even the type of airline (Gillen and Lall, 2004 or Castillo-Manzano and López-Valpuesta, 2015).

The present paper seeks to complement the prior literature by using a new approach to evaluate whether the fact that the terminal provides a broad and varied commercial and F&B offer has any effect on passengers' spending levels and the likelihood that they will make purchases at airport stores and establishments. The aim is to highlight the role of certain terminals as generators of non-aeronautical revenue and their transformation into *de facto* airport shopping malls (Appold and Kasarda, 2006; Geuens et al., 2004) with a high concentration, high volume and wide range of retail and F&B establishments. Given the space requirements and high passenger volumes needed to be cost-effective, airport shopping malls are typically located inside the terminal at larger airports and especially at hubs. Hubs are airport network nodes where traffic from several origins can be consolidated and distributed to a diverse range of final destinations (Button, 2002), thus enabling airlines to improve connectivity and increase the number of markets that can be served.

Strategic changes to ensure the viability of many airports (Freathy and O'Connell, 2012) and the success of shopping malls in terminals (Appold and Kasarda, 2006) have led to many airport operators considering enlarging the area devoted to shopping and F&B facilities. We focus on two Spanish airport hubs that have carried out multi-million Euro refurbishment programs, specifically Madrid-Barajas, with approx. €6200 m, and Barcelona, with over €3000 m (see Castillo-Manzano et al., 2015; Castillo-Manzano et al., 2017; Dobruszkes et al., 2017). Barcelona even advertises its new terminal's shopping area with the name of shopping center (http://www.aena.es/es/aeropuerto-barcelona/todas-tiendas.html). These two airports have greatly increased their commercial offer by giving passengers the choice of a vast array of stores carrying the main international brands that are very unlike traditional convenience stores, as well as a large food court with an assortment on offer that ranges from fast food, including McDonalds, Burger King and similar, to thematic restaurants with different degrees of sophistication and price.

Thus, this paper is structured as follows: after this introduction,

* Corresponding author.
E-mail addresses: jignacio@us.es (J.I. Castillo-Manzano), lolopez@us.es (L. López-Valpuesta), asb@us.es (A. Sánchez-Braza).

https://doi.org/10.1016/j.jairtraman.2018.07.003
Received 25 April 2018; Received in revised form 20 July 2018; Accepted 21 July 2018
Available online 02 August 2018
0969-6997/

Table 1
Main characteristics of the sample.

Airport		Almeria	Alicante	Barcelona-El Prat	Madrid-Barajas	Santiago	Seville	Tenerife Sur	Valencia
Information gathering	Questionnaire	Available in five languages		Available in six languages	Available in five languages	Available in six languages	Available in five languages		
Commercial offer	Total no. stores (stores in passenger-only areas)	1 (1)	18 (16)	58 (46)	68 (60)	11 (3)	15 (13)	17 (13)	4 (3)
	Total no. restaurants (restaurants in passenger-only areas)	3 (1)	17 (12)	44 (34)	64 (51)	3 (2)	6 (3)	13 (8)	9 (4)
	Total sq.meters, stores (sq. meters airside stores)	305 (305)	2157 (2092)	10941 (9501)	19260 (18234)	723 (159)	1366 (1216)	2832 (2467)	743 (611)
	Total sq.meters, restaurants (sq. meters airside restaurants)	1084 (548)	2362 (1571)	12773 (7178)	20866 (18234)	1305 (335)	7165 (5852)	3863 (1525)	2566 (667)
Sampling	General information	Departing passengers (> 15 years of age) in departure lounges							
	Sample size	1808	3202	6931	9096	3530	6027	3092	3540
	Sampling method	Stratified by traffic segments in which a selection of flights was made for each route, and a group of passengers was selected by means of systematic sampling.							
	Sampling error	± 2,1%	± 1,7%	± 1,2%	± 1,0%	± 1,6%	± 1,2%	± 1,8%	± 1,7%

Section 2 describes the data. Section 3 details the methodology used. Section 4 presents the main empirical outcomes while Section 5 offers conclusions.

2. Data

We use a broad database collected through personal interviews conducted by the Spanish Public-Private (51% vs. 49%) Airport Authority (AENA) during summer 2010. Our research uses a database of 37,226 passengers who were interviewed in departure lounges at eight different Spanish airports (namely Almeria, Alicante, Barcelona, Madrid, Santiago, Seville, Tenerife Sur and Valencia). Table 1 gives the main characteristics of the sample. As can be seen, the greatest care was taken to obtain a broad simple random sample.

Table 1 also shows the great differences that exist between the retail offer at hub and regional airports, with respect to both floor space allocated to commercial areas and number of establishments, both stores and F&B. For example, the number of F&B establishments ranges from 3 at Santiago and Almería airports to 44 at Barcelona and 64 at Madrid-Barajas.

These data make it easy to conclude that what can be found in the case of both Madrid and Barcelona airports is a shopping area that clearly imitates real shopping malls. It should therefore come as no surprise that some new terminology has been used on official maps of Barcelona airport's new terminal; rather than the traditional terminology applied to maps at all other Spanish airports, which only distinguishes between three areas for passengers (Public zone, Passenger-only zone and Boarding area), a fourth has been included to define the area in the terminal where the retail stores are concentrated, called "shopping center".

We focus on 20 different variables that were available in their entirety for 36,271 passengers: one indicator variable, 3 dependents and 16 explanatory variables or covariates that were identified in the academic literature as factors that may be major determinants of airport retail demand.

To be specific, the group of most important variables includes passenger Socio-demographics, with age and gender (Castillo-Manzano, 2010; Chung et al., 2013; Freathy and O'Connell, 2012; Geuens et al., 2004; Lin and Chen, 2013; Lu, 2014) and nationality and place of residence (Freathy and O'Connell, 2012; Geuens et al., 2004; Lin and Chen, 2013) standing out. In this category, income plays a major role in determining passengers' airport shopping intentions (Chung et al., 2013; Lin and Chen, 2013; Lu, 2014). As this is a variable that is not easy to obtain in personal interviews, proxies are usually used, such as education level (education), work status (in work) and means of travel to the airport (taxi), as used in Castillo-Manzano (2010). The passenger's travel behavior or frequency of travel is a factor that should be taken into account (Chung et al., 2013; Freathy and O'Connell, 2012; Geuens et al., 2004; Lin and Chen, 2013; Lu, 2014), and can also be interpreted as an indicator of the passenger's level of income.

Focusing on trip characteristics, other factors that can also affect the passenger's purchase decision and volume of purchases have been highlighted by previous research, including:

a) The reason for traveling, whether leisure, business or VFR (Visiting Friends & Relatives) (Appold and Kasarda, 2006; Chung et al., 2013; Freathy and O'Connell, 2012; Fuerst et al., 2011; Lin and Chen, 2013; Lu, 2014; Torres et al., 2005)
b) In general terms, the duration of the trip, as it is more likely that F&B will be consumed before a long flight (Freathy and O'Connell (2012); Appold and Kasarda (2006) and, in particular, the type of route (Fasone et al., 2016; Fuerst et al., 2011). As very short trips of 0–1 day are clearly correlated with short distances, while connecting passengers at a hub usually fly to more distant destinations, three control variables have been included that help to correct for

these factors, specifically: Non-Eurozone international destination; Connecting flight and Trip Duration.

c) Type of airline, with special emphasis on LCCs (Fasone et al., 2016; Freathy and O'Connell, 2012; Lei and Papatheodorou, 2010; Yokomi et al., 2017).
d) Time spent in the airport (Chung et al., 2013; Torres et al., 2005) including ticket processing time, the time needed to check in and clear security, and free time available to make purchases, as highlighted by Appold and Kasarda (2006). Some studies have measured this time by the punctuality and delays of airlines using the airport (Fuerst et al., 2011).

Lastly, the prior literature highlights the growing importance of the passenger's social interaction factors, and considers these to be even more important than many of the previous factors. These include travel company size (Chung et al., 2013; Lin and Chen, 2013), which is captured by a number of different variables, such as whether the passenger is accompanied on his/her trip by friends, family members or work colleagues (Freathy and O'Connell, 2012; Lu, 2014). Whether the passenger is being accompanied by children is especially important (Chung et al., 2013; Freathy and O'Connell, 2012) as is whether family members/friends have traveled to the airport to see the passenger off (Castillo-Manzano, 2010).

These 16 explanatory variables allow us to analyze the factors that define the profile of passengers and travel features.

All of these variables are presented in Table 2 along with their main descriptive statistics.

With respect to binary variables, the mean value indicates the percentage of individuals in the sample set who meet the indicated characteristic. For example, the explanatory variable "gender" is a binary variable defined as 1 if male and 0 if female, so the 0.528 mean value indicates that 52.8% of the sample are men while the remaining 47.2% are consequently women. With respect to categorical variables, which are ordered categories in our case, the mean value can be useful as an approximate indicator of sample distribution among the various considered ordered categories. So, if we take the explanatory variable "age" as an example, as the mean is 2.015 and there are 4 categories, this would indicate that the sample is slightly biased toward the two population groups under 49 years old.

3. Methodology

The proposed methodology is framed within statistical causal inference (Dawid, 2000; Pearl, 2000). The "Rubin causal model" (Rubin, 1974) as it was initially developed and the subsequent contributions made by Holland (1986) have been used as the starting point for developing this model. Statistical causal inference methods allow researchers to estimate the causal effect induced by a fact that is to be evaluated (the cause) on one or more variables of interest (the effect). This methodology enables the attainment of consistent estimators of the effect induced by the considered fact by controlling for the possible influence of other factors on the variables of interest. Thus, the purpose of this methodology is to isolate the effect of this fact on the variable or variables of interest by maintaining control over other factors that may affect these variables. If the conditions are not the same then the effects cannot be attributed exclusively to the cause, hence the relevance of adequately controlling for these possible contaminating variables.

Causal inference-based techniques are currently widely used in multiple scientific disciplines, ranging from medicine (Hirano and Imbens, 2001), where they were originally developed in medical experimentation, to different areas of the Social Sciences, such as the political sciences (Imai, 2005); sociology (Morgan and Harding, 2006); and the economic evaluation of public policies (Sánchez-Braza and Pablo-Romero, 2014), to cite but a few examples. Recently, the application of this methodology has spread further to include the economic evaluation of transportation policy-related actions and behaviors (see, for example, Canavan et al., 2015; Whitehead et al., 2015); and user-related behaviors regarding different modes of transport (see Oliveira et al., 2015 or Castillo-Manzano et al., 2016; among others).

Initially, a participation variable (indicator variable) must be defined as a binary indicator of participation in the fact to be evaluated. Thus, starting with a sample of size N, the binary variable D {0, 1} is defined ("hub" variable) that captures whether the observation corresponds to a passenger at a hub airport ($D_i = 1$) or at a regional airport ($D_i = 0$). Sample observations are thus divided into n_1 (treatment group) and n_0 (control group).

Next, the response variables are defined (Y). These are the variables on which the causal effect of the analyzed fact will be evaluated. In this case, three response variables are defined, as explained in Table 2: "expenditure at airport", "consumes food/drink at the airport" and "purchase at airport". These variables are specified in terms of potential results.

$$Y_i \begin{cases} Y_{1i} & if \quad D_i = 1 \\ Y_{0i} & if \quad D_i = 0 \end{cases} \tag{1}$$

The evaluated fact's causal effect is captured by defining the "*Average Treatment Effect on the Treated" (ATET)* as the difference between the mean values of the response variable for passengers corresponding to hub and non hub airports, conditioned on the treatment group.

$$ATET = E(Y_1 - Y_0|D = 1) = E(Y_1|D = 1) - E(Y_0|D = 1) \tag{2}$$

Nevertheless, the quality of the average effects could be diminished if individuals of both the treatment and control groups differ in characteristics other than those arising from participation in the evaluated fact. So, the contaminant effects of other variables that could impact on the effect have to be controlled for in order to obtain the *ATET*. A k-dimensional vector formed of a set of covariates therefore has been created according to specifications in Table 2 (16 explanatory variables). The covariates must be independent of variable D for each and every one of the observations. The condition of independence should therefore be guaranteed that ensures that variable D, which is conditioned on these predetermined variables, is independent of the potential results:

$$D \perp (Y_1, Y_0)|X \tag{3}$$

Thus the average effect will be evaluated of the fact that the airport is a hub on the likelihood of in-airport expenditure, consumption and purchase, conditioned on the possible values of the vector of the covariates, X. This evaluation procedure will follow a two stage process (Hahn et al., 2011; Heckman and Vytlacil, 2005).

However, the existence of a large number of covariates might make a comparison of treated and control individuals a very complicated process. For this reason, the *propensity score* is deployed. The first stage is now to calculate the so-called *propensity score*, $\varepsilon(X)$, defined by Rosenbaum and Rubin (1983) as the likelihood that an observation of the sample belongs to the evaluated fact's treatment group (in this case, individuals from hub airports), conditioned on the values taken by an X vector of predetermined covariates.

$$\varepsilon(X) = P(D = 1|X = x) = E[D|X = x] \tag{4}$$

The calculation of the *propensity score* thus streamlines the comparison process by reducing a large number of covariates to a single, one-dimensional variable (see also Imbens and Wooldridge, 2009). The aforementioned condition of independence is therefore formulated as:

$$D \perp (Y_1, Y_0)|\varepsilon(X) \tag{5}$$

Different binary response models can be specified to estimate the *propensity score* depending on the hypothesis adopted regarding the form of its distribution function (F):

$$\varepsilon(X) = P(D = 1|X) = F(\beta X) \tag{6}$$

Table 2
Description of variables and descriptive statistics.

Variable	Explanation		No. obs.	Mean	Stand. dev.
Indicator variable: *D*					
Hub	1 if a hub airport (Madrid or Barcelona)		15284	0.421	0.494
	0 otherwise		20987		
Response variables: *Y*					
Expenditure at airport	Euros spent by passengers at stores and catering establishments		36271	8.493	19.001
Consumes food/drink at the airport	1 if the passenger consumes food/drink		17074	0.471	0.499
	0 otherwise		19197		
Purchase at airport	1 if the passenger makes a purchase		8564	0.236	0.425
	0 otherwise		27707		
Explanatory variables: *X*					
Socio-demographic characteristics of the passenger	Gender	1 if male	19145	0.528	0.499
		0 if female	17126		
	Age	1 < 30	9986	2.015	0.827
		2 = 31-49	17882		
		3 = 50-64	6268		
		4 > 65	2135		
	Spanish	1 if passenger is Spanish	20622	0.569	0.495
		0 if passenger is foreign	15649		
	Education	1 = no formal or only primary education	2876	2.541	0.638
		2 = completed secondary education	10903		
		3 = holds university degree	22492		
	In work	1 if the passenger is in some form of work (self-employed or salaried)	26281	0.725	0.447
		0 otherwise	9990		
	Taxi	1 if passenger has traveled to the airport by taxi	8666	0.239	0.426
		0 otherwise	27605		
	Frequent passenger	Number of flights taken by passenger in previous twelve months:		2.396	0.960
		1 = no flights	7016		
		2 = 1 to 3 flights	13264		
		3 = 4 to 12 flights	10616		
		4 = more than 12 flights	5375		
Travel features	Vacation	1 if trip is for vacation	17464	0.471	0.499
		0 otherwise	18807		
	Duration of trip	1 = 0–1 days	2790	2.521	0.987
		2 = 2–7 days	20255		
		3 = 8–14 days	6642		
		4 = 15–30 days	4705		
		5 > 30 days	1879		
	LCC(Low-cost carrier)	1 if passenger is flying on an LCC	16464	0.454	0.498
		0 otherwise	19807		
	Non-Eurozone international destination	1 if final destination outside the Eurozone	3943	0.109	0.311
		0 otherwise	32328		
	Connecting flight	1 if not a direct flight (the passenger must transfer to at least one additional connecting flight to reach his/her final destination)	4160	0.115	0.319
		0 otherwise	32111		
	Waiting time prior to boarding	1 < 1 hour	1510	2.812	0.862
		2 = 1–2 hours	12999		
		3 = 2–3 hours	12578		
		4 > 3 hours	9184		
Social interaction	Accompanied	1 if passenger has travel companion/s	17804	0.491	0.500
		0 otherwise	18467		
	Children	1 if passenger is flying with children	2946	0.081	0.273
		0 otherwise	33325		
	Seen off	1 if passenger has been seen off at the airport	6809	0.188	0.391
		0 otherwise	29462		

The two most commonly used are the probit and logit models (Cameron and Trivedi, 2009). The probit model considers that F is the standard normal cumulative distribution function:

$$P(D_i = 1|X) = \int_{-\infty}^{X_i'\beta} \left(\frac{1}{\sqrt{2\pi}}\right)^{\frac{-z^2}{2}} dz \quad \text{for } -\infty < z < \infty \tag{7}$$

While in the logit model, F is specified as the cumulative distribution function of the logistic distribution:

$$P(D_i = 1|X) = \frac{e^{X_i'\beta}}{1 + e^{X_i'\beta}} \tag{8}$$

The regression parameters β can be obtained by maximizing the log-likelihood function (Cameron and Trivedi, 2005):

$$L_N(\beta) = \sum_{i=1}^{N} D_i \ln[F(X_i'\beta)] + (1 - D_i)\ln[1 - F(X_i'\beta)] \tag{9}$$

There are no defined selection criteria for choosing one or other model to estimate the *propensity score*. Hence, the selection is made merely for operational reasons. After estimating both models, the model that maximizes the value of the likelihood function is chosen (Wooldridge, 2002).

Subsequently, each individual in the participant group with a specific value of $\varepsilon(X)$ is assigned one (or various) individual/s from the control group with a value that equates to, or approaches, $\varepsilon(X)$. The distribution of the X vector of covariates is thus similar for the two groups. In this way, any possible contamination from the covariates is isolated and matching provides an unbiased estimate of the effect of the analyzed fact.

The *propensity score matching* technique is subsequently used to calculate the estimator using the expression:

$$\hat{\alpha}_{ATET} = \frac{1}{n_1}\sum_{i=1}^{n_1}(Y_i - Y_{m(i)}) \tag{10}$$

where $Y_{m(i)}$ is the value of the response variable Y for the control individual assigned as the pair of the participating individual i.

Radial and kernel matching methods are used for the assignment process. The first of these assigns participating individuals a weighted average *propensity score* for the control individuals within a certain bandwidth (b). The weighting term (w) is defined as follows, where the function k (·) is a function of the kernel.

$$w_{ij} = \frac{k\left(\frac{\varepsilon_{i1}-\varepsilon_{j0}}{b}\right)}{\sum_{j=1[j\in(D=0)]}^{n_0} k\left(\frac{\varepsilon_{i1}-\varepsilon_{j0}}{b}\right)} \tag{11}$$

On the other hand, the radial matching method establishes a radius (r) that enables each participant with a certain *propensity score* (ε_{i1}) to be assigned all the control individuals with a *propensity score* (ε_{j0}) within the radius formed by ε_{i1} and r. Thus the pairing condition for the radial method is expressed as:

$$c_{ij} = \{\varepsilon_{j0} \mid \ \|\varepsilon_{i1} - \varepsilon_{j0}\| \ < r\} \tag{12}$$

where c_{ij} indicates the control individual that meets the pairing condition for individual i.

4. Results

First, the *propensity score* is estimated. This allows individuals in the treatment group (hub airports) and the control group (regional airports) to be homogenized with the covariates for the two groups now being comparable. In this case, the variable used as the dependent variable in the estimation of the logit and probit models is the "hub" variable (the indicator variable of participation in the fact to be evaluated). Table 3 summarizes the obtained results from the estimations of the propensity score in the context of the 16 explanatory variables in Table 2. Estimations have been made using both the logit and probit models. In this case, we opted for a logit specification as it maximized the log pseudo-likelihood (−16253531 v. probit −16264596). The corresponding results are given in Table 3. Column (a) includes the obtained results when the whole sample is used for the estimation. Next, propensity score values are assigned according to the specifications of this model. Column (b) includes the obtained results but the total sample is restricted to passengers who really do make a purchase and consume F&B (when "expenditure at the airport" variable > 0).

The resulting coefficients indicate the degree to which each of the 16 considered covariates contributes to the *propensity* score. When a covariate is not significant this implies that there are no significant differences in the attribute representing the variable between individuals in the treatment and control groups. However, when the covariate is shown to be significant, this means that differences in the attribute do exist between the individuals in the two groups. The use of the *propensity score* specifically enables these significant differences to be controlled for, so that individuals in the two groups, treatment and control, can ultimately be compared with respect to the attribute set used. However, the significance of each of the individual covariates is not important for our analysis. As explained above, the purpose of the *propensity score* is solely to make treatment group individuals as homogeneous as possible as far as the 16 covariates are concerned.

Next, the *propensity score matching* estimators are calculated. This is done using both the radial and kernel *matching* methods. Three degrees of radius/bandwidth (0.05, 0.10 and 0.15) are used to test the sensitivity of the obtained estimators to changes in the level of proximity required in terms of *propensity score* between the individuals in the participant and control groups. Tables 4, 5 and 6 show the obtained results for each of the considered response variables.

The obtained results demonstrate that mean expenditure per passenger at the Madrid and Barcelona hub airports is significantly higher than at other airports. To be specific, according to Table 4, mean expenditure at hub airports exceeds that at regional airports by between 3.52€ and 3.54€ on average, calculated using kernel matching and

Table 3
Estimation of the propensity score using the logit model (Dependent variable: Hub).

Variable	Logit Model (whole sample) (a)	Logit Model (subsample: Expenditure > 0€) (b)
Gender	−0.049	−0.074*
	(0.032)	(0.042)
Age	−0.153***	−0.159***
	(0.019)	(0.026)
Spanish	0.264***	0.358***
	(0.033)	(0.043)
Education	0.564***	0.577***
	(0.027)	(0.037)
In work	0.037	−0.035
	(0.038)	(0.051)
Taxi	0.080**	0.008
	(0.038)	(0.051)
Frequent passenger	0.067***	0.109***
	(0.019)	(0.026)
Vacation	−0.062*	−0.062
	(0.036)	(0.049)
Duration	0.016	−0.016
	(0.017)	(0.023)
LCC	0.559***	0.588***
	(0.019)	(0.027)
Non-Eurozone	1.828***	1.988***
	(0.071)	(0.094)
Connecting flight	−1.464***	−1.335***
	(0.066)	(0.086)
Waiting time	−1.337***	−1.462***
	(0.033)	(0.044)
Accompanied	0.269***	0.290***
	(0.037)	(0.050)
Children	−0.852***	−0.900***
	(0.062)	(0.078)
Seen off	−0.167***	−0.170***
	(0.043)	(0.059)
Constant	−0.855***	−0.857***
	(0.129)	(0.178)
No. obs.	36271	20678
Max. log-likelihood	−16253531	−8557974.9
Pseudo-R^2	0.164	0.185

Robust standard deviation is given in brackets.
One, two and three asterisks indicate 10%, 5%, and 1% significance, respectively.

Table 4
Estimators for the radial and kernel *matching* methods for response variable "Expenditure at the airport".

Response variable "Expenditure at the airport"		$\hat{\alpha}_{ATET}$	Std. dev	Lik.
Radial matching				
Radius	0.05	3.534***	0.236	0.000
	0.10	3.559***	0.231	0.000
	0.15	3.518***	0.228	0.000
Kernel matching				
Bandwidth	0.05	3.407***	0.304	0.000
	0.10	3.570***	0.180	0.000
	0.15	3.577***	0.220	0.000

*** Three asterisks indicate 1% significance.

Table 5
Estimators for the radial and kernel *matching* methods for response variable "Consumes food/drink at the airport".

Response variable "Consumes food/drink at the airport"				
		$\hat{\alpha}_{ATET}$	Std. dev.	Lik.
Radial matching				
Radius	0.05	0.041***	0.006	0.000
	0.10	0.042***	0.006	0.000
	0.15	0.039***	0.006	0.000
Kernel matching				
Bandwidth	0.05	0.030***	0.008	0.000
	0.10	0.040***	0.007	0.000
	0.15	0.041***	0.008	0.000

*** Three asterisks indicate 1% significance.

Table 6
Estimators for the radial and kernel *matching* methods for response variable "Purchase at the airport".

Response variable "Purchase at the airport"				
		$\hat{\alpha}_{ATET}$	Std. dev.	Lik.
Radial matching				
Radius	0.05	0.013***	0.005	0.000
	0.10	0.013***	0.005	0.000
	0.15	0.010***	0.005	0.000
Kernel matching				
Bandwidth	0.05	0.013***	0.003	0.000
	0.10	0.014***	0.005	0.000
	0.15	0.013***	0.006	0.000

*** Three asterisks indicate 1% significance.

radial matching respectively. These values come from the average of the results obtained for each method with the application of the three degrees of radius/bandwidth considered. In addition, Fasone et al. (2016) consider that a greater availability of total surface area for commercial activities (as found at the analyzed hubs) seems to impact positively on average expenditure per passenger. One explanation for this greater spending at hub airports might lie in the Fuerst et al. (2011) assertion that high-quality outlets such as designer boutiques are typically only found at larger airports, or Graham's (2009) observation that larger airports like hubs tend to have more international (and especially intercontinental) passengers, who spend more.

It can be concluded from Tables 5 and 6 that an airport mall increases the likelihood of food or drink being consumed at F&B establishments and of a purchase being made at an airport store. To be precise, the former rises by between 3.7% (mean value obtained by the kernel matching method) and 4.1% (mean value obtained using the radial matching method). The latter, the likelihood of making a purchase at an airport store, increases between 1.3% and 1.2% depending on whether the mean value considered is calculated using kernel or radial matching. In this sense, Appold and Kasarda (2006) analyzed two factors that influence sales per passenger, both of which are found at the two hub airports considered in this paper. On the one hand, the amount of retail space provided at an airport has a positive influence on sales per passenger, which is in line with the results here. However, for these authors the number of passengers at an airport would have a negative effect, due to congestion discouraging sales. Thus, as commented in the introduction, the increase in the number of hub passengers should be accompanied by an increase in the size of the airport's commercial area to offset what Fasone et al. (2016) refer to as the "overcrowding effect" on individual spending that characterizes larger sized airports.

This greater likelihood of making a purchase and consuming food and/or drink would explain the previously-mentioned greater mean expenditure of 3.52€-3.54€. However, the increase in mean

Table 7
Estimators for the radial and kernel *matching* methods for response variable "Expenditure at the airport" (subsample Expenditure > 0).

Response variable "Expenditure at the airport"				
		$\hat{\alpha}_{ATET}$	Std. dev.	Lik.
Radial matching				
Radius	0.05	5.400***	0.382	0.000
	0.10	5.442***	0.371	0.000
	0.15	5.411***	0.364	0.000
Kernel matching				
Bandwidth	0.05	5.450***	0.407	0.000
	0.10	5.440***	0.328	0.000
	0.15	5.418***	0.421	0.000

*** Three asterisks indicate 1% significance.

expenditure could also be due to an average increase in the mean expenditure of the people who are making the purchase and/or consuming F&B.

We repeated the whole procedure in two stages to test this last supposition, but restricting the total sample only to passengers who really do make a purchase and a consume F&B, i.e., to passengers who have a value greater than 0 in the "*Expenditure at the airport*" variable. The *propensity score* was estimated once again, but this time only considering the above-stated subsample. Estimations were made using both the logit and probit models. Once again we opted for a logit specification as it maximized the log pseudo-likelihood (−8557974.9 v. probit −8558790.8). Results for the model are given in Table 3, column (b). Table 7, below, gives the results of estimator matching for this subsample.

According to the results in Table 7, passengers that make a purchase and/or consume F&B at the airport (Expenditure at the airport > 0€) effectively spend more at the hubs. The mean expenditure of "real" consumers at hub airports is between 5.40€ and 5.45€ more than at regional airports.

5. Conclusions

Using six regional airports and two main hubs in the Spanish airport system, irrespective of the matching method used this paper provides clear, robust empirical proof, significant at 1%, that passengers change their consumer behavior at bars and stores in malls at hub airports compared to how they would behave at regional airports, where the commercial and F&B offers are clearly more limited.

Specifically, once corrected for a set of 16 standard explanatory variables justified by prior analyses (that range from waiting time to journey motivation and duration), it can be observed in our analysis that the greater commercial offer of airport malls increases the likelihood that passengers make a consumption at an F&B establishment by 3.7–4.1%, and a purchase at an airport store by 1.2–1.3%. Furthermore, passengers who really do make an expenditure, increase their spending by some 5.40€ at hubs over regional airports. The final mean increase in expenditure/per passenger made by all passengers at hub airports generated by this set of effects is approximately 3.53€.

The increase in passenger expenditure is even more relevant when it is taken into account that not only do the hubs offer a wider range of commercial outlets, but also cheaper consumption choices (e.g., fast food chains such as McDonalds and Burger King, see Castillo-Manzano and López-Valpuesta (2013), and the main low cost clothes stores, such as H&M, Mango and, especially, some of brands belonging to the enormous INDITEX multinational company), compared to the limited and, generally, expensive offer at regional airports, which is usually limited to an occasional overpriced souvenir shop or the odd bar or cafeteria franchise.

In other regards, these robust results clearly support the strategy of developing large shopping centers at hubs, which accumulate high volumes of potential customers. Specifically, in the Spanish case, Madrid's 53.4 million and Barcelona's 47.2 million passengers in 2017 demonstrate a distinct opportunity for maximizing non-aeronautical revenues.

The success of these shopping centers represents a significant source of non-aeronautical revenue for hubs that enables them to charge lower airport dues and so attract new connections and new airlines, especially those that, theoretically, are most sensitive to airport charges, such as the low cost companies. It is no surprise, therefore, that, in the wake of the refurbishments that vastly increased their commercial offer, it was, precisely, the Spanish hubs that quickly became benchmark international hub airports where all the main European low-cost airlines operate, including Ryanair.

Pressure on airport management to lower airport charges and thus maximize non-aeronautical revenue cannot only come from the airlines. For example, in the Spanish case, acting on a request from the National Commission for Markets and Competition (CNMC) the Government has opted for year-on-year reductions to airport dues (and will reduce them by a further 20% during the 2018–2020 period) rather than the price freeze called for by AENA.

Finally, as this study has been developed with a database of 37,226 passengers interviewed at eight different Spanish airports, one possible future line of research would be to test the robustness of these results in other airport systems to confirm whether consumer behavior differs with a change in geographic circumstances. In this regard, the fact that a large part of the sample, over 44% (specifically, 15649 passengers), are foreigners makes us quite optimistic about the likelihood that the results can be extrapolated to other airport systems, within Europe, at least.

Acknowledgments

The authors would like to express their gratitude to the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (ECO2015-64996-P MINECO/FEDER) for their support.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jairtraman.2018.07.003.

References

Appold, S.J., Kasarda, J.D., 2006. The appropriate scale of US airport retail activities. J. Air Transport. Manag. 12, 277–287.

Button, K., 2002. Debunking some common myths about airport hubs. J. Air Transport. Manag. 8, 177–188.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometric. Methods and Applications. Cambridge University Press, New York.

Cameron, A.C., Trivedi, P.K., 2009. Microeconometrics Using Stata. Stata Press, College Station, TX.

Canavan, S., Graham, D.J., Melo, P.C., Anderson, R.J., Barron, A.S., Cohen, J.M., 2015. Impacts of moving-block signaling on technical efficiency: application of Propensity score matching on urban metro rail systems. Transport. Res. Rec.: J. Transport. Res. Board (2534), 68–74.

Castillo-Manzano, J.I., 2010. Determinants of commercial revenues at airports: lessons learned from Spanish regional airport. Tourism Manag. 31 (6), 788–796.

Castillo-Manzano, J.I., López-Valpuesta, L., 2013. Analysing passenger behaviour towards the catering industry: implications for airport management. Int. J. Hosp. Manag. 35, 258–260.

Castillo-Manzano, J.I., López-Valpuesta, L., 2015. Who is the charter passenger? Characteristics and attitudes of the least known passenger. Tourism Econ. 21, 1079–1085.

Castillo-Manzano, J.I., Pozo-Barajas, R., Trapero, J.R., 2015. Measuring the substitution effects between High Speed Rail and air transport in Spain. J. Transp. Geogr. 43, 59–65.

Castillo-Manzano, J.I., Castro-Nuño, M., López-Valpuesta, L., Pedregal, D.J., 2017. Measuring the LCC effect on charter airlines in the Spanish airport system. J. Air Transport. Manag. 65, 110–117.

Castillo-Manzano, J.I., López-Valpuesta, L., Sánchez-Braza, A., 2016. Going a long way? On your bike! Comparing the distances for which public bicycle sharing system and private bicycles are used. Appl. Geogr. 71, 95–105.

Chung, Y.-S., Wu, C.-L., Chiang, W.-E., 2013. Air passengers' shopping motivation and information seeking behavior. J. Air Transport. Manag. 27, 25–28.

Dawid, A.P., 2000. Causal inference without counterfactuals. J. Am. Stat. Assoc. 95 (2), 407–448.

Del Chiappa, G., Martin, J.C., Roman, C., 2016. Service quality of airports' food and beverage retailers. A fuzzy approach. J. Air Transport. Manag. 53, 105–113.

Dobruszkes, F., Givoni, M., Vowles, T., 2017. Hello major airports, goodbye regional airports? Recent changes in European and US low-cost airline airport choice. J. Air Transport. Manag. 59, 50–62.

Fasone, V., Kofler, L., Scuderi, R., 2016. Business performance of airports: non-aviation revenues and their determinants. J. Air Transport. Manag. 53, 35–45.

Freathy, P., O'Connell, F., 2012. Spending time, spending money: passenger segmentation in an international airport. Int. Rev. Retail Distrib. Consum. Res. 22 (4), 397–416.

Fuerst, F., Gross, S., Klose, U., 2011. The sky is the limit? The determinants and constraints of European airports commercial revenues. J. Air Transport. Manag. 17 (5), 278–283.

Geuens, M., Vantomme, D., Brengman, M., 2004. Developing a typology of airport shoppers. Tourism Manag. 25 (5), 615–622.

Gillen, D., Lall, A., 2004. Competitive advantage of low-cost carriers: some implications for airports. J. Air Transport. Manag. 10 (1), 41–50.

Graham, A., 2008. Managing Airports: an International Perspective, third ed. Butterworth Heinemann-Elsevier, Oxford.

Graham, A., 2009. How important are commercial revenues to today's airports? J. Air Transport. Manag. 15, 106–111.

Hahn, J., Hirano, K., Karlan, D., 2011. Adaptative experimental design using the propensity score. J. Bus. Econ. Stat. 29 (1), 96–108.

Heckman, J.J., Vytlacil, E., 2005. Structural Equations, treatment effects, and econometric policy evaluation. Econometrica 73 (3), 669–738.

Hirano, K., Imbens, G.W., 2001. Estimation of causal effects using propensity score weighting: an Application to data on right heart catheterization. Health Serv. Outcome Res. Meth. 2 (3–4), 259–278.

Holland, P.W., 1986. Statistics and causal inference (with discussion). J. Am. Stat. Assoc. 81 (396), 945–970.

Imai, K., 2005. Do get-out-the vote calls reduce turnout? The importance of statistical methods for field experiments. Am. Polit. Sci. Rev. 99 (2), 283–300.

Imbens, G.M., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. J. Econ. Lit. 47 (1), 5–86.

Lei, Z., Papatheodorou, A., 2010. Measuring the effect of low-cost carriers on regional airports' commercial revenue. Res. Transport. Econ. 26, 37–43.

Lin, Y.-H., Chen, C.-F., 2013. Passengers' shopping motivations and commercial activities at airports - the moderating effects of time pressure and impulse buying tendency. Tourism Manag. 36, 426–434.

Lu, J.-L., 2014. Investigating factors that influence passengers' shopping intentions at airports. Evidence from Taiwan. J. Air Transport. Manag. 35, 72–77.

Morgan, S.L., Harding, D.J., 2006. Matching estimators of causal effects: prospects and pitfalls in theory and practice. Socio. Meth. Res. 35 (1), 3–60.

Oliveira, R., Moura, K., Viana, J., Tigre, R., Sampaio, B., 2015. Commute duration and health: empirical evidence from Brazil. Transport. Res. Pol. Pract. 80, 62–75.

Pearl, J., 2000. Causality: Models, Reasoning and Inference Applications. Cambridge University Press, New York.

Rosenbaum, P.R., Rubin, D.B., 1983. The Central role of the propensity score in observational studies for causal effects. Biometrika 70 (1), 41–55.

Rubin, D.B., 1974. Estimating causal effects of treatments in randomized and non-randomized studies. J. Educ. Psychol. 66 (5), 688–701.

Sánchez-Braza, A., Pablo-Romero, M.P., 2014. Evaluation of property tax bonus to promote solar thermal systems in Andalusia (Spain). Energy Pol. 67, 832–843.

Torres, E., Domínguez, J.S., Valdés, L., Aza, R., 2005. Passenger waiting time in an airport and expenditure carried out in the commercial area. J. Air Transport. Manag. 11, 363–367.

Whitehead, J., Franklin, J.P., Washington, S., 2015. Transitioning to energy efficient vehicles: an analysis of the potential rebound effects and subsequent impact upon emissions. Transport. Res. Pol. Pract. 74, 250–267.

Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge (USA).

Yokomi, M., Wheat, P., Mizutani, J., 2017. The impact of low cost carriers on non-aeronautical revenues in airport: an empirical study of UK airports. J. Air Transport. Manag. 64, 77–85.