Applying Cleeren et al. (2008) to the US airline sector

Benjamin Clouet[*]
Toulouse School of Economics

Bruce Ravey[†]
Toulouse School of Economics

April 16, 2023

1 Introduction

The airline industry in the United States has gone through various subsequent changes after deregulation of the industry in 1978. While the large legacy carriers (American Airlines, Eastern Airlines, TWA and United Airlines) continued their expansion, a large number of new smaller airlines entered the market in 1980's. The importance of cost optimization and profit maximization increased during that period, while the increased overall competition helped in reducing the average fare. During that period, airlines seeking economies of scale to reduce their cost of operation inforced on enlarging their networks, operating on an increasing number of routes and optimizing the occupation of their aircrafts. Through the end of the 1980's, the economic recession hit airlines that were too ambitious following the liberalization of the industry. A subsequent number of firms were forced to file for bankruptcy, not being able to support the costs that rose in the attempt of acquiring large market shares and to expand on as many routes as possible.

At the end of the 1990's and in the beginning of the 2000's, the US airline industry was greatly challenged by large decreases in the demand for air travel. The economic downturn in the end of the 1990's made airlines suffer from a sharp decrease in passenger demand, while the unexpected shock of 9/11 changed the perception of security that surrounded this travel mode. The demand for business air travel also went under structural changes, as fares were continuously rising with no compensation in terms of additional services provided, greater amount of security checks were applied following 9/11 which consumed additional time in the timetables of business customers, and the creation of other modes of communication which did not require business passengers to travel as much as before. The inability of legacy carriers to assess those changes in the demand patterns for air travel and to adapt their yield management strategies greatly hurt their financial situation

This modification of the demand for air travel in the beginning of the 2000's gave rise to a change in the market structure of the airline industry. Four of the six major carriers at that time had to file for Chapter 11 bankruptcy. Although these budgetary constraints were not considered to be indefinite, those legacy carriers were forced financially to take a step back and re-organize their production schemes to sustain their activity, mainly cutting on labour-related costs. This allowed smaller carriers to gain great shares of the market as their own.

While the low cost carriers only represented ten percent of the air traffic in the US in 1990, this market share rose to about 25% in 2005. Using a more flexible organization of labour, where a large share of their employees are cross-utilized across various tasks, they were able to increase greatly the productivity of their workers. Low cost carriers were also pioneers in their utilization of new technologies, integrating the use of internet for the booking of the tickets, making customers loyal through fidelity clubs organized around emails, adding other air-related services such as car rentals or hotel booking to their full-package of supply.

[*]Available through benjamin.clouet@ut-capitole.fr.
[†]Available through bruce.ravey@ut-capitole.fr.

Helped greatly by fierce competition in the airline manufacturing industry, low cost carriers were able to buy entire fleets of medium-haul aircrafts for extremely low prices. Their quality of service was deemed high by the public, while maintenance costs were kept low thanks to the acquisition of brand new models of airplanes. Low cost carriers also operated with a very different organizational structure as the one used by legacy carriers, namely the point to point organization of aircraft flows. Legacy carriers, when they profited off large market shares in the 1980's, first decided to implement a hub and spoke policy, centering the flow of aircrafts to large airports and thus optimizing their utilization, at the expense of investing large sunk costs to keep such operations afloat. The revolution of low cost carriers was to depart from such an organizational management, and to rather focus on point-to-point connections, where the low cost carriers would cherry-pick the most profitable lines across the US and exploit them, rather than concentrating all their routes towards specific hubs. Although the discrimination of activities between hub-and-spoke for legacy carriers and point-to-point for low cost carriers is not as extreme as a binary situation, and some legacy carriers operate point-to-point on some routes as low cost carriers also created small structures similar to hubs, it still helps explain greatly the organizational efficiencies low cost carriers brought to the airline market in the US.

In this paper, we will study the competitive effects the legacy carriers (FSC) and low-cost carriers have on each other. Using the framework of analysis of Cleeren et al. (2008), we try to estimate how the intra-format and inter-format competition influences airlines' profits and we provide elements of analysis on the relationship between market size and competition.

2 Data Analysis

Our sample of data is an extract of a randomly chosen number of airplane tickets chosen in the second quarter of 2017. We transform the data as follows ; by grouping the extracted tickets by departure and arrival cities we compose markets from the original data. For each market, we know how many legacy and low-cost carriers operate and we also observe some other control variables such as the mean of the population between the two cities or the distance between the two cities. Our final database is composed of 3983 markets.

We provide in the annex two summary tables of our data.

Insert Table 1 here

Insert Table 2 here

The first table describes the counts of market structure we observe in our data. On the rows is the count of legacy carriers, and on the columns the counts of low-cost carrier. The second summary table describes the explanatory variables we will use in our estimation strategy, as well as their mean and dispersion.

3 Econometrics analysis

In this section, we detail how our model of entry for the US airline sector is specified, the different assumptions used to help simplify the optimization procedures, and we provide interpretation material for the parameters estimated.

The estimation strategy we use here follows the methodology used in Cleeren et al. (2008). It is a model of entry with two types of formats, or quality of service, with which the authors study the impact of intraformat and interformat competition on firm's profits. In the original paper, it is through the lense of the German's supermarket sector that this analysis is conducted, with "regular" supermarkets and discounters competing against each other to ideally extract positive profits. The core argument of the paper is that competition within formats does not have the same impact on firm's profits as competition between formats, and that the strength of the within and between competition differs for the different formats as well.

In this section, we will study those nuanced competition effects between and within formats in the US airline sector. The two types of formats firms belong to are the "Full Service Carriers" (FSC), which are considered to be the high quality type, and the "Low Cost Carriers" (LCC) will be our low-quality type.

In a first step of the estimation method, we define the profit functions of the firms studied in the following way. We study a latent variable $\boldsymbol{\pi}_f(\boldsymbol{N}_f, \boldsymbol{N}_{-f}) = \pi_f(\boldsymbol{N}_f, \boldsymbol{N}_{-f}) - \epsilon_f$, which represents the profits for a firm of format f. More particularly, we have :

$$\pi_f(\boldsymbol{N}_f, \boldsymbol{N}_{-f}) = \alpha_f ln(marketsize) + \beta_f X_m - \delta_{f,f} log(\boldsymbol{N_f})\mathbf{1}_{N_f>0} - \delta_{f,-f} log(1 + \boldsymbol{N}_{-f})/\boldsymbol{N}_f$$

The index f is for a given format, and $-f$ relates to the other format. In this functional form, we make the profits of a given firm vary with the $marketsize$ and some other market characteristics X_m, namely the distance between the two cities considered and its square. The intra-format competitive effects will be estimated by the parameter $\delta_{f,f}$, and we allow for the inter-format competition to be of a different magnitude with the parameter $\delta_{f,-f}$. It is important to note here that, for a given equilibrium and market situation, the payoff functions are the same for firms of the same type.

In this framework, an equilibrium $(\boldsymbol{N}_{FSC}, \boldsymbol{N}_{LCC})$ arises in a given market depending on the value of the unobserved firm characteristics ϵ_{FSC} and ϵ_{LCC}. Using the definition of our latent variable previously defined, the equilibrium $(\boldsymbol{N}_{FSC}, \boldsymbol{N}_{LCC})$ arises as a Nash equilibrium if :

$$\pi_f(\boldsymbol{N}_{f+1}, \boldsymbol{N}_{-f}) < \epsilon_f \leq \pi_f(\boldsymbol{N}_f, \boldsymbol{N}_{-f}), forf = FSC, LCC$$

It will thus be the characterization of the distribution of those format specific error terms that will allow us to derive the probabilities for the different equilibria that arise in our framework of analysis. We will consider that the error terms individually follow a normal distribution, and thus when grouped together they follow a joint bivariate normal distribution. To solve for the parameters of interest of the model, we will use a Maximum Likelihood Estimation approach. The parameters chosen are the ones that maximize the joint probability of our sample arising given the chosen distribution of the error terms.

We write the probability of a given market situation $(\boldsymbol{N}_{FSC}, \boldsymbol{N}_{LCC})$ arising as Nash Equilibrium the same way as in Cleeren et al. (2008) :

$$P(N_S = n_S, N_D = n_D) = \int_{\pi_S(n_S+1, n_D)}^{\pi_S(n_S, n_D)} \int_{\pi_D(n_S, n_D+1)}^{\pi_D(n_S, n_D)} \phi(u_S, u_D) du_S du_D$$

$$- \int_{\pi_S(n_S+1, n_D)}^{\pi_S(n_S+1, n_D-1)} \int_{\pi_D(n_S+1, n_D)}^{\pi_D(n_S, n_D)} \phi(u_S, u_D) du_S du_D,$$

In their specification, the indexes S and D would refer respectively to FSC and LCC in our case. The first term of this expression is purely the probability of $(\boldsymbol{N}_{FSC}, \boldsymbol{N}_{LCC})$ arising as Nash Equilibrium, and the second term represents the substraction of the overlapping regions of other equilibria situations that arise for the same set of values of unobservables. For a given equilibrium situation determined by the values of ϵ_{FSC} and ϵ_{LCC}, Cleeren et al. (2008) make various assumptions to characterize and restrict the set of other equilibria that could satisfy the Nash criterion as well. Following this methodology, they conclude that substracting the area of overlap of the $(\boldsymbol{N}_{FSC-1}, \boldsymbol{N}_{LCC+1})$ and $(\boldsymbol{N}_{FSC+1}, \boldsymbol{N}_{LCC-1})$ equilibria is sufficient to get rid of the multiplicity problematic. They state that all other $(\boldsymbol{N}_{FSC-m}, \boldsymbol{N}_{LCC+m})$ and $(\boldsymbol{N}_{FSC+j}, \boldsymbol{N}_{LCC-j})$ where $m, j > 1$, are subsets of the $(\boldsymbol{N}_{FSC-1}, \boldsymbol{N}_{LCC+1})$ and $(\boldsymbol{N}_{FSC+1}, \boldsymbol{N}_{LCC-1})$ area. Substracting the area of $(\boldsymbol{N}_{FSC-1}, \boldsymbol{N}_{LCC+1})$ and $(\boldsymbol{N}_{FSC+1}, \boldsymbol{N}_{LCC-1})$ thus gets rid of the situation of the multiplicity of equilibria. We follow in our econometric analysis the same assumptions and approach.

To write our log-likelihood function, we simply take the sum of the log transformation of these probabilities for our sample of 3983 markets. We specify the likelihood function of a given observation depending on the situation of entry in the market, namely that the individual likelihood function takes alternative functional forms for cases where $(\boldsymbol{N}_{FSC} > 0, \boldsymbol{N}_{LCC} > 0)$, or $(0, \boldsymbol{N}_{LCC} > 0)$, etc.

In order to facilitate estimation, we force some inequality restrictions on the parameters of interest.

1. $\delta_{f,f} \geq 0$, $\delta_{f,-f} \geq 0$ for $f = FSC, LCC$
2. $\delta f, f \geq \delta_{f,-f}$ for $f = FSC, LCC$

These assumptions help ensure that the competition effects on the profits of firms of another entry is negative no matter the format of the entrant, as the δ parameters enter with negative sign in our specification of firms' profits. Also, we make sure that the intra-format competition is at least as hard as the inter-format competition. This helps in restricting the space of optimization and allowed us to obtain meaningful results for our parameters. In terms of the optimization procedure, we use the package *nloptr* on Rstudio to obtain our estimates. The optimization algorithm used is the $COBYLA$ one, and we fix some stopping criterions on the tolerance level and the maximum number of iterations of the optimization procedure. For the starting values of the parameters to include in the optimization algorithm, we ran two separate ordered probit models on the two dependent variables $\boldsymbol{N}_{FSC}$ and $\boldsymbol{N}_{LCC}$ using as regressors the controls of the model, namely $marketsize$, $distance$ and $distance^2$. For the starting values of the competitive effects, we scrabbled about and tried different sets of values until the algorithm would provide final results.

We now turn to our estimation results, displayed in Table 2. The standard errors reported are computed inversing the hessian of the log-likelihood function at the final parameter values.

[Insert Table 2 about here]

As in Cleeren et al. (2008), we find that for both formats the impact of intra-format competition on profits is of a larger magnitude compared to the inter-format one. The distance has a positive impact on the profits for both types while the square of the distance has an inverse impact, although both these coefficients have a quite neglibeable magnitude. We also observe that the impact of competition on profits in both intra an inter-format cases, reported in the Table 3, increases with the number of entrants on the market, which is a result in line as well with the literature. The profits of firms of both formats decreases with the size of the market, which is a result that is quite counter-intuitive after taking away the effect of competition.

[Insert Table 3 about here]

4 Thresholds, Counterfactual and Robustness Checks

In this final section, we want to compute interesting quantities from the results of the non-linear optimization of our sample log-likelihood function.

First, we begin by computing for both types the intra and inter-format entry thresholds. These quantities allow us to infer on how market size has to adjust in the case of an additional entry. The intra-format ratios compute how much market size has to change following the entry of a same type competitor, while the inter-format ratio computes the same quantity for an entry of a competitor of a different type.

[Insert Table 4 about here]

We find a surprising result for the intra-threshold ratios of the FSC type. They are always below one, which means that an additional entry of a FSC type in a market with no LCC competitor would require the market size to decrease for the incumbent to make profits. This counter intuitive result might be put in perspective with the fact that we found in our estimation results that market size was negatively correlated with the profits of both types. For the LCC types, we find an intra threshold ratio that is always superior to one, but which does not seem to increase much with the number of competitors. That is, the increase in the market size for an additional entrant to break even in profits when there is already four other LCC competitors (and no FSC competitor) is the same in the proportion as the increase of market size it would require for an additional entrant to break even in profit in a market with say seven LCC competitors already established.

[Insert Table 4 about here]

In terms of the inter-format threshold ratios, we observe some more explainable results for the FSC type. These ratios are always above one, which means that the additional entry of a competitor of a different type requires a higher market size so that it can sustain positive profits. As the largest value for these inter-format threshold ratios is highest for the first entry of a LCC competitor, it seems that it is just the presence of another LCC carrier (and not their number) that seems to hurt more FSC carriers in our sample. For the LCC carriers, these inter-format ratios have a similar interpretation. The entry of the first FSC carrier seems to be the most damaging to LCC carriers, as this is the situation where the market size need to be the most increased to sustain positive profits. Then, any additional entry of new FSC competitors still needs an increase in the market size to be sustained, but less and less compared to the first entry of a FSC-type competitor.

We continue this section with the analysis of a counterfactual estimation situation. Since our estimated parameter for the logarithm of the market size meant that an increase in this variable would decrease the profits of the airlines, no matter the format, we want to evaluate if a different evaluation of the market size would lead to similar results. We thus proceed again with the non-linear optimization of our sample log-likelihood function now using another measure named *popmean*, which is computed as the mean of the logarithm of the population in millions of the departure and arrival cities. We want to infer on the robustness of our $\delta_{f,f}$ and $\delta_{f,-f}$ estimates as well as understanding why the population seemed negatively correlated with the profits in our analysis.

[Insert Table 5 about here]

We find that our estimated intra-format and inter-format parameters do not change much in this counterfactual situation. For both formats, the intra competition is still stronger compared to the inter-format analog, and we keep a similar magnitude for those estimates in comparison to our initial results. The distance and the square of the distance have very similar effects as well as those computed earlier. A deceiving result is that in this estimation scenario, population is still significant but it's still negatively correlated with the profits of the airlines in our sample. The change of functional form for this variable does not seem to alter this relationship.

Finally, as a robustness check, we wanted to expand our analysis to the framework where the low-cost carriers would be considered as the ones entering a market first rather than the usual case where the FSC carriers enter first. To do so, we thought that changing the order of integration in the probability functional form for an equilibrium $(\boldsymbol{N}_{FSC}, \boldsymbol{N}_{LCC})$ arising would be sufficient to change the order of entry in this scenario. We tried many different ways to estimate conveniently this change of the order of entry, but our optimization program was not able to produce any results. We are thus unable to present results for this robustness check.

5 Bibliography

Ben-Yosef, Eldad. "The Evolution of the US Airline Industr Olution of the US Airline Industry: Technology Echnology, Entry, and Market Structur Et Structure - Thr e - Three Revolutions Olutions ." Journal of Air Law and Commerce 72, no. 2 (n.d.).

Global Airline Industry Program. Accessed April 16, 2023.
https://web.mit.edu/airlines/analysis/analysis$_a irline_i ndustry.html$.

Cleeren, Kathleen and Dekimpe, Marnik and Gielens, Katrijn and Verboven, Frank, Intra- and Inter-Format Competition Among Discounters and Supermarkets (September 2008). CEPR Discussion Paper No. DP6964, Available at SSRN: https://ssrn.com/abstract=1308035

6 Annex

Counts of market structures

	0	1	2	3	4	5	6	7	8	9	10	11
0	0	396	107	50	17	4	7	0	0	0	0	0
1	365	250	144	69	38	12	5	2	1	1	0	0
2	287	167	132	80	40	17	4	2	1	0	0	0
3	140	90	77	70	37	19	9	2	2	0	0	0
4	59	71	44	38	36	16	13	1	1	1	0	0
5	38	41	50	33	33	18	10	10	2	1	0	0
6	26	35	41	33	22	23	14	13	4	2	0	0
7	21	26	34	20	20	15	13	2	7	5	0	0
8	14	25	26	29	23	20	15	10	4	2	1	1
9	7	16	20	25	21	17	7	8	3	1	1	0
10	7	14	10	8	5	10	4	6	1	1	1	0
11	1	8	7	6	6	3	2	2	1	1	2	1
12	1	1	3	4	1	6	0	2	2	1	1	0
13	0	0	1	3	2	5	1	2	1	1	0	0
14	0	0	0	0	1	0	2	0	0	0	1	0
15	0	0	0	0	0	0	0	1	0	0	0	0
16	0	0	0	0	0	0	0	1	0	0	0	0

Table 1: Descriptive Statistics

Variable	**Description**	**Mean**	**Standard_Deviation**
dist	Distance between the two airports	1220.62	692.27
dist_2	Squared distance between the two airports	1969053.26	1980049.97
marketsize	Root of the product of population of departure and arrival	3.46	2.02

Table 2: NL Optimization estimates

	Dependent variable:
cst_{Leg}	3.025^{***}
	(-0.000)
distance_{Leg}	0.001^{***}
	(-0.000)
distance^2_{Leg}	-0.00000^{***}
	(-0.000)
marketsize_{Leg}	-2.963^{***}
	(-0.000)
$\delta_{Leg,Leg}$	0.709^{***}
	(-0.000)
$\delta_{Leg,LCC}$	0.108^{***}
	(-0.000)
cst_{LCC}	1.153^{***}
	(-0.000)
distance_{LCC}	0.002^{***}
	(-0.000)
distance^2_{LCC}	-0.00000^{***}
	(-0.000)
marketsize_{LCC}	-0.756^{***}
	(-0.000)
$\delta_{LCC,Leg}$	0.146^{***}
	(-0.000)
$\delta_{LCC,LCC}$	1.167^{***}
	(-0.000)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 3: Competitive effects

	Dependent variable:	
	FSC	LCC
	(1)	(2)
$\delta^1_{f,f}$	0.000	0.000
$\delta^2_{f,f}$	0.491	0.809
$\delta^3_{f,f}$	0.779	1.282
$\delta^4_{f,f}$	0.982	1.617
$\delta^5_{f,f}$	1.141	1.878
$\delta^6_{f,f}$	1.270	2.090
$\delta^1_{f,-f}$	0.075	0.101
$\delta^2_{f,-f}$	0.118	0.160
$\delta^3_{f,-f}$	0.149	0.202
$\delta^4_{f,-f}$	0.173	0.234
$\delta^5_{f,-f}$	0.193	0.261
$\delta^6_{f,-f}$	0.210	0.283

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4: Threshold ratios

	Dependent variable:	
	FSC	LCC
	(1)	(2)
$INTRA^{1}$		
$INTRA^{2}$	0.588	1.008
$INTRA^{3}$	0.733	1.005
$INTRA^{4}$	0.802	1.003
$INTRA^{5}$	0.843	1.003
$INTRA^{6}$	0.870	1.002
$INTRA^{7}$	0.889	1.002
$INTRA^{8}$	0.903	1.002
$INTRA^{9}$	0.914	1.001
$INTRA^{10}$	0.922	1.001
$INTRA^{11}$	0.930	1.001
$INTER^{1}$	Inf.000	Inf.000
$INTER^{2}$	1.025	1.092
$INTER^{3}$	1.015	1.053
$INTER^{4}$	1.010	1.037
$INTER^{5}$	1.008	1.029
$INTER^{6}$	1.007	1.023
$INTER^{7}$	1.006	1.020
$INTER^{8}$	1.005	1.017
$INTER^{9}$	1.004	1.015
$INTER^{10}$	1.004	1.013
$INTER^{11}$	1.003	1.012

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 5: NL Optimization estimates with popmean

	Dependent variable:
cst_{Leg}	2.640^{***}
	(0.000)
distance_{Leg}	0.0002^{***}
	(0.000)
distance^2_{Leg}	-0.00000^{***}
	(0.000)
popmean_{Leg}	-1.285^{***}
	(0.000)
$\delta_{Leg,Leg}$	0.743^{***}
	(-0.000)
$\delta_{Leg,LCC}$	0.085^{***}
	(0.000)
cst_{LCC}	2.203^{***}
	(0.000)
distance_{LCC}	0.004^{***}
	(0.000)
distance^2_{LCC}	-0.00000^{***}
	(0.000)
popmean_{LCC}	-0.980^{***}
	(-0.000)
$\delta_{LCC,Leg}$	0.142^{***}
	(-0.000)
$\delta_{LCC,LCC}$	0.973^{***}
	(0.000)
Note:	*p<0.1; **p<0.05; ***p<0.01