Entry games for the airline industry

Christian Bontemps, Bezerra Sampaio

Entry games for the airline industry

Christian Bontemps* Raquel Menezes Bezerra Sampaio †

February 25, 2018

Abstract

In this paper we review the literature on static entry games and show how they can be used to estimate the market structure of the airline industry. The econometrics challenges are presented, in particular the problem of multiple equilibria and some solutions used in the literature are exposed. We also show how these models, either in the complete information setting or in the incomplete information one, can be estimated from i.i.d. data on market presence and market characteristics. We illustrate it by estimating a static entry game with heterogeneous firms by Simulated Maximum Likelihood on European data for the year 2015.

Keywords: entry, airlines, multiple equilibria, estimation, industrial organization.

JEL codes:

1 Introduction[1]

Being able to evaluate the nature of competition between firms in a given sector and how it has changed over time or with a new economic environment is an important feature of the empirical IO literature. The use of the Herfindhal index has been proved (see Sutton, 1991) to be a poor proxy of the degree of competition. Ideally, one would like to get data where both supply and demand could be estimated. However, such data are often difficult to get, to the exception of the data provided by the US Department of Transportation for the US interior market. Entry games are therefore very popular in the empirical Industrial Organization (IO) literature because they allow to study

*Universiy of Toulouse, ENAC & Toulouse School of Economics. christian.bontemps@tse-fr.eu

†Universidade Federal do Rio Grande do Norte. raquelsampaio@ect.ufrn.br

[1]We thank Alex Pedurand for his research assistance. We also thank Mariane Bontemps for help in the editing process. The usual disclaimer applies.

the nature of competition between firms from data that are generally easy to collect. Observing how firms operate in different markets with different characteristics allows the empirical economist to estimate how these characteristics affect the profitability of each potential player and how the entry of an additional competitor harms everyone's profit.

In this chapter, we seek to explain how static entry games can be used to estimate the decision of a firm to operate or not in a market by endogeneizing everyone's decision. These models have been estimated in the empirical IO literature in many sectors including the airline industry (Bresnahan and Reiss, 1991a, 1991b, Berry, 1992, Mazzeo, 2002, Seim, 2006, Cleeren et al. 2009, Grieco, 2014, among plenty other contributions).[2]. Although the estimation of an entry model raises some econometrics issues, that are exposed in this chapter, the alternative option of estimating standard ordered choice models is a bad idea because it implicitly considers that the firms' decisions to operate or not are independent, an assumption ruled out by the data.

Additionally, most of the entry games are games with multiple equilibria unless some order of entry is assumed.[3] It creates potential identification issues and requires adapted econometric techniques to handle it. We present some of the solutions used in the literature and introduce briefly the recent advances on moment inequality models that permits to consider more general entry games.[4]. In this review, we focus on static games with simultaneous moves. Contrary to dynamic games, one should view an entry as a decision of being active in a market rather than entering /exiting a market in which there are incumbents and potential entrants. Estimating a dynamic game poses other econometrics' challenges which are out of the scope of this chapter.[5]

First, in Section 2, we present the literature on static entry games with complete information where all firms observe the profitability of all other competitors. We introduce in particular the specification of Berry (1992) for heterogeneous firms, which encompasses Bresnahan and Reiss (1991a) homogeneous version. Understanding this seminal paper is sufficient to understand the further developments of this literature. Some of them are introduced at the end of the section. Then, in Section 3, we extend our presentation to the case of models with incomplete information where firms decide to enter before observing everybody's profit shocks. In Section 4, we briefly review applied papers which have estimated entry games for the airline industry. Finally, in Section 5, we propose a short illustration of the potentialities of an entry game on European data collected for 2015 from the Official Airline Guide. Section 6 concludes.

[2]See Berry and Reiss (2007) for a survey.

[3]Assumptions often combined with additional minor assumptions.

[4]See de Paula, 2013, for a full survey

[5]Aguirregabiria and Mira (2010) and Bajari et al., 2013, propose a survey including dynamic games.

2 Entry games of complete information

In this section, we present the standard entry game with complete information and simultaneous move widely used in empirical Industrial Organization. In these type of models, firms, homogeneous or not, decide to be active in a given market if the profits collected after the entry decision are positive. Here, all the firms observe their profitability shocks as well as the ones of their competitors. However, the applied economist does not observe them and, usually, postulates a parametric distribution for these shocks. We focus on Berry (1992), which encompasses as a special case the homogeneous version of Bresnahan and Reiss (1991a). These two models are the basis of more complex models introduced later in the section. As in any structural or semi-structural economic model, the question of the identification of the profit function is at stake. A discussion about the identification of entry games can be found in Tamer (2003) and Berry and Tamer (2005).

2.1 The standard entry game with heterogeneous firms

Let N_p be the number of firms potentially active in market m. In the standard game, we assume that the profit of an active firm, labeled i, in a market with N^* active firms, is equal to

$$\pi_{i,m}(N^*) = \pi(X_m, Z_{i,m}; \theta) - h(N^*; \delta) + \varepsilon_{i,m}. \tag{1}$$

The usual normalization is to assume that a non-entrant firm gets a profit of 0. In Equation (1), X_m is a vector of market characteristics, $Z_{i,m}$ is a vector of firm i specific characteristics for market m, $\pi(X, Z; \theta)$, a parametric profit function known up to a vector of parameters θ. $h(N; \delta)$ is a strictly increasing in N, positive, parametric function that drives the degree of competition. It captures the property that more competitors harms profits.[6] Note that we can easily extend the model by allowing both $\pi(X, Z; \theta)$ and $h(N; \delta)$ to be firm specific, i.e. $\pi_i(X, Z; \theta)$ and $h_i(\delta, N)$ without changing any of the derivations made in this section. Finally, $\varepsilon_{i,m}$ represents a profit shock for firm i in market m. Profit shocks are correlated within each market. These shocks are observed by all firms who could potentially enter (complete information) and who decide, simultaneously, to enter or not.

In Berry (1992), the profit $\pi(X_m, Z_{i,m}; \theta)$ is linear in the explanatory variables, i.e. $\pi(X_m, Z_{i,m}; \theta) = X_m^\top \beta + Z_{i,m}^\top \gamma$. The market characteristics, X_m, are the product of the population at the two end-points served by the airline, the distance between the two end-points, its square and a dummy variable for tourist markets. The firm characteristics, $Z_{i,m}$ are some measures of airport presence, i.e. a dummy if the airline is operating from the two end-points the period before and a measure of market share at these two end-points. We consider similar characteristics in Section 5.

[6]With the normalization $h(1; \delta) = 0$ when there is the constant term in X_m. Berry (1992) uses $h(N; \delta) = \delta \log N$ whereas Bresnahan and Reiss (1991a) estimate each value $h(2; \delta), h(3; \delta)$, *etc.* separately.

In this model of complete information, firm i decides to enter in a market m with N^* active firms if $\pi_{i,m}(N^*) \geq 0$, otherwise it does not enter. The action $y_{i,m}$ taken by firm i in market m is therefore $y_{i,m} = \mathbf{1}\{\pi_{i,m}(N^*) \geq 0\}$ and, obviously,

$$N^* = \sum_{i=1}^{N_p} y_{i,m}.$$

Consequently, for each market m, the actions $y_{i,m}$, $i = 1, \ldots, N_p$ are the solutions of the simultaneous equation system:

$$y_{i,m} = \mathbf{1}\{\pi(X_m, Z_{i,m}; \theta) - h(y_{1,m} + y_{2,m} + \ldots + y_{N_p,m}; \delta) + \varepsilon_{i,m} \geq 0\},\ i = 1, 2, \ldots, N_p. \quad (2)$$

Therefore, the entry game makes the market structure endogeneous. First, the simultaneous equations in (2) prevent us from estimating the parameters of the model with the econometric procedures designed for standard binary models because of the endogeneous variables on the right hand side of Equation (2).[7] Additionally, estimating a probit or a logit while omitting the actions $y_{i,m}$ on the right hand side implicitly makes the assumption that more competitors don't change each firm's profit, an assumption rejected in most of the cases. Consequently, the estimators would be inconsistent. Therefore it is crucial to characterize the solutions of this system of equations. This is the objective of the next part.

2.1.1 The problem of multiple equilibria

Assume for the sake of simplicity that there are only two potential entrants, 1 and 2, and that the equilibrium concept is the Nash equilibria in pure strategies. Adapting Equation (1), our entry game can be summarized as follows:

$$y_{1,m} = \mathbf{1}\{\beta_{1,m} - \delta y_{2,m} + \varepsilon_{1,m} \geq 0\}, \quad (3)$$

$$y_{2,m} = \mathbf{1}\{\beta_{2,m} - \delta y_{1,m} + \varepsilon_{2,m} \geq 0\}. \quad (4)$$

Here, with the notations above, $\beta_{i,m} = \pi(X_m, Z_{i,m}; \theta)$ for $i = 1, 2$. Figure 1 plots the different regions of interest.

In region I, the profit shocks are too negative and no firm is profitable even in monopoly, nobody enters, i.e., $y_{1,m} = y_{2,m} = 0$. In region IV, the shocks are very high and both firms can be profitable even with one competitor. Both enter, i.e., $y_{1,m} = y_{2,m} = 1$. In region II, we have two subcases (separated by the dashed line). In region II-A, firm 1 is not profitable even in monopoly and firm

[7] One could estimate this system using instrumental variables, but estimating binary models with endogeneous explanatory variables is notoriously more complicated than for the linear case. See Blundell and Powell (2003), for example.

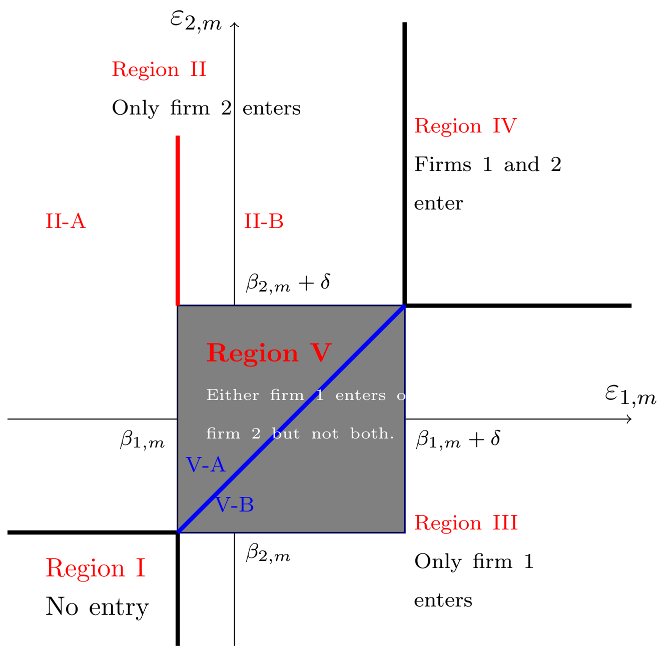

Figure 1: The multiple equilibria problem with 2 potential entrants

2 is profitable. Consequently, $y_{2,m} = 1$ and $y_{1,m} = 0$. In region II-B, firm 1 is profitable only in monopoly, whereas firm 2 can sustain a competitor. Consequently, firm 2 enters and firm 1, knowing that, does not. Consequently, $y_{2,m} = 1$ and $y_{1,m} = 0$. By symmetry, in region III, $y_{1,m} = 1$ and $y_{2,m} = 0$. Finally, in region V, each firm is profitable in monopoly but is not in duopoly. Without further information or another equilibrium concept, the model can not predict the outcome. Either $y_{1,m} = 1$ and $y_{2,m} = 0$ or $y_{2,m} = 1$ and $y_{1,m} = 0$. This entry game with heterogeneous agents is a game with multiple equilibria.

The multiple equilibria problem prevents the econometrician from estimating the model with the usual techniques because there is no one-to-one mapping between the regions and the four possible outcomes. Region V predicts either $(1, 0)$ (i.e. $y_{1,m} = 1$ and $y_{2,m} = 0$) or $(0, 1)$. The model is said incomplete because we don't know what the mechanism of outcome selection in region V is.[8] There are different solutions to tackle this problem.

2.1.2 Solutions to the problem of multiple equilibria

First, one can add some assumptions to complete the model, like in standard problem of missing information in economic models. Bajari et al. (2010) assume large support for the profit shocks, leading to point identification and the possibility to estimate the equilibrium selection mechanism

[8] See also Heckman (1978).

of this model.[9] One can also make assumptions about the order of entry of the different firms. Bjorn and Vuong (1985) assume that the order is purely random whereas Berry (1992) proposes one estimator where the most profitable firm enters first. For the latter, Region V-A now predicts the outcome $(0, 1)$ whereas Region V-B predicts $(1, 0)$.

Consequently, the probability to observe the outcome $(1, 0)$ is the sum of the areas of the Regions III and V-B with respect to the true distribution of $(\varepsilon_{1,m}, \varepsilon_{2,m})$ and, we can calculate similarly the probability of observing the outcome $(0, 1)$. This solution is ad-hoc, of course, and suffers from the risk of misspecification, i.e., adding a wrong assumption that may bias the estimators.

Another solution consists in using the recent literature on moment inequalities and set identification (see Tamer, 2009 or Bontemps and Magnac, 2017, for a survey). It is out of the scope of this chapter to introduce formally this econometric solution. The idea is to bound the probability of each outcome.[10] Ciliberto and Tamer (2009) are the first to apply this literature on entry games. For example, in Figure1, the probability of observing the outcome $(1, 0)$ is bounded by the probability that the shocks belong to region III and the probability that they belong to region III or V. Similarly, the probability of observing the outcome $(0, 1)$ is bounded by the probability that the shocks belong to region II and the probability that they belong to region II or V. Using such bounds allows to exploit all the information provided by the structure of the model and leads to more precise estimates of the parameters of interest. However, generalizing this to more players is cumbersome and requires more sophisticated inference techniques.

Bresnahan and Reiss (1991b) observe that, in the case above, despite the presence of a multiple equilibria region, the number of firms entering is always the same. The inference can therefore be conducted from the number of firms rather than the set of actions. Here, the probability of observing 0 firm entering is the area of region I, 2 firms entering is the area of region IV, and 1 is the sum of the areas of region II, III and V. Nevertheless, this simplification is at the cost of having less precise estimators because of some informational loss (the probability to observe the outcome $(1, 0)$ is at least greater than the area of region III, an information no longer taken into account in the new strategy). But it allows to estimate the model by standard techniques, here maximum likelihood, as the one-to-one mapping between regions of shocks and number of firms is restored.

[9] See, also Tamer (2003).

[10] Moment inequality procedures are now well developed in the partial identification literature, which includes contributions by, among others, Chernozhukov et al. (2007), Andrews and Soares (2010) and Beresteanu et al. (2011). It becomes complicated and numerically challenging for entry games with more than 5 players. See Bontemps and Kumar(2018) who propose a computational solution for games with more players.

2.1.3 A simulated estimator

In his seminal paper, Berry (1992) generalizes the last example to the entry games with more than two potential entrants. The number of multiple equilibria regions is much higher than with two potential entrants. It increases exponentially (8 regions with 3 entrants, 81 regions with 4, 2008 regions for 5 entrants, etc. Bontemps and Kumar (2018) provide, in particular, an exact counting). Berry (1992) proves that, nevertheless, the number of active firms is constant in these regions.

Therefore, like for the case with two entrants, one can estimate the entry game from the number of entrants in each market. We assume that we observe M independent markets, $m = 1, \ldots, M$ in which the N_p firms compete. The characteristics $X_m, Z_{1,m}, \ldots, Z_{N_p,m}$ and the profit shocks $\varepsilon_{1,m}, \ldots, \varepsilon_{N_p,m}$ are assumed to be identically distributed and independent across markets[11] Our goal is to estimate the parameters θ of the profit function, δ of the function $h(N;\delta)$ and, the correlation matrix of the within market profit shocks, $\varepsilon_{1,m}, \ldots, \varepsilon_{N_p,m}$ (in its simplest form, one can assume, like in Berry, that $\varepsilon_{i,m} = \rho u_{0,m} + \sqrt{1-\rho^2} u_{i,m}$, where all the shocks $u_{j,m}$ are independent and standard normally distributed ; $u_{0,m}$ is the market shock common to all firms). Let us denote α the vector gathering all parameters in θ, δ and ρ. If the number of active firms is constant despite the presence of multiple equilibria region, one can estimate α by maximum likelihood:

$$\hat{\alpha}_{MLE} = \arg\max_{\alpha} \left(L(\alpha) = \sum_{m=1}^{M} \log P(N_m^*; \alpha) \right).$$

If one estimates a probit model without common market shock (assuming $\delta = 0$ and $\rho = 0$), we have a closed form for $P(N_m^*; \alpha)$. This assumption is generally rejected bu the data. The homogeneous version of Bresnahan and Reiss (1991a) is also a particular example of the model for $\rho = 1$ and homogeneous firms (i.e., no $Z_{i,m}$) in which we obtain a closed form for the likelihood (it's an ordered model, see Bresbahan and Reiss (1991a) for further details). Generally, due to the huge number of regions involved and their difficulty to be characterized explicitly, it is not possible to get such an explicit form. However, one can use simulated methods to estimate (across simulations) $P(N^*; \alpha)$ for each α using the following algorithm proposed by Berry (1992):

Step 1: Choose the number of simulations S (at least 500).

Step 2: For each s in $1, \ldots, S$, and for each m, draw iid sequences $u_{0,m}^{(s)}$, $u_{1,m}^{(s)}, \ldots, u_{N_p,m}^{(s)}$ from the standard normal distribution. This sequence is now fixed.

Step 3: Choose a value for the vector of parameters α. The next three steps estimate $P(N_m^*; \alpha)$.

[11] There may be correlation within markets for the profit shocks like in Berry (1992). Jia (2008) is the first one to depart from the iid assumptions. She analyzes competition between K-Mart and Wal-Mart with network effects. Her model is applicable for games with two players only.

Step 4: For each market m and each s, compute $\Phi_{i,m}^{(s)} = \pi(X_m, Z_{i,m}; \theta) + \rho u_m^{(s)} + \sqrt{1-\rho^2} u_{i,m}^{(s)}$, for $i = 1, \ldots, N_p$.

Step 5: For each market m and each s, calculate from these $\Phi_{i,m}^{(s)}$'s, the number of active firms at the equilibrium in market m for simulation s, $N_m^{*,(s)}$, by counting how much firms are profitable in a market with 1, 2, 3, ... active firms.

Step 6: For each market m, calculate the empirical frequency, $\widehat{P(N_m^*; \alpha)}$, of observing the outcome $N = N_m^*$ across simulations:

$$\widehat{P(N_m^*; \alpha)} = \frac{1}{S} \sum_{s=1}^{S} \mathbf{1}\{N_m^* = N_m^{*,(s)}\}.$$

The previous procedure estimates $P(N_m^*; \alpha)$ with $\widehat{P(N_m^*; \alpha)}$. One can now estimate α from observing market characteristics and decision of firms to enter or not by maximizing the Simulated Maximum Likelihood (SML):

$$L_{sim}(\theta) = \sum_{m=1}^{M} \log \widehat{P(N_m^*; \alpha)}.$$

The SML estimator is consistent but asymptotically biased. When S is sufficiently large, this bias can be neglected. The standard errors can be computed using the classical variance formula, assuming S is large enough to neglect the simulation impact.

2.1.4 The results of Berry (1992)

Berry estimates the decision to enter for airlines in markets whose end-points are two of the 50 largest cities of the United States.[12] He considers the year 1980. The simulation method proposed above is slightly modified by adding in Step 4, an order of entry. It is therefore possible to estimate on top of the number of active firms, the probability of entering for each firm in each market.

One particular empirical finding is the importance of airport presence at both end-points to explain the decision to operate between two end-points and a numerical evaluation of this impact. This result can be put in perspective, for example, with the empirical findings of Goolsbee and Syverson (2008), where it is shown that the threat to entry (i.e. being present at both end-points) would put a pressure on the active firms even if the airline does not serve the corresponding market.

[12] There is an opened debate about what is the right market, i.e. city pairs or airport pairs and whether the markets are directional or not. Each different assumption can be handled within this framework.

Also, the larger the market size, the higher the probability to enter. This effect is quantified and one can estimate which markets are under exploited. Besides, based on estimated parameters of the model, it is also possible to simulate the number of entrants in different market configurations.

Furthermore, Berry's model encompassing the homogeneous model of Bresnahan and Reiss (1991a) and also the probit model where the competitor's presence does not harm each firm profit ($\delta = 0$), these restrictions can be tested empirically. They are rejected by the data.

2.2 Extensions

The key assumption which allows to solve the problem of multiple equilibria in Berry (1992) is to assume that the impact of the competitors on each firm profit does not depend on their identity (the function $h(N;\delta)$ does depend only on N in Equation (1) and does not vary when we switch one active firm and one non entrant). This assumption may seem extreme in the case of the airline sector where at leat two type of firms are present, Full Service carriers and Low-cost carriers (even 3 in Europe with the regional version of the Full Service carriers). However, relaxing this assumption is difficult because of the problem of multiple equilibria. This problem is recurrent in this literature.

Cleeren et al. (2010) assume that there are two types of firms who decide whether to enter or not in a version of the game where firms within each type are exchangeable. Their model is applied on competition between supermarket chains with traditional supermarket chains and discounters. Multiplicity of equilibria is dealt with by adding assumptions about the nature of competition between types and by selecting the equilibria with the highest number of traditional supermarket chains, which are argued to have entered the market first. Mazzeo (2002) proposes an equivalent version in which the format (or the quality) is also endogeneized. He considers two or three quality levels. Again, the multiple equilibria problem is solved by assuming a sequence of quality choice. Both models can therefore be estimated by maximum likelihood.

Bajari et al. (2010) goes one step further into estimation of complete information games with equilibrium selection, by, instead of picking one selection rule, explicitly modeling and estimating the probability that each one of several equilibrium selection possibilities will be played on a game.

A number of studies (Mazzeo, 2002; Jia, 2008; Cleeren et al., 2010; Sampaio, 2011) that specified a sequential order of movement to solve for the multiple equilibria also investigated the robustness of their results to alternative order of movements, and found fairly similar coefficients to profits functions under the alternative specifications.

Finally, Ciliberto and Tamer (2009) introduce much more heterogeneity on firms' profit functions. They also apply their model to the US domestic airline industry and allow each variable (market and firm observed characteristics) and competition effect to potentially differ among airlines. Not surprisingly, such flexibility comes at a cost of more general multiple equilibria structure. They estimate their model using moment inequalities.

3 Entry games of incomplete information

The previous section have analysed entry models under the assumption that part of airlines' profits are unobserved by the econometrician, but are common information detained by all airlines that could potentially enter the market.

We now investigate how to recover airlines' profits assuming that information about a specific airline profitability is privately held by that airline, thus not being shared by their competitors as well as the econometrician.

With incomplete information, the equilibrium concept used is the Bayes-Nash equilibrium, in which airlines base their entry decisions on expected profits. The models presented here were developed in the context of static entry models by Seim (2006) and Bajari et al (2010). Also, we report the reader to Bajari et al. (2013) for a survey of the econometrics with game interactions. In this section, we first look into an entry game with only two players. Then, we extend this entry game to multiple competitors.

3.1 The case of 2 firms

Let $y_{i,m}$ denote a dummy variable representing the entry decision of airline i in market m. Profits of not entering a market are also normalized to zero. The profit function of an active firm is the same than in the complete case, i.e., in Equation (3) or (4). Firm i gets the following profit if it enters :

$$\pi_{i,m} = \pi_i(X_m, Z_{i,m}; \theta) - \delta y_{-i,m} + \varepsilon_{i,m}, \tag{5}$$

where $-i$ denotes the rival's index ($-i = 1$ if $i = 2$ and vice versa). In the usual linear setting, like the one considered above, $\pi_i(X_m, Z_{i,m}; \theta) = X_m^\top \beta + Z_{i,m}^\top \gamma$ but, again, any parametric specification is possible. Like above, δ represents the decrease in airline i profits caused by the actual entry of airline $-i$. Now, $\varepsilon_{i,m}$, the profit shock, is unknown to both airline $-i$ and the econometrician. It is assumed that $\varepsilon_{i,m}$ has an extreme value distribution to guarantee some analytical tractability to the computation of the equilibrium of the game. This distribution is common knowledge of the players and the econometrician.[13]

[13]Up to some parameters for the econometrician, eventually.

Then, since airline 1 does not observe airline 2's profit shock, but knows its distribution, its entry strategy is based on its own expected profits on market m:

$$\begin{aligned} y_{1,m} = 1 &\iff \mathbb{E}\left[\pi_1\left(X_m, Z_{1,m}; \theta\right) - \delta y_{2,m} + \varepsilon_{1,m} \middle| \text{ Information of firm 1}\right] \geq 0 \\ &\iff \varepsilon_{1,m} \geq -\pi_1\left(X_m, Z_{i,m}; \theta\right) + \delta \mathbb{E}\left[y_{2,m}\right]. \end{aligned}$$

In the derivation above, the only uncertainty comes from the non observation of $\varepsilon_{2,m}$ and its impact on the expected value of $y_{2,m}$. Since $y_{2,m}$ is a dummy variable, $\mathbb{E}\left[y_{2,m}\right]$ corresponds to the probability that airline 2 actually enter the market, $Prob(y_{2,m} = 1)$. Let $\sigma_{2,m}$ represent this probability.

The probability that airline 1 enter market m, also known as airline 1 best response function, depends on the probability that airline 2 will also enter market m. With the extreme value distribution for $\varepsilon_{1,m}$:

$$\begin{aligned} \sigma_{1,m} = Prob\left(y_{1,m} = 1\right) &= Prob\left(\varepsilon_{1,m} \geq -\left(\pi_1\left(X_m, Z_{1,m}; \theta\right) - \delta \sigma_{2,m}\right)\right) \\ &= \frac{\exp\left(\pi_1\left(X_m, Z_{1,m}; \theta\right) - \delta \sigma_{2,m}\right)}{1 + \exp\left(\pi_1\left(X_m, Z_{1,m}; \theta\right) - \delta \sigma_{2,m}\right)}. \end{aligned} \tag{6}$$

Similarly, airline 2 best response function is given by:

$$\sigma_{2,m} = Prob\left(y_{2,m} = 1\right) = \frac{\exp\left(\pi_2\left(X_m, Z_{2,m}; \theta\right) - \delta \sigma_{1,m}\right)}{1 + \exp\left(\pi_2\left(X_m, Z_{2,m}; \theta\right) - \delta \sigma_{1,m}\right)}. \tag{7}$$

The equilibrium of the game, $(\sigma^*_{1,m}, \sigma^*_{2,m})$, corresponds to the solution of equations (6) and (7) for each market m. The problem of multiple equilibria present in the entry games with complete information still arises here. To solve this problem, we have to assume some form of uniqueness of equilibrium across markets or over time. As discussed in Pesendorfer and Schmidt-Dengler (2003), if we have data on the same market in multiple periods of time, it is easier to support the assumption of uniqueness of equilibria. However, in airlines entry models, we typically gather information of multiple origin-destination markets. In this case, we need the same equilibria $\left(\sigma^*_{1,m}, \sigma^*_{2,m}\right)$ to emerge in markets having the same market share and the same airlines characteristics, $(X_m, Z_{1,m}, Z_{2,m})$.

Following the uniqueness assumption, Seim (2006) and Bajari et al. (2010), propose a two-step method of estimation of the parameters of the model, which is easy to implement using standard commands of statistical softwares. In the first step, the entry probability of each airline, σ_1 and σ_2, is estimated as depending on all available information on market and airlines' characteristics, $(X_m, Z_{1,m}, Z_{2,m})$, but ignoring the interactions between players decisions.[14] Estimation can be

[14]Tamer (2003) also proposes it for the complete model with two players.

performed using a parametric model, such as a logit model, or a nonparametric model, such as a local linear model.

In the second step, interactions between airlines decision are brought back, by plugging the estimates of $\widehat{\sigma}_1(X_m, Z_{1,m}, Z_{2,m})$ and $\widehat{\sigma}_2(X_m, Z_{1,m}, Z_{2,m})$ obtained from the first step in equations (6) and (7), so that:

$$\sigma_1 = \frac{\exp\left(\pi_1\left(X_m, Z_{1,m}; \theta\right) - \delta\widehat{\sigma}_2(X_m, Z_{1,m}, Z_{2,m})\right)}{1 + \exp\left(\pi_1\left(X_m, Z_{1,m}; \theta\right) - \delta\widehat{\sigma}_2(X_m, Z_{1,m}, Z_{2,m})\right)}$$

and

$$\sigma_2 = \frac{\exp\left(\pi_2\left(X_m, Z_{2,m}; \theta\right) - \delta\widehat{\sigma}_1(X_m, Z_{1,m}, Z_{2,m})\right)}{1 + \exp\left(\pi_2\left(X_m, Z_{2,m}; \theta\right) - \delta\widehat{\sigma}_1(X_m, Z_{1,m}, Z_{2,m})\right)}.$$

Finally, the parameters of interest (θ, δ) can be recovered with the estimation of the following pseudo log-likelihood:

$$L\left(\theta, \delta\right) = \sum_{m=1}^{M}\sum_{i=1}^{2} y_{i,m} \log\left(\frac{\exp\left(\pi_i\left(X_m, Z_{i,m}; \theta\right) - \delta\widehat{\sigma}_{-i}(X_m, Z_{1,m}, Z_{2,m})\right)}{1 + \exp\left(\pi_i\left(X_m, Z_{i,m}; \theta\right) - \delta\widehat{\sigma}_{-i}(X_m, Z_{1,m}, Z_{2,m})\right)}\right),$$

which is actually just a conditional logit likelihood estimation.

3.2 The case of multiple firms

The last entry model can be very easily extended to multiple players. Suppose now that there are N_p potential players in each market.

In a simple model with symmetric competition effects of different rival airlines, if airline i decides to enter in market m, it collects profits given by:

$$\pi_i\left(X_m, Z_{i,m}; \theta\right) - \delta \sum_{j \neq i} y_{jm} + \varepsilon_{i,m},$$

where, once again, $\varepsilon_{i,m}$ is private information to airline i and is assumed to have an extreme value distribution. Profits of not entering a market are also normalized to zero. The decision to enter a market is based on expected profits, which alongside the extreme value distribution, yields the following best response function for airline i:

$$\sigma_{i,m} = \frac{\exp\left(\pi_i\left(X_m, Z_{1,m}; \theta\right) - \delta \sum_{j \neq i} \sigma_{jm}\right)}{1 + \exp\left(\pi_i\left(X_m, Z_{1,m}; \theta\right) - \delta \sum_{j \neq i} \sigma_{jm})\right)}, \ \forall\ i = 1, \ldots, N_p.$$

The Bayes-Nash equilibrium is now implicitly defined by a system with N_p unknowns $(\sigma^*_{i,m},\ i = 1, \ldots, N_p)$ and N_p equations.

The multiple equilibria problem still arises and is solved by assuming that there exists a unique equilibrium in the same market over time or in a cross-section of markets, like in the case with two players. The estimation of the model is done likewise in two steps.

In a first step, $\sigma_{i,m}$ is estimated based on market and all airlines characteristics $(X_m, Z_{1,m}, \ldots, Z_{Nm})$ in this market. Given this first step estimates, $\widehat{\sigma}_{i,m} = \widehat{\sigma}_{i,m}(X_m, Z_{1,m}, \ldots, Z_{N_p,m})$, the parameters of interest can once again be estimated by maximum likelihood on the log-likelihood:

$$L(\theta, \delta) = \sum_{m=1}^{M} \sum_{i=1}^{N_p} y_{i,m} \log \left(\frac{\exp\left(\pi_i\left(X_m, Z_{i,m}; \theta\right) - \delta \sum_{j \neq i} \widehat{\sigma}_{jm}\right)}{1 + \exp\left(\pi_i\left(X_m, Z_{i,m}; \theta\right) - \delta \sum_{j \neq i} \widehat{\sigma}_{jm}\right)} \right). \tag{8}$$

Note that the model above can easily accommodate the case where an airline might decide between K options. The unobserved profit shocks now are action specific, that it, each possible action is associated to a different error term, and it is unknown to their competitors. As before, competitors know the distribution of these error terms, but not the specific realization that occurred in a market. The solution to this problem follows the same two steps of the simple entry game.

3.3 Discussion and extensions

This section has presented a simple version of an entry game with incomplete information, with the main goal of illustrating the traditional assumptions made in the literature. As discussed, the biggest advantage of the private framework is that the estimation of game parameters are much simpler than in complete information set up, since it can be easily performed on available software commands at low time cost, without any simulations being necessary. This simplification is also driven by the fact that it is assumed explicitly that the same strategy is played across observations when there may be multiple equilibria. This is a strong assumption and deriving bounds like in the complete case would be a solution that could be investigated more systematically though it is not the case yet.

One argument that has been put forward in the literature in favor of the complete information assumption is the following: since static entry models are supposed to represent firm's long run decisions, there should be no ex-post regret, which is a property of Nash equilibria. In games with incomplete information, since decision are made ex-ante, before knowing the real realization of their competitors unobserved profits, airlines may be surprised by the entry (or lack of entry) of their competitors and end up with ex-post realized profits that do not correspond to their most profitable action.

Einav (2010) incorporate sequential moves into entry games with incomplete information to study competition in movies release dates. This model derives a unique perfect Bayesian equilibrium, even for quite heterogeneous profit functions. Conditional on a fixed order of moves, estimation can be easily done with maximum likelihood. The author argues that, with sequential choices, there is less ex-post regret, since firms that move later on have more information on which to base their decisions.

There is no consensus in the literature on what is the most appropriate assumption about the degree of information on the unobserved profitability shared by competitors on entry models of airlines. To the best of our knowledge, no unifying framework has been developed.

The set-up of games with private information also provided a nice framework to dynamic considerations on entry models. In the dynamic setting, ex-post regret is no longer an issue, since airlines are allowed to freely enter and exit markets as long as they pay some sunk costs. Aguirregabiria and Mira (2007) and Bajari et al. (2013) offer excellent surveys of dynamic structural econometric models with more in depth details about estimation, inference and efficiency. We review some of the empirical papers related to the airline sector in the next section.

4 Entry games applied to the airline sector

In this section, we present some of the "structural" empirical entry games papers applied to the airline industry in the recent years. There is a huge and vast literature of reduced form estimations (see, for example, Bogulaski et al., 2004, or de Oliveira, 2008)).

4.1 Complete Information

Sampaio (2011) uses Cleeren et al. (2010)'s model to study competition between low-cost and full-service airlines. The model is estimated with the assumption that full-service carriers enter the market first and then low-cost carriers make their decisions. Results suggest the existence of strong competition effects on this industry. Entry of same type rivals has a large significant effect on profits. Besides, entry of low-cost carriers seems to affect more profits of full-service airlines than vice-versa.

Dunn (2008) also introduces product quality to entry models. He studies competition between airlines offering non-stop and one-stop routes to investigate cannibalization and business stealing competition effects. One-stop routes passing through hubs are taken as fixed, and entry with non-stop services is modeled with a game theoretical model à la Berry (1992). However, the inclusion of a cannibalization effect in the model implies multiple equilibria in the total number of firms with non-stop operations in a market, and, to deal with this problem, the author selects the equilibria with the highest number of entrants. Their results suggest both the presence of strong competition effects as well as relevant cannibalization effects.

On the same direction of Mazzeo (2002) and Cleeren et al. (2010), Blevins (2015) investigates entry models with complete information and sequential movement in the US domestic industry. Now, the order of moves is not observed (or assumed) by the econometrician. Instead it is estimated alongside the coefficients of the profit function. In contrast to Mazzeo (2002) and Cleeren et al. (2010), and similar to Berry (1992), the model focus on pure entry, without product differentiation, but it allows firm observed heterogeneity.

Ciliberto and Tamer (2009) introduce much more heterogeneity on firms' profit functions. They also apply their model to the US domestic airline industry and allow each variable (market and firm observed characteristics) and competition effect to potentially differ among airlines. The complicated multiple equilibria problem (there is no obvious outcome invariant in the multiple equilibria regions) is tackled by using moment inequalities. They found evidence of a greater impact of an entry of Southwest on the three major airlines' profits compared to the impact of the entry of another major airline.

4.2 Incomplete Information

An important contribution of dynamic entry models to the airline industry is the ability to analyse network effects of entry decisions and to rationalize hub-and-spokes networks. Static models usually incorporate network effects via a hub dummy variable or market presence. In contrast, Aguirregabiria and Ho (2010) and Aguirregabiria and Ho (2012) create the concept of local managers that maximize profits of several interconnected routes inside an airline's network, so that the airlines recognize that entering in one route affect profits on other routes. This also creates scope for entry deterrence effects, since airlines might enter a route with negative profits, as long as they are compensated by positive profits on the rest of the network. The entry threat is credible to potential most efficient competitors in one specific route, since no profits loss is actually incurred by the network airline at each period of time. In the before mentioned papers, they document relevant network effects on the sunk entry cost as well as large entry deterrence effects.

Benkard et al. (2018) advocates the need of introducing dynamic considerations on merger analysis to understand medium and long run effects of a proposed merger, since potential negative short run effects of a merger implied by an increase in concentration might be offset by the entry of other potential competitors. They use a simplified version of Bajari et al. (2007) estimation method to investigate the effects of three proposed mergers in the airlines industry between 2000 and 2010. Interestingly, their long run analysis would motivate a competition authority conduct at odds with what was actually adopted in practice. Their model is however much less flexible than Aguirregabria and Ho (2012) and does not allow for entry deterrence effects. It thus remains the question of how entry deterrence effects might affect medium and long run competition after mergers.

Finally, Gillen et al. (2015) analyse the relationship between network and regional airlines in a game with two stages. In the first stage, network airlines decide the number of regional airlines they are going to sign contracts with. In a second stage, they chose if they are also going to operate the market with their own fleet. Their model is an extension of the classical incomplete information entry game where, instead of modeling a 0/1 decision, one models the choice of an integer $n_{i,m}$, here, the number of regional airlines airline i is contracting with.

Except for the inclusion of terms related to the number of contracts signed by all network airlines with regional airlines, the best response function of network airlines can be estimated following the procedure described in the previous section. Their results for both stages corroborate the existence of negative network airlines competition effects on each other in both stages ($\delta > 0$). They also indicate a positive correlation between its own number of contracts with regional airlines and its decision to also operate in a market.

5 Example on European Data

We now illustrate the approach by estimating a static entry game with complete information on western European data for the year 2015. The data were collected from the Official Airline Guide database (OAG) which contains all posted flights for the year 2015, their schedules and eventually, their code share agreements.

A market is defined as an Origin/Destination flight (non directional) between two European cities. We therefore group the different airports of a given city (London, Paris).[15] We restrict our investigations to the decision to operate between the 50 largest cities[16] of Western Europe. The distance between these cities and the socioeconomic variables such as GDP and population of the metropolitan areas are collected from additional statistical sources.[17] We eliminate markets for which the distance between the cities is less than 150 kilometers. 1204 potential markets remain.

Contrary to the US market, the western European market has more players. Each country has its own flag carrier which plays usually a significant role in the industry[18]. We group the different airlines as follows. The first fifteen airlines in number of routes, see Table 1, are kept as they are. All the other airlines are gathered according to their alliance membership (if any), otherwise according

[15]We also grouped Alicante and Valencia, Malaga and Sevilla.

[16]See Table 8 for the list of cities

[17]Eurostat.

[18]British Airways, Lufthansa, AirFrance, KLM and Alitalia are among the top 15 airlines counting for 85% of the traffic

to their type.[19]

5.1 Descriptive statistics

In 2015, 900 million passengers were transported. The most frequented routes were Paris-Toulouse and Barcelona-Madrid with respectively 3.5 and 2.5 millions of passengers. Out of the possible 1204 markets, we observe 670 existing connections. Table 1 and 2 present some descriptive statistics of the final dataset.

Most of the growth in markets served from 2014 to 2015 is due to an increased activity of Low-cost carriers who opened 95 markets out of 148 new markets. They serve around 41% of the routes with a share in seats offered of approximately 29.4%. On average, there are 2 airlines per market served. Less than 6% of the markets have more than 4 competitors.

Following Berry (1992), we consider the following explanatory variables for the profit function:

- *Pop* is the geometric mean of the population of the metropolitan areas of the two cities of the market in millions.
- *Gdppercap* is the mean GDP per capita of the metropolitan areas of the two cities of the market in thousands of euros per capita.
- *Dist* is the distance in thousand of kilometers, *Dist2* is its square.
- *Sun* is a dummy equal to 1 if one of the end points is on the Mediterranean Sea.
- *LC* is a dummy for Low Cost airlines
- *City2* is a dummy if the airline operates on the two cities out of the considered market.
- *Nbroutes* is the mean number of cities served by the airline out of the two endpoints.

Table 3 displays some descriptive statistics of these variables. Table 4 presents the results of the regression of the number of firms on the market characteristics. All variables but the dummy *Sun* are significant.

[19] We have the first fifteen airlines Ryanair, Easyjet, Vueling, Lufthansa, Germanwings, Norwegian, SAS, Air Berlin, British Airways, Air France, TAP, Alitalia, Aer Lingus, Iberia and KLM. We have therefore 5 additional fictive airlines, StarAlliance2 which gathers Brussels Airlines, Aegean Airlines, Austrian airlines mainly, Skyteam2 which gathers Transavia, Garuda, Air Europa, Hop and Czech Airlines, Oneworld2 which gathers Finnair, Niki, Main_others which gathers all classic airlines, and LC2 which gathers all the other Low Cost airlines.

5.2 Results of the estimation and comments

We use a Simulated Maximum Likelihood method to estimate the static entry game with complete information. The profit function has the form of Equation (1), more specifically,

$$\pi_{i,m}(N) = \beta^\top X_m + \gamma^\top Z_{i,m} - \delta \log N + \underbrace{\rho u_m + \sqrt{1-\rho^2} u_{i,m}}_{=\varepsilon_{i,m}},$$

where the market variables X^m are *Pop*, *Gdppercap*, *Dist*, *Dist2* and *Sun* whereas the firm specific variables are *LC*, *City2* and *Nbroutes*. The results are displayed in Table 5. We also put the results of a probit estimation without the interaction term (i.e. assuming $\delta = 0$) and assuming $\rho = 0$. The results are totally different as expected. δ, the interaction parameter is significantly positive. Assuming (wrongly) that the value is equal to zero leads to biased estimates.

The percentage of good predictions[20] is equal to 78.1%.

Like in Berry (1992), we get the expected signs. When the size of the cities grows, it is more likely to serve the corresponding market. The effect of distance is first increasing than decreasing from 1357 kms. It captures the fact that people fly more to distant cities but less when it starts to be too far.

More importantly, the coefficient of City2 is positive and very important. Airport presence is a good predictor of the decision to enter in a market. The more important is the presence, the more likely the airline will enter. The interaction coefficient δ is positive as expected. Finally ρ is positive and captures the fact that profits shocks within market are correlated across firms.

We can now use these estimates to evaluate in which markets, not served in 2015, airlines are more likely to enter by checking in which markets the discrepancy is the highest. We report in Table 6 the top ten markets distant from at least 500 kms. Among these 10 markets, 7 are actually served in 2019. Alternatively, we report in Table 7 the top ten markets for which the model predicts too much entry. We report the number of competitors for the year 2019. 7 markets out of these ten markets has experienced a decrease in the number of competitors. The model can obviously be better adapted to the European context by taking into account the fragmentation of countries and their relation to their flag carriers. More specific characteristics could be tested. This is left for additional research. The goal of this section is to show what kind of results we can expect and how they can be used to predict some market structure change in markets between top european cities.

[20] We set that the prediction is right when the true number of competitors is either equal to the integer part of our mean predicted number or to the next integer.

6 Conclusion

In this review, we present the static entry games with complete or incomplete information. These models are useful to estimate market structure from data that are relatively easy to obtain. The problem of multiple equilibria has been exposed. It complicates the estimation procedure. Despite the fact that it occurs in both information cases, complete or incomplete, it has been treated more seriously in the games with complete information.

The recent literature on moment inequality models could propose a unified treatment of the two cases. By a revealed preference argument a choice is made because any other one would have been less optimal. These inequalities are however difficult to handle in a game with a medium number of players. Nevertheless this is an interesting way to work on.

Most of the empirical work on entry games, applied to airlines, have considered the US market. The data are easy to collect, available online and the market has only a few large players. We hope that the online access to flights proposed in other continents would allow researchers to study more systematically the market structure of the airline sector in Europe and Asia.

References

Aguirregabiria, V., and Mira, P. (2010), "Dynamic discrete choice structural models: A survey". *Journal of Econometrics*, **156(1)**, 38-67.

Aguirregabiria, V., and Ho, C. Y. (2010), "A dynamic game of airline network competition: Hub-and-spoke networks and entry deterrence", *International Journal of Industrial Organization*, **28(4)**, 377-382.

Aguirregabiria, V., and Ho, C. Y. (2012), "A dynamic oligopoly game of the US airline industry: Estimation and policy experiments", *Journal of Econometrics*, **168(1)**, 156-173.

Andrews, D.W.K., and G. Soares (2010), "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection", *Econometrica*, **78**, 119-157.

Andrews, D.W.K., and X. Shi (2013), "Inference Based on Conditional Moment Inequalities", *Econometrica*, **81**, 609-666.

Bajari, P., Benkard, C. L., and Levin, J. (2007), "Estimating dynamic models of imperfect competition", *Econometrica*, **75(5)**, 1331-1370.

Bajari, P., H. Hong, and S. Ryan (2010), "Identification and Estimation of Discrete Games of Complete Information", *Econometrica*, **78**, 1529-68.

Bajari, P, H. Hong, and D. Nekipelov, (2013), "Game Theory and Econometrics: A Survey of Some Recent Research", In D. Acemoglu, M. Arellano, & E. Dekel (Eds.), *Advances in Economics and Econometrics: Tenth World Congress (Econometric Society Monographs)*, 3-52.

Benkard, C. L., Bodoh-Creed, A., and Lazarev, J. (2018), "Simulating the dynamic effects of horizontal mergers: Us airlines", Working Paper, Standford University.

Beresteanu, A., I. Molchanov and F. Molinari (2011), "Sharp Identification Regions in Models with Convex Moment Predictions", *Econometrica*, **79**, 1785-1821.

Berry, S. (1992), "Estimation of a Model of Entry in the Airline Industry", *Econometrica*, **60**, 889-917.

Berry, S.,and P. Reiss (2007), "Empirical Models of Entry and Market Structure", in M. Armstrong and R. Porter, Editor(s), *Handbook of Industrial Organization*, Elsevier, Volume 3.

Berry, S., and E. Tamer (2007), "Identification in Models of Oligopoly Entry", in *Advances in Economics and Econometrics: Theory and Applications, Ninth World*

Congress, vol. 2, R. Blundell, W.K. Newey and T. Persson, eds., Cambridge Univ. Press.

Bjorn, P., and Q. Vuong (1984), "Simultaneous equations models for dummy endogenous variables: a game theoretic formulation with an application to labor force participation", Working Paper, California Institute of Technology.

Blundell, R. and J. L. Powell (2003), "Endogeneity in Nonparametric and Semiparametric Regression Models", In: M. Dewatripont, L. P. Hansen, and S. J. Turnovsky (eds.), *Advances in Economics and Econometrics*, Eighth World Congress, vol. 2, Cambridge University Press.

Boguslaski, C., Ito, H. and D. Lee (2004), "Entry Patterns in the Southwest Airlines Route System", *Review of Industrial Organisation*, **25**, 317350.

Bontemps, C. and T. Magnac (2017), "Set Identification, Moment Restrictions and Inference", *Annu. Rev. Econ.*, **9**, 103-179.

Bontemps, C. and R. Kumar (2018), "A Geometric Approach to Inference in Set-Identified Entry Games", Working Paper 18-943, Toulouse School of Economics.

Bresnahan, T. and P. Reiss (1991a), "Entry and Competition in Concentrated Markets", *Journal of Political Economy*, **99(5)**, 977-1009.

Bresnahan, Timothy F., and P. Reiss (1991,b), "Empirical models of discrete games", *J. Econometrics*, **48(1-2)**, 57-81.

Ciliberto, F., and E. Tamer (2009), "Market Structure and Multiple Equilibria in Airline Markets", *Econometrica*, **77**, 1791-1828.

Chernozhukov, V., H. Hong, and E. Tamer (2007), "Inference on Parameter Sets in Econometric Models", *Econometrica*, **75**, 1243-1284.

Cleeren, K., Verboven, F., Dekimpe, M. G., and Gielens, K. (2010), "Intra-and interformat competition among discounters and supermarkets", *Marketing science*, **29(3)**, 456-473.

Dunn, A. (2008), "Do low-quality products affect high-quality entry? Multiproduct firms and nonstop entry in airline markets", *International Journal of Industrial Organization*, **26(5)**, 1074-1089.

Einav, L. (2010), "Not all rivals look alike: Estimating an equilibrium model of the release date timing game", *Economic Inquiry*, **48(2)**, 369-390.

Grieco, P.L.E. (2014), "Discrete games with flexible information structures: an application to local grocery markets", *RAND Journal of Economics*, **45**, 303-340.

Heckman, J. J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System", *Econometrica*, **46**, 931959.

Jia, P.(2008), "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry", *Econometrica*, **76(6)**, 1263-1316.

Manski, C. (1995), *Identification problems in the social sciences*, Harvard University Press.

Mazzeo, M.J. (2002), "Product Choice and Oligopoly Market Structure", *Rand Journal of Economics*, **33(2)**, 221-242.

de Oliveira, A.V.M. (2008), " An empirical model of low-cost carrier entry", *Transportation Research Part A Policy and Practice*, **42**, 673-695.

de Paula, A. (2013, "Econometric Analysis of Games with Multiple Equilibria", *Annual Review of Economics*, **5**, 107-131.

Pesendorfer, M., and Schmidt-Dengler, P. (2008), "Asymptotic least squares estimators for dynamic games", *The Review of Economic Studies*, **75(3)**, 901-928.

Reiss, P. (1996), "Empirical Models of Discrete Strategic Choices", *American Economic Review*, **86**, 421-426.

Sampaio, R. M. B. (2011), "Competition Between Low-Cost Carriers and Traditional Airlines: An Empirical entry Model", Working Paper, Toulouse School of Economics.

Seim, K. (2006), "An Empirical Model of Firm Entry with Endogenous Product-type Choices", *Rand Journal of Economics*, 37, 619-640.

Sutton, J. (1991), *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*, MIT Press.

Tamer, E. (2003), "Incomplete simultaneous discrete response model with multiple equilibria", *Rev. Econ. Stud.*, **70(1)**, 147-165.

Tamer, E. (2010), "Partial Identification in Econometrics", *Annu. Rev. Econ.*, **2**, 167-95.

List of Tables

Airline	Type	Nb Routes Served 2015	2014	Tot_seat	Share cumul.
Ryanair	L	195	173	36036441	11.5%
Easyjet	L	158	139	31244528	9.9%
Vueling	M	88	76	15875178	5.1%
Norwegian	L	73	68	11164021	3.6%
LH	M	71	71	39218702	12.5%
G-wings	L	69	58	9592148	3.1%
SK	M	64	61	16155919	5.1%
AB	M	47	46	12324708	3.9%
BA	M	39	38	22270455	7.1%
Air France	M	38	38	22086929	7.0%
TP	M	38	38	8253202	2.6%
Alitalia	M	36	29	10334054	3.3%
EI	M	36	32	7594926	2.4%
KLM	M	34	34	13495804	4.3%
IB	M	34	30	11890206	3.8%
Main_Others	M	87	71	7875822	2.5%
StarAlliance2	M	76	71	16017272	5.1%
SkyTeam2	M	70	61	10841036	3.5%
LC2	L	58	32	4381525	1.4%
OneWorld2	M	33	30	7461873	2.4%

Type= M for mainline and L for Low-cost.

Table 1: Number of markets served between the top 50 cities of Western Europe.

N 2015	Freq.	Percent	Cum.
0	534	44.35	44.35
1	268	22.26	66.61
2	223	18.52	85.13
3	108	8.970	94.10
4	55	4.570	98.67
5	11	0.910	99.58
6	4	0.330	99.92
7	1	0.0800	100
Total	1,204	100	

Table 2: Number of competitors per market.

VARIABLES	N	mean	sd	min	max
N_2015	1,204	1.116	1.283	0	7
Pop	1,204	2.4	1.2	1.1	12.8
Gdppercap	1,204	38.8	10.5	14.4	72.7
Dist	1,204	1,167	617	156	3,362
Sun	1,204	0.370	0.483	0	1
City2	1,204	0.212	0.409	0	1
Nbroutes	1,204	4.66	8.48	0	75.5

Table 3: Explanatory variables.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-1.3590	0.1782	-7.63	0.0000
Pop	0.5782	0.0244	23.65	0.0000
Gdppercap	0.0203	0.0031	6.50	0.0000
Dist	0.3987	0.1753	2.27	0.0231
Dist2	-0.1052	0.0621	-1.69	0.0903
Sun	0.0597	0.1654	0.36	0.7181

Table 4: Regression of N on the market characteristics

	Estimate (Std. Error)	Probit
(Intercept)	-2.616 (0.044)	-3.163 (0.123)
Pop	0.090 (0.009)	0.060 (0.013)
Gdppercap	-0.0003 (0.0001)	-0.004 (0.002)
Dist	0.342 (0.033)	0.560 (0.119)
Dist2	-0.126 (0.012)	-0.178 (0.042)
Sun	0.0002 (0.002)	-0.068 (0.101)
LC	-0.121 (0.058)	-0.001 (0.040)
City2	1.518 (0.056)	0.783 (0.040)
Nbroutes	0.047 (0.002)	0.076 (0.002)
δ	0.274 (0.030)	
ρ	0.306 (0.025)	

Table 5: Simulated Maximum Likelihood estimates.

N_{2015}	Predicted	Diff	Market	Distance	2019 (airline)
0	1.58	1.58	Brussel-Glasgow	808	$\surd$ (Ryanair)
0	1.7	1.7	Bordeaux-Manchester	955	$\surd$ (Ryanair& Easyjet)
0	1.86	1.86	Berlin-Marseille	1177	$\surd$ (Easyjet)
0	1.74	1.74	Lille-Roma	1200	No airline
0	1.68	1.68	Birmingham-Stockholm	1360	$\surd$ (SAS)
0	1.68	1.68	Glasgow-Milan	1504	No airline
0	1.9	1.9	Birmingham-Lisbon	1623	No airline
0	1.54	1.54	Athens-Toulouse	1996	$\surd$ (Aegean Airlines)
0	2.48	2.48	Athens-Valencia	2119	$\surd$ (Aegean Airlines)
0	2.12	2.12	Athens-Malaga/Sevilla	2615	$\surd$ (Aegean Airlines)

Table 6: Market where entry is predicted.

N_{2015}	N_{model}	Diff	Market	Distance	N_{2019}
5	1.24	-3.76	Roma-Vienna	777	3
5	1.68	-3.32	Paris-Palermo	1471	2
6	2.74	-3.26	Porto-Paris	1224	6
4	0.96	-3.04	Copenhagen-Dublin	1239	3
7	4.04	-2.96	Lisbon-Paris	1454.5	5
5	2.3	-2.7	Copenhagen-Roma	1535.4	3
4	1.48	-2.52	Barcelona-Birmingham	1273	2
4	1.5	-2.5	Glasgow-Sevilla	1966.5	4
4	1.5	-2.5	Marseille-Roma	608	3
4	1.54	-2.46	Madrid-Porto	436	4

Table 7: Market with too many entries.

Amsterdam, Antwerp, Athens, Barcelona, Berlin, Birmingham, Bilbao, Bordeaux, Bristol, Brussels, Copenhagen, Cardiff, Dresden, Dublin, Duesseldorf, Frankfurt, Glasgow, Goteborg, Hannover, Hamburg, Helsinki, LeedsBradford, Lille, Lisbon, London, Liverpool, Lyon, Madrid, Manchester, Milan, Marseille, Munich, Naples, Newcastle, Nottingham, Nantes, Porto, Oslo, Paris, Palermo, Roma, Rotterdam, Thessaloniki, Stockholm, Stuttgart, Sevilla, Torino, Toulouse, Valencia, Vienna.

Table 8: List of cities selected for the analysis